

Proceedings of
2nd International Conference on
Smart Computing and Communication
for Sustainable Convergence
(ISCCSC 2025)



Edited by:

Dr. Shalli Rani, Chitkara University, Punjab, India

Dr. Ayush Dogra, Chitkara University, Punjab, India



River Publishers

Table of Contents

Chapter no.	Paper ID	Title
01	143	Performance Evaluation of AI Models for Multi-Class Brain Disease Prediction
02	203	VirtuAero: A CNN-LSTM and Random Forest- Driven Deep Learning Framework for Adaptive VR-Based Air Force Pilot Training
03	215	Plant Disease Detection and Pesticide Spraying Robot Using IoT
04	224	Footprint Recognition Using Deep Learning and Ensemble Machine Learning Models
05	268	Secret communication using steganography and cryptography
06	271	A Firefly-Driven Approach for Complex Background Resilient Sign Language Recognition
07	278	Deep Learning Neural Based System for Automated case Analysis in Homeopathic healthcare
08	279	End to End Deep Neural Networks for Automated Cardiac Arrest Detection and Risk Stratification
09	308	Crop Disease and Pest Detection Using Deep Learning and Transformer Models
10	317	
11	325	Deep learning Based Symptom Classification for Homeopathic Management of Respiratory Disorders
12	326	Fake News Detection: A Survey on Multivariate Feature Selection and Hybrid Deep Learning
13	327	AI-Based Public Safety Mobile application
14	345	IoT type Blynk App based Voice Pellet Medicine Dispenser
15	350	Human Queue Management System
16	351	Deep learning For Mapping Physiological Variability in Chronic Disease Patients using Homeopathic Care
17	352	CONSOLE LESS HAND GESTURE CONTROL DRONE
18	361	Crime Pattern Analysis and Prediction Using Machine Learning
19	363	IoT BASED UNDERGROUND CABLE FAULT DETECTION SYSTEM
20	379	Downlink Channel Prediction in FDD Massive MIMO Using Deep Transfer Learning Algorithms
21	389	Smart Walking Aid for Elderly Safety Using IoT Sensors and Real-Time Alert System
22	404	DISPLAY UNIT FOR COST ANALYSIS BETWEEN GRID AND SOLAR ENERGY
23	418	Autonomous Delivery Drones for LightWeight Payload Transpotation a Comprehensive Review
24	420	AI-Search for E-Learning Platforms

25	438	A Review On A Fault Detection System In Low Voltage Transmission Lines Using Relay
26	439	Automated Fruit Calorie Estimation Using 3D Image Reconstruction and Machine Learning
27	453	SmartEye: Self-Supervised AI System for Refractive Error Detection Using Fundus Images
28	133	Integration of Deep Learning and Edge-Cloud for Big Data Analytics in IOT
29	218	Facial Expression Recognition using Deep Learning Techniques
30	247	A Dual-Task Machine Learning Framework for Air Quality Index Prediction in India
31	284	A Hybrid CAE–MOCNN Architecture for Efficient Multiple-output Plant Species and Disease Classification
32	298	A Dual-Task Machine Learning Framework for Air Quality Index Prediction in India
33	304	AI Based Question Paper Generation
34	311	Automated Group Attendance System Using NFC and Facial Recognition for Enhanced Identity Verification
35	340	Eco-Farming Agriculture Method Using Artificial Intelligence Techniques: Applications and Challenges
36	342	A Survey on Explainable Artificial Intelligence(XAI) Techniques
37	359	Enhanced Comparative Study of Summarization Methods for Legal Assistants
38	378	AI Career Path Recommendation System
39	384	Integration of AI, Sensors, and Smart Devices for Monitoring Plastic Induced Health Hazards
40	397	A review on IoT Based Systems for Fault Detection in Power Line and Utility Poles
41	402	AI POWERED RETINAL SCREENING FOR GLUACOMA DETECTION USING DEEPLARNING TECHNIQUES
42	403	Smart Helmet Using Embedded Iot System
43	406	Explainable AI for E-Commerce: Interpretable Transformer Models for Negative Review Analysis and Feedback Summarization
44	407	Attention Enhanced Hybrid CNN – LSTM Framework for Realtime Mental Health Monitoring and Support
45	437	AI Based Legal Reference and Case Retrieval System
46	297	Design and Implementation of a Self-Sustaining Greenhouse using Internet of Things
47	353	Vehicle Prioritization and Modified GPSR for Congestion-Adaptive Routing in VANETs
48	354	A Robust Metaheuristic Approach for Parameter Extraction of Complex Solar Cell Models: Validation on Industrial PV Panels

49	355	An Enhanced Artificial Hummingbird Algorithm Outperforms Established Metaheuristics in DMFC Parameter Optimization
50	356	An IoT-Enabled Smart Cap for the Blind: Echo Vision with Obstacle Detection, Haptic Feedback & Emergency SOS
51	357	Cognitive Modeling of Smile Perception Using Reaction Time and PCA-Based Feature Selection
52	424	A REVIEW ON DETECTION OF GRASSHOPPER DISEASE IN SILKWORM
53	227	Predictive Modeling and Survival Analysis of Primary Biliary Cirrhosis: A Comparative Study of Machine Learning Approaches
54	228	Metaheuristic-Tuned Convolutional Networks for Plant Disease Recognition: Waterwheel Plant Algorithm as a Standout Optimizer
55	226	Earthquake Risk-Level Classification Using Machine Learning and Somersaulting Spider Optimizer
56	230	Comparative Analysis of Machine Learning Models for Rice Leaf Disease Classification: A Focus on Sensitivity, Specificity, and Accuracy
57	231	Skin Lesions Classification: A Comparative Analysis of SGD and SVC for Optimal Accuracy
58	232	AI-Biruni Earth Radius Algorithm-Based Hybrid Metaheuristic for Efficient Electric Vehicle Path Planning
59	234	Advancing Cytogenetic Diagnostics with PO-CNN: Robust Population-Based Optimization for Chromosome Classification
60	235	Metaheuristic Optimization in Wind Power Forecasting: A Comprehensive Literature Review
61	236	Ensemble and Tree-Based Machine-Learning Models for Customer Churn Prediction
62	229	An Enhanced Dipper-Throated Optimization Framework for Efficient Engineering Design Problem Solving
63	425	Advancing Skin Cancer Diagnosis: A Deep Learning Approach with EfficientNetB3
64	426	Thyroid Disease Identification via EfficientNetB2:Leveraging Deep Neural Networks for Early Detection
65	427	True vs Fake News Classification using Long Short Term Memory Networks
66	428	Knee Osteoporosis Diagnosis at an Early Stage Using a Fine-Tuned VGG19 Model on X-rays
67	429	Tire Condition Classification Using EfficientNetB0 for Improved Road Safety
68	431	Transfer Learning-Based CNN Models for Vehicle Image Classification
69	432	Deep Convolutional Feature Scaling with EfficientNetB3 for Robust Early Detection of Smoking Patterns in Health Informatics
70	433	Optimized EfficientNetB3 Architecture for High-Precision Multi-Class Weather Condition Classification
71	237	Deep Learning Approaches for Multiclass Plant Disease Detection: A Comparative Study of Custom and Pretrained Models

72	238	Optimizing Diabetes Prediction through Machine Learning and Feature Selection on Health Indicators
73	239	Metaheuristic Optimization in Space Science and Engineering: A Comprehensive Literature Review
74	240	BER-CNN: Bio-Inspired Evolutionary Refinement of CNNs for Rapid and Accurate Tuberculosis Detection
75	241	Optimizing Crop Recommendations Using Advanced Machine Learning Techniques
76	242	Metaheuristic Optimization in Monkeypox detection: A Comprehensive Literature Review
77	243	Multi-Layer Malware Defense—A Survey: From Metaheuristic-Driven Detection to Adversarially Robust Operations
78	244	Fast and Accurate Apple Scab Diagnosis via Dipper Throated Optimizer Enhanced CNNs
79	245	Earthquake Forecasting through Machine Learning: Predicting Magnitudes from USGS Seismic Data
80	246	Machine Learning-Based Diagnosis of Parkinson’s Disease Using Vocal Features: A Comparative Analysis of Model Efficacy and Sensitivity
81	321	Design and Implementation of an AI-Based Smart Traffic Framework

Performance Evaluation of AI Models for Multi-Class Brain Disease Prediction

Aakriti Khanna , Shubham Kumar , Saruchi Kukkar

Department of Computer Science and Engineering Chandigarh University , Department of Computer Science and Engineering, Chandigarh University , Department of Computer Science and Engineering, Chandigarh University
aakritikhanna1720@gmail.com , shuh5225@gmail.com, ganpati.saruchi@gmail.com

Abstract.

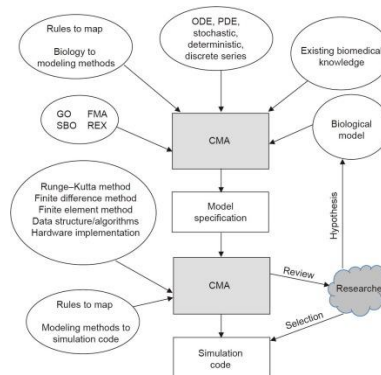
Machine learning and artificial intelligence are revolutionizing medical research, development and care. Open-source platforms, collaborative projects, and information exchange forums have democratized AI and ML, enabling healthcare innovation and knowledge sharing. The proposed work is divided into three parts that collect the data sets, train different models, and compare the results. We have taken five brain diseases dataset that include depression, brain stroke, Alzheimer, Parkinson disease, and Epilepsy. Also we have trained six different ML models that are random forest that is RF, support vector machine that is SVM , linear discriminant analysis that is LDA, NB that is Naive Bayes , Gradient Boost that is XGboost and logistic regression that is LR. The highest accuracy for logistic regression is 90% for parkinsons disease , 97% for XGboost model of depression , 97% for for random forest of epilepsy , 87% for LDA model of Parkinsons disease, 96% for naive bayes of epilepsy and 97% for SVM models of epilepsy disease. We have also used the heatmaps and bar graphs to show the comparison between them all.

Keywords. Artificial Intelligence, Machine Learning, LR model , Naive Bayes model, XGboost model , SVM model, LDA model , Random forest model , Brain diseases

1. INTRODUCTION

AI can improve healthcare efficiency, accessibility, and cost. Figure 1 shows how AI with proper regulation can improve equity, care, and society's access to new technology, cures, and pharmaceuticals. This ensures timely resource availability, reducing waste and improving care. AI may solve healthcare's biggest problems: cost, efficiency, and quality. AI-powered automation and optimization can cut costs and simplify patient scheduling, billing, and EHR management. AI changes pharmaceutical R&D, pharmacokinetics, assessment, production, marketing approval, and pharmacovigilance. AI assists patient stratification, digital twins, clinical trial simulations, and pharmacokinetic dose recommendations. Simplifying research, streamlining processes, and avoiding clinical trial failures saves time and money for safe and effective pharmaceuticals. Personalized medicine diagnoses, predicts, and analyzes patients using enormous healthcare data. Imaging, drug combinations, population-wide patient health outcomes, genetic or family data, medical papers natural language analysis [3]. Researchers and medics want to reduce worldwide illness mortality. Due to expanding healthcare data from varied and incompatible sources, predictive analytic models are becoming more relevant in medicine. General database storage cannot manage massive healthcare historical and streaming data. Medical diagnostics fix issues. Observational diagnosis is disease. Diseases have chemical and data anomalies. ML algorithms can learn categories or anticipate unknown circumstances from data[6]. Many old brain data-processing processes work. ML detects Alzheimer's, dementia, schizophrenia, MS, cancer, viral, and degenerative diseases in brain data. Sectioning and identifying sick tissue and brain structure. Diagnosis, surgery, postoperative examination, and chemo/radiotherapy require tissue identification. AI-powered virtual assistants and chatbots offer individualized health advice, prescription reminders, and real-time support. Despite progress, healthcare AI/ML integration is difficult. To improve healthcare using AI and ML, we must address ethics, data protection, regulatory compliance, and the digital divide. Fig. 1.

Artificial intelligence in medical field



2. LITERATURE REVIEW

The substantial healthcare literature on AI and ML reflects its many effects on patient care, clinical decision-making, and health system optimization. Early and present healthcare AI and ML studies demonstrate its benefits and drawbacks. Early 2000s studies showed that computer-aided diagnosis systems can diagnose numerous diseases, laying the framework for AI and ML medical imaging interpretation. ML method by Smith et al. detects CT lung nodules with 92% sensitivity and 96% specificity. AI-driven mammography interpretation for breast cancer diagnosis was proven by Lee et al.[4]. Faster, more accurate, and scalable AI diagnostics beat traditional approaches. A meta-analysis by Johnson et al. evaluated radiologists and AI systems for CT pulmonary angiography embolism detection. Deep learning enhances healthcare AI/ML. Deep learning algorithms, inspired by the brain, excel in photo identification and natural language processing, making them ideal for medical data analysis[1]. Deep learning models can improve brain and heart diagnosis and therapy utilizing MRI and CT. AIML and deep learning customize brain, heart health. AI-driven models improve treatment results and side effects using genetics, biomarkers, and clinical history. A personalized healthcare plan could personalize patient care[7].

3. Methodology

3.1. Dataset: Our study used Kaggle's Parkinson's, Alzheimer's, epilepsy, brain stroke, and depression datasets. Logistic Regression (LR), Gradient Boosting (GBoost), SVM, and Random Forest were created and tested using these datasets. Every dataset includes 5000 people's names, genders, and medical histories (heart disease, hypertension, and average glucose). All datasets had binary targets of (0) or (1). All data were.csv. Spreadsheets create CSV files using commas [2]

3.2.Data Preprocessing: Performing categorization quality, consistency, and sustainability preprocessing before model training and testing. BMIs missing were imputed using all entries' mean. One-hot encoded multi-class features like work type and smoking status, while label encoded categorical data. Numbers become 0 and 1. A lightweight, high-performing convolutional neural network, EfficientNet B0, extracted features to improve representation. All extracted features go to machine learning models following pretrained data extractor architecture. Mixed handmade and learning components increase model capture of complex patterns in the hybrid method.

3.3. Disease

3.3.1 Alzheimer's disease: Alzheimer's causes most dementia. Alzheimer's begins with amyloid plaques and neurofibrillary tangles[8]. Brain cell death lowers it gradually. Protein accumulation around brain cells may cause Alzheimer's. Amyloid surrounds brain cells. Tau protein causes brain tangles.

3.3.2. Brain Stroke: Stroke results from brain blood flow stoppage. In an emergency. Brain function requires continual oxygen and nourishment. Even a brief blood shortage can be problematic[4]. Two main factors induce stroke. A clogged brain artery causes ischemic stroke. Hemorrhagic strokes result from brain blood vessel ruptures. Some people experience brief ischemia attacks, which disrupt cerebral blood flow.

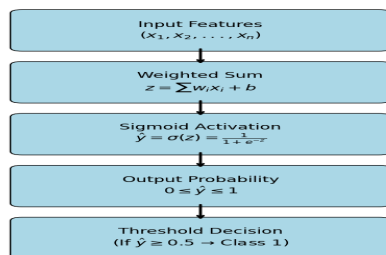
3.3.3.Depression: Depression is common and affects everyone. Loss of interest or long-term depression describe it. This differs from mood swings and daily feelings. No single reason causes depression. It has several causes and triggers.

3.3.4.Parkinson Disease: The neurological illness PD impacts movement, mental health, sleep, pain, and more. Over time, PD worsens. Therapies and drugs reduce symptoms but do not cure. Speech issues, tremors, and strong muscle contractions are prevalent. Brain chemical messenger neuron loss causes many Parkinson's symptoms. The messenger is dopamine. Brain activity changes with reduced dopamine. This causes Parkinson's mobility difficulties.

3.3.5.Epilepsy: Epilepsy affects all ages. Globally, 50 million have epilepsy. Epileptics are common in poor countries. Neurological epilepsy causes seizures[5]. Different epilepsy types exist. Some know why. Unknown causes exist elsewhere. Epilepsy is prevalent. The CDC reports 1.2% of Americans have active epilepsy. All sexes, races, ethnicities, and ages have epilepsy. The symptoms of seizures differ. Seizures knock you out. A few seizure victims stare blankly. Some jerk legs or arms.

3.4 Classifiers

3.4.1.Logistic regression (LR): Linear regression is the accurate correspondence between dependent and independent variables with few or more variables[6]. Since linear regression accurately predicts the change in dependent class variables to independent class variables. The algorithm of LR is shown in figure 2.



3.4.2.Gradient boost (GBoost): Gradient boosting (GBoost) creates powerful prediction models from weak learners like decision trees. Simple models predict first. Improved accuracy is achieved by training new trees each iteration to forecast the aggregate ensemble's residual errors.

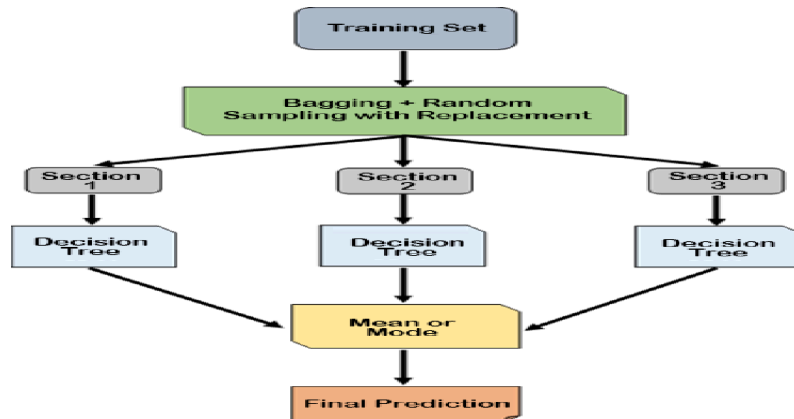
3.4.3.Naïve Bayes: Naive Bayes classifies statistically and supervisedly. Bayesian classifier. A value is assumed. Unrelated feature any other feature. Before probability, likelihood evaluated posterior probability[9]. The method for highest posterior probability varmeter estimate. The method is easy. enough training data to estimate parameters .; Required categorization. Training time and classification drops.

3.4.4. Linear Discriminant Analysis (LDA): A popular natural language processing generative probabilistic topic model is Linear Discriminant Analysis. LDA uses word distributions to identify corpus document latent themes. The "bag-of-words" assumption ignores document word order and emphasizes frequency and co-occurrence. LDAs are created by choosing a subject distribution, topic, and word from the topic's word distribution for each word.

3.4.5.Support vector machines (SVM): SVM classifies data points uniquely using an N-dimensional hyperplane. Supervised machine learning for classification, regression, outliers. Margin maximisation aids new-data model generalization. SVM handle linear and nonlinear data. For linearly separable data, model will find out the straight line

3.4.6. Random forest (RF)

Random forests use randomized data and several separate decision trees (DT). These trees grow throughout training, compile the output of the decision tree. Voting methods determine the prediction of the algorithm. This paradigm requires DTs to choose stroke sign indications or absences. Forecast comes out. Working Random Forest chooses the most-voted class as shown in figure 3.



3.5. Validation

3.5.1.Accuracy: Accuracy (the percentage of correctly identified examples among all predictions) can be misleading in imbalanced datasets.

Accuracy = Total positive value/Total Value

3.5.2.Precision: Precision measures positive case prediction accuracy.

Precision = True positive / Total positive

3.5.3.F1 Score: The harmonic mean of accuracy and recall knowns as the F1 score, adjusts for imbalanced classes or large false positives and negatives.

F1 score = $2 * PV * RV / PV + RV$, PV = Precision Value , RV = Recall Value

3.5.4. Recall : Recall (or sensitivity) measures how many positive cases the model correctly discovered.

Recall = Original real detection value / Original real detection value + False wrong detection value

3.5.5. Confusion Matrix : Verified real positives, verified real negatives, verified fake positives, and verified fake negatives are listed in the confusion matrix to detail prediction outcomes.

3.5.6 Receiver Operating Characteristic Curve : The ROC curve demonstrates how well the model detects classes by comparing recall and false real detection rates across the limiting value . More AUCs mean better performance.

True Positive Rate (TPR): Real detection / Real detection + Fake detection.

The fake real detection rate is the sum of the fake real detection and original fake detection. ROC Curve Algorithm (Binary Classification)

3.5 Results

3.5.1 Alzheimer's disease: We have trained six machine learning approaches using the Alzheimer dataset in the form of csv file and their performance indicators were presented through bar graphs, a comparison table and heatmaps. Gradient Boosting exhibited the highest performance metrics, obtaining a good accuracy value of 95.58%, with precision of 95.61%, recall score of 95.58%, and an F1 score of 95.59%. Random Forest classifier also performed well, with an accuracy of 93.49%. Moderate performance was observed with linear discriminant analysis (81.16%), logistic regression (81.63%), and support vector machine (83.02%). Although Naive Bayes recorded the lowest metrics, it still achieved a respectable accuracy of 76.98%, which is acceptable to some existing models. The results obtained after training all the models are shown in Table 1. The correlation and effectiveness of each model across the four evaluation metrics were effectively visualized using the heat map and the bar graph, which is shown in Figures 4 and 5.

Model	Accuracy	Precision	F1 Score	Recall
SVM	0.83	0.82	0.83	0.82
LR	0.81	8.81	0.81	0.81
LDA	0.81	0.81	0.81	0.81
RF	0.93	0.93	0.93	0.93

Naive bayes	0.76	0.77	0.76	0.77
Gradient boost	0.95	0.95	0.95	0.95

3.5.2. Brain Stroke: We enhanced performance measurements with brain stroke datasets and modified machine learning models. Gradient Boosting model outperformed others with high metrics and 90.51% accuracy during evaluation. Linear Discriminant Analysis and Random Forest models produced balanced results with 80.64% and 88.16% accuracy. The dataset's class imbalance hurt Naive Bayes. The heatmap, bar graphs, and comparison table below highlight these models' performance differences.

Model	Accuracy	Precision	Recall	F1 Score
LR	0.84	0.14	0.44	0.21
SVM	0.86	0.11	0.28	0.16
RF	0.88	0.14	0.28	0.18
Naive bayes	0.39	0.06	0.88	0.12
LD	0.85	0.19	0.44	0.26
Gradient boost	0.90	0.20	0.32	0.24

3.5.3. Depression: The depression dataset showed the best accuracy of 99% in logistic regression, suggesting faultless categorization. Gradient boost achieved amazing 97.5% accuracy, followed by SVM (96.25%) and Random Forest (95%). Improved accuracy of 92.05% is achieved with Naïve Bayes. The figure 08 shows heatmaps of all models , Table 3 shows comparison tables, and figure 09 shows bar graphs of all the six trained models results . These results are further used in comparison with all other diseases. The bar graph shows the comparison between all the trained models results. Further these results are compared with other four diseases result.

Model	Accuracy	Precision	Recall	F1 Score
LR	1.00	1.00	1.00	1.00
SVM	0.96	0.95	1.00	0.97
RF	0.95	0.94	1.00	0.97
Naive bayes	0.92	0.92	1.00	0.95
XGradient boost	0.97	0.97	1.00	0.98

3.5.4. Epilepsy: Random Forest , gradient-boosting and SVM perform well after model training. Random Forest achieved the highest performance metrics with 97% accuracy. SVM and Gradient Boosting obtained excellent accuracy of 97% and 96.65%. Naive Bayes was balanced, while LR and LDA were more precise. Machine learning algorithms had the top metrics. The accuracy , precision , F1 score and recall value obtained after training model is shown in Table 04. Along with this their heatmap is shown in figure 10 and figure 11 represents the bar graph of all models.

Model	Accuracy	Precision	Recall	F1 Score
LR	0.81	0.98	0.10	0.19
LDA	0.82	0.91	0.13	0.23
RF	0.97	0.95	0.94	0.94
Gradient boost	0.96	0.96	0.86	0.91
Naive bayes	0.96	0.90	0.91	0.91
SVM	0.97	0.96	0.92	0.94

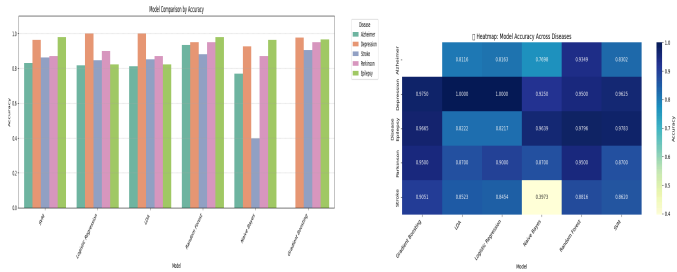
3.5.5. Parkinson Disease: The accuracy of Random Forest and XGboost performed is 95% each . LR with 90% accuracy was among many other successful models. Naive Bayes and LDA successfully improved accuracy to 87% apiece. The SVM accuracy was poor at 87%. The calculated results are shown in Table 05, 12 and 13.

Model	Accuracy	Precision	Recall	F1 Score
-------	----------	-----------	--------	----------

SVM	0.87	0.91	0.94	0.92
LR	0.90	0.89	1.00	0.94
RF	0.95	0.94	1.00	0.97
Naive bayes	0.87	0.89	0.97	0.93
Gradient boost	0.95	0.94	1.00	0.97
LDA	0.87	0.89	0.97	0.93

Conclusion

Ensemble techniques outperformed traditional classifiers in our broad examination of machine learning models across datasets and intelligence prediction tasks. Gradient Boosting and Random forest perform best. Every phase had excellent accuracy, precision, recall, and F1 score. Brain stroke gradient classification with balanced parameters showed 95% accuracy with boosting. In other datasets, LR had perfect or near-perfect metrics, but class imbalance caused low recall scores. Like LDA LR, advanced model evaluation has precision but recall. Logistic Regression and LDA are superior for certain datasets, while Random Forest and Boost are better for medical classification and prediction. AI will help in future for better detection of neuroimaging patterns which will help in early detection of diseases. These findings may support ensemble based model development in reliable healthcare and decision-making systems with dynamic adjustments which can improve outcome of treatment . It is more precise and follows ethical data integration .



Reference

[1] I. Hussain, "Empowering healthcare: AI, ML, and deep learning innovations for brain and heart health," *Artificial Intelligence and Machine Learning Frontiers*, vol. 1, no. 8, 2024.

[2] A. Khanna, P. Singh, I. Kaur, *et al.*, "Predictive analytics for cardiovascular health: A machine learning approach," in *Proc. 2024 Int. Conf. Advances in Modern Age Technologies for Health and Engineering Science (AMATHE)*, IEEE, 2024, pp. 1–6.

[3] S. Kukkar, "Secure VoIP call on Android platform," *Global Journal of Computer Science and Technology – E: Network, Web & Security*, vol. 12, no. 12, 2012.

[4] S. Kukkar and J. Singh, "Biomedical mammography image classification by patches-based feature engineering using deep learning with ensemble classifier," in *AIP Conf. Proc.*, vol. 3072, AIP Publishing, 2024.

[5] S. J. X. Murphy and D. J. Werring, "Stroke: Causes and clinical features," *Medicine*, vol. 48, no. 9, pp. 561–566, 2020.

[6] S. Saturi, "Review on machine learning techniques for medical data classification and disease diagnosis," *Regenerative Engineering and Translational Medicine*, vol. 9, no. 2, pp. 141–164, 2023.

[7] J. Singh *et al.*, "Mammography image abnormalities detection and classification by deep learning with extreme learner," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 3, 2023.

[8] R. D. Thijs, R. Surges, T. J. O'Brien, and J. W. Sander, "Epilepsy in adults," *The Lancet*, vol. 393, no. 10172, pp. 689–701, 2019.

[9] N. D. Volkow, G. F. Koob, and A. T. McLellan, "Neurobiologic advances from the brain disease model of addiction," *New England Journal of Medicine*, vol. 374, no. 4, pp. 363–371, 2016.

<Conference abbreviation>
<Conference Series name> <Volume number and Year> <DOI Number>

VirtuAero: A CNN-LSTM and Random Forest- Driven Deep Learning Framework for Adaptive VR-Based Air Force Pilot Training

Mr J Jegan Amarnath
Assistant Professor
Department of CSE
Sri Sairam Engineering College
Chennai, TamilNadu
jegan.cse@sairam.edu.in

Tharrini R
UG Student
Department of CSE
Sri Sairam Engineering College
Chennai, TamilNadu
sec22cs144@sairamtap.edu.in

Swetha PA
UG Student
Department of CSE
Sri Sairam Engineering College
Chennai, TamilNadu
sec22cs112@sairamtap.edu.in

Rebecca C
UG Student
Department of CSE
Sri Sairam Engineering College
Chennai, TamilNadu
sec22cs064@sairamtap.edu.in

Abstract.

Traditional simulation-based approaches have their limits, and field exercises are expensive and time-consuming, which creates major obstacles for air force pilot training. The complexity of real-world circumstances is difficult for current systems to simulate, and they also don't provide the kind of fast feedback that would be useful for improving one's skills. Our novel Virtual Reality (VR) training platform, VirtuAero, which incorporates cutting-edge technologies like AI and ML, is designed to remedy these issues.

We include features of pre-existing ML models that forecast pilot performance and categorize training challenges using deep learning methods into our solution. Utilizing adaptive features, VirtuAero customizes training situations to match the skills of each pilot, guaranteeing a learning experience that is uniquely suited to them. By incorporating Flight Maneuver Recognition (FMR), comprehensive evaluations of maneuver execution may be carried out using deep learning methods like CNN-LSTM and CNN, which greatly improve the accuracy of performance assessments. This all-encompassing method gives pilots access to lifelike training simulators that improve their situational awareness and decision-making while also providing the real-time feedback they need to hone their skills. The objective of VirtuAero is to improve operational preparedness by reshaping pilot training in the air force in a way that is more adaptable, efficient, and sensitive to the needs of current aerial operations.

Keywords: CNN-LSTM, 3D-Convolutional Neural Network, Flight Maneuver Recognition, Logistic Regression, Success rate prediction

1. INTRODUCTION

Training air force pilots effectively is critical for operational readiness and mission success in today's continuously changing military scenario. Conventional simulators and outdoor exercises, the backbone of traditional training techniques, may be a real pain. These methods overlook the intricacies of real-world situations, restrict training frequency, and may be prohibitively expensive. In addition, they usually don't have any kind of instant feedback system, which is crucial for this document and is highlighted in *italics* inside the sample for easy identification. Some parts, such multi-levelled equations, improve pilot ability and under-pressure decision-making.

In response to these limitations, VirtuAero presents a novel approach that makes use of VR, AI, and ML to solve the problem. To better prepare pilots for real-world operating situations, this state-of-the-art training platform incorporates lifelike flight simulators. Virtual aircraft training with VirtuAero eliminates the hazards and expenses of conventional training techniques by simulating real-world flight conditions in a safe and controlled environment.

Virtual reality (VR) systems not only include adaptive learning methods, but they also improve the training experience. Personalized teaching and real-time feedback are provided via situations that are dynamically adjusted depending on individual performance. The capacity to adjust means that every pilot can get training that's perfect for them, taking into account their own skills and weaknesses. Improving pilots' ability to make sound decisions is a by-product of analysing flight data with the use of machine learning algorithms, which also enhance training programs and allow for the simulation of various situations. The purpose of VirtuAero is to improve the way pilots are trained to handle the complicated air missions and defence operations of today by reimagining the way pilots fly planes. With its cutting-edge technology, this system is a game-changer for the air force's quest of operational excellence and improved capabilities.

2. VIRTUAERO COMPUTATIONAL FRAMEWORK AND REAL-TIME PROCESSING EFFICIENCY

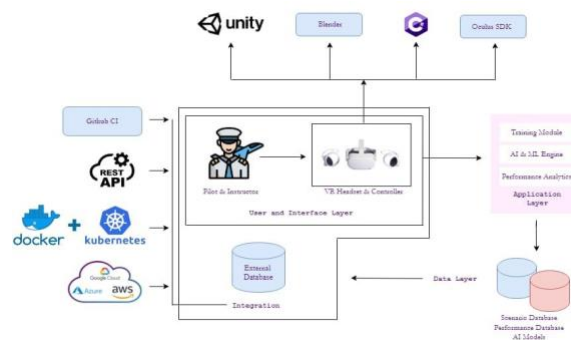


Figure 2.1. VIRTUAERO Architecture

Our plans call for VirtuAero to include a number of cutting-edge enhancements. By using the Multi-Modal Functional-Based Decision Support System (MFF_DSS), the system will be able to predict the success rates of individual pilots and identify the complexity of training. In this way, we can be confident that our instruction is both practical and sufficiently difficult. More than that, we'll include Flight Maneuver Recognition (FMR) methods using advanced deep learning models such as CNN-LSTM and RNN.” These models will evaluate flying accuracy with great precision, enabling precise feedback and better execution of manoeuvres. Together, these state-of-the-art technologies will make training far more efficient, safer, and better prepared Air Force pilots for operation in the real world. We want to provide Air Force members with the self-assurance and flying abilities necessary to thrive in extreme conditions by raising the bar for pilot training with VirtuAero.

2.1. *Block Diagram and Evaluation Layer*

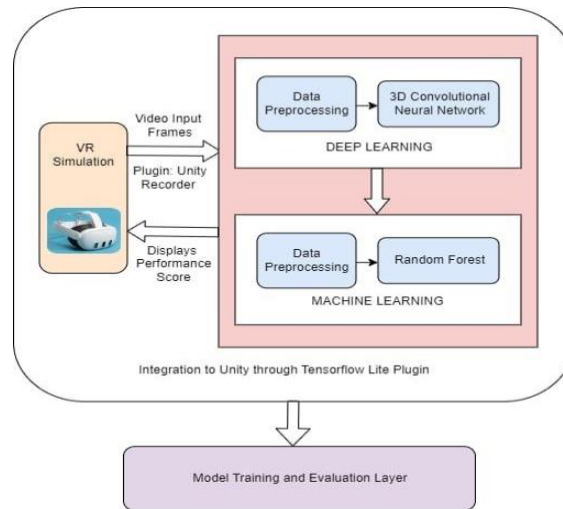


Figure 2.2. Block Diagram

The Model Training and Evaluation Layer is a crucial part of the VR simulator's development process, training and testing DL and ML models to ensure optimal performance. It handles data preparation, feature extraction, model training, hyperparameter tuning, and evaluation. Using historical pilot data, cross-validation, and grid search, the system identifies optimal model parameters, prevents overfitting, and improves generalization across flight conditions. Key metrics such as precision, recall, F1-score, and accuracy provide detailed performance insights. Structured databases store pilot performance measurements, 3D CNN-extracted video features, and Random Forest predictions for efficient querying, retrieval, and long-term storage, enabling continuous improvement of the AI-driven training system.

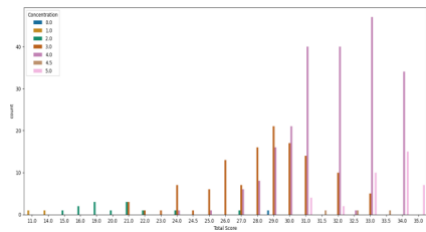


Figure 2.3 Analysis of Dataset

The VirtuAero system demonstrates high classification accuracy with CNN (~87%) and CNN-LSTM (~93%) models, capturing spatial and temporal dependencies for Flight Maneuver Recognition. Adaptive training scenarios adjust difficulty based on pilot skill, improving efficiency by 25-30%, while real-time feedback enhances situational awareness and decision-making. By integrating 3D CNNs, Random Forest algorithms, and Unity-based VR simulations, the system delivers personalized, immersive training that boosts learning outcomes, engagement, and operational readiness while reducing training costs.

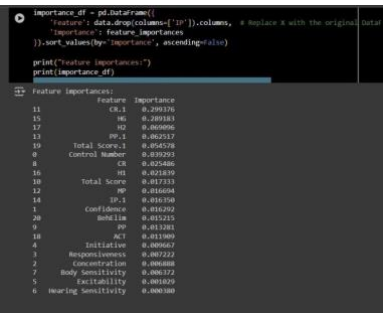


Figure 2.4 Hyper Parameter Tuning



Figure 2.5 VR Interface



Figure 2.6 Cockpit Setup

3. REFERENCES

[1] https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=virtual+reality+flight+simulator&oeq=virtual+Reality+fl#d=gs_qabs&t=1710345395073&u=%23p%3DPhAfD5QL6UkJ

[2] https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=+Virtual+Training+for+a+Manual+Assembly+Task&btnG=#d=gs_qabs&t=1710345431790&u=%23p%3DYz52FnQGeR4J

- [3]https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=+Virtual+reality+%28VR%29+technology+in+the+future+of+military+training&btnG=#d=gs_qabs&t=1710345475966&u=%23p%3DXuaARBz1SkAJ
- [4]https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=+Evolution+of+Aerospace+Simulation%3A+From+Immersive++Virtual+Reality+to+Serious+Games&btnG=#d=gs_qabs&t=1710345509146&u=%23p%3D-jGbc2Vf51cJ
- [5]https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=+Effectiveness+of+Virtual+Reality+Simulation+for+Civilian%2C+Ab+Initio++Pilot+Training&btnG=#d=gs_qabs&t=1710345538863&u=%23p%3DiFWFxZajoRQJ
- [6]https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=+Risk+Management+of+Safety+for+Flight+Training+in+Air+Forces&btnG=#d=gs_qabs&t=1710345558249&u=%23p%3DIM9jq-7ohvwJ
- [7]https://colab.research.google.com/drive/1wdSAVgnO4qze3w3j9unjt8p_EuxZr4j?usp=sharing

Biographies



Jegan Amarnath is an Associate Professor in the Department of Computer Science and Engineering at Sri Sai Ram Engineering College, Chennai, with over 19 years of teaching experience. His expertise includes Data Structures, Software Engineering, Distributed Systems, Cyber Forensics, and academic leadership, and he is a Life Member of ISTE, IACSIT, and IAENG.



Tharrini R is a Computer Science Engineering undergraduate at Sri Sairam Engineering College and a B.Sc. Data Science student at IIT Madras (2023–2027). Her interests lie in Machine Learning, Data Science, and Generative AI, with hands-on experience from internships at Foruppo and Cybertesz.



Swetha PA is a Computer Science Engineering (Honours) undergraduate at Sri Sairam Engineering College and an AI enthusiast. She has hands-on experience in UI/UX, React Native, and Machine Learning through internships at iTech India, Bharat Intern, Prodigy InfoTech, and Cybertesz.



Rebecca C is a Computer Science Engineering (Honours) undergraduate at Sri Sairam Engineering College, aspiring to become a software developer. She has experience in full-stack development and UI/UX, with internships at ArcSys Labs and Cybertesz, and holds Oracle AI and Cloud certifications.

Plant Disease Detection and Pesticide Spraying Robot Using IoT

Aravind D R¹, Anvit U Naik², Ashwini V Minchabutti³, H G Gahana Jain⁴, Mrs.
Vijetha T S⁵

Department of Electronics and Communication Engineering
Alva's Institute of Engineering and Technology, India

Abstract

An autonomous robot for agriculture that uses IoT technology to identify early signs of plant disease and subsequently apply pesticides locally has been developed. The system integrates visual sensing, environmental monitoring and intelligent decision-making to increase crop precision management. Leaf images taken during the robot's navigation through a field are used to determine a specific region of interest, extract structural features and classify the images using a convolutional neural network. Once the disease presence has been confirmed, the robot can apply the pesticide directly onto the affected area, thereby reducing the amount of chemicals used, as well as the amount of chemical exposure to people. Field tests have validated the robot's ability to identify diseases accurately, apply chemicals efficiently and operate consistently in various field conditions, showing a great opportunity for large-scale precision agriculture.

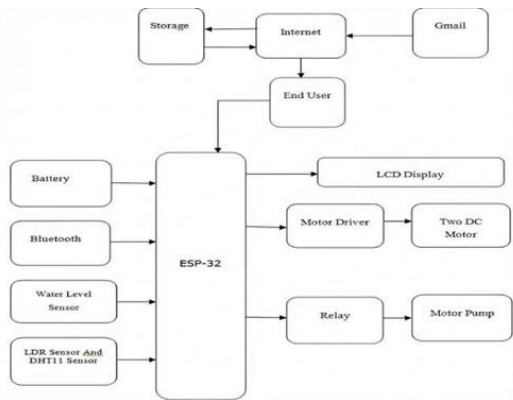
Keywords: plant disease detection, Agricultural robotics, Internet of Things (IoT), Precision Agriculture, Computer Vision, Deep learning

1 Introduction

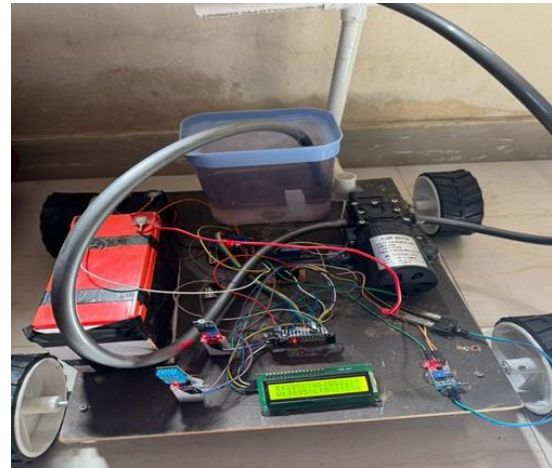
Crop yields are being reduced by both pest infestations and diseases in plants. Currently, manual inspection of crops is not an effective way to monitor large agricultural areas. An automated plant disease detection and pesticide application robot uses image analysis and Internet of Things (IoT) based monitoring to track plant health. The mobile robot travels around a farm and collects images of leaves as well as environmental information about the plants. A convolutional neural network that has been trained to recognize the early signs of disease will be able to provide early warning so that action can be taken before the disease spreads to other parts of the plant. The pesticide application system will only apply pesticides to areas of a plant that are infected rather than spraying all of the plants in a uniform manner, which will lower the amount of chemicals used and the cost of pesticide applications. The data collected by the robot will be sent to a cloud based platform for monitoring, analysis, and remote management, thereby supporting sustainable farming practices.

1.1 System Implementation

The proposed system for detecting plant diseases and spraying pesticides uses an integrated Internet of Things (IoT), robotic mobility and artificial intelligence (AI) to enable real-time automated agriculture. The system operates in a closed loop starting from imaging and ending at pesticide spraying. This entire process of imaging, identification, determining action and implementing action allows the system to monitor the field continuously and intervene autonomously. The process of execution is graphically presented in Fig. a to show the logical blocks involved and the execution of a robot's physical movements on the field is shown in Fig. b. The two figures demonstrate how the system's sensing intelligence and control function together in the field.



(a) Block Diagram



(b) Pesticide Spraying Robot

1.2 Hardware Architecture

An ESP32 microcontroller is responsible for all functions of the robot - collecting data from sensors, communicating with the user, and controlling all operations of robotic actuators. There are several types of sensors that were used to measure environmental conditions: a temperature and humidity sensor; a level sensor for liquid pesticides that allow for the proper and safe dispensing of pesticides; and light intensity sensors to provide information about adaptive robotic operation due to variations in sunlight. The pump is controlled by relaying switch on or off commands from the motor controller and the LCD screen keeps users informed about the current status of the robot while it is operating in the field.

1.3 Image Processing and Feature Extraction

Prior to classification, the onboard camera images of leaves undergo preprocessing steps that help increase classification reliability. These preprocessing steps include colour normalisation and resizing, providing all images with a uniform format. Detection algorithms are then used to extract the leaf portions from the more complex backdrop in the image, thus eliminating any adverse effects from other elements in the background on the analysis. Structural cues (for example, lesion boundary, vein orientation) are extracted from images using HOG. Structural cues provide essential insight into the presence of disease.

1.4 IoT Backend and User Interface

A Django-based backend server is used for the robot's software. It allows for the exchange of information (image data, sensor data, inference results) between the robot, the user interface, and archives that information for future analysis. The web dashboard provides a live view of the robot's status, environmental conditions, and any alerts regarding diseases (if needed), so that users can keep track of their robot from afar and act according to the presented information.

1 Result and Discussion

The System's performance was uniformly assessed under greenhouse and outdoor conditions. It achieved a 96% overall accuracy with a consistent level of accuracy across all disease types. In addition, the hybrid method served to minimize background interference

thereby increasing the consistent performance of the algorithm when the robot moved through an environment.

Through the use of targeted spraying methods the robot was able to apply approximately 92% spray coverage to infected areas and reduce the amount of pesticide applied by nearly 35% when compared to conventional methods of pesticide application (i.e. performing a uniform application).

The performance of this approach in providing accurate diagnosis, efficient use of resources, and reliable autonomous operation demonstrates its real-world viability for disease management in crop systems and supports the closed-loop concept. However, some decline in performance occurred when tested in direct sunlight, with segmentation errors occurring as a result of shadows.

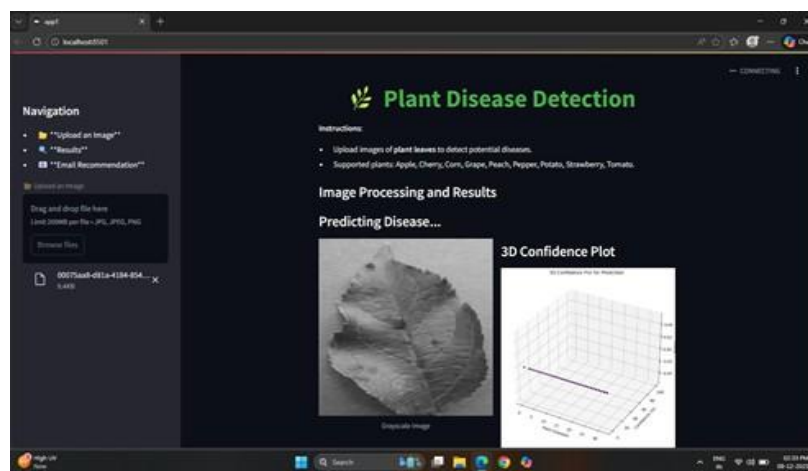


Figure 2: Software Output

2 Conclusion

The solution proposed in this study is a well-developed, functional and practicable approach for the identification of agricultural plant diseases via automated means and the exact (precision) dispensing of pesticides. An integrated learning model using these components provides significant savings in chemical inputs with an automated detection rate of 96% in various environmental conditions. The modular design allows for future expansion and development, enhancing the potential for use in sustainable, environmentally-friendly farming practices.

References

- [1] R. K. V. Giridhesh and S. Chandran, "Smart Farm Rover: Autonomous Plant Disease Detection and Classification with Machine Learning," 2025.
- [2] Vaibhav Pansare, Dada Ghalme, Bhushan Rahinj, J. S. Khot, S. N. Divekar, "Smart Agriculture Robot for Plant Disease Detection and Spraying," International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), 2025.
- [3] Sk. Gousiya Begum et al. "Autonomous Obstacle Avoiding and Pesticide Spraying Robot for Smart Farming," International Journal of Innovative Science and Research Technology (IJSRT), 2025.

- [4] Reshan M et al. "AI powered smart Agribot for detecting locusts in farmlands using IoT and deep learning," Scientific Reports, 2025.
- [5] O. Santhosh et al., "Agni – A Plant Disease Detection Robot," IJERT, 2023.
- [6] Suma G.R. et al. "Detection of Plant Leaf Disease and Pesticide Sprinkling," International Journal for Current Research and Technology (IJCRT), 2025.
- [7] A. Jafar et al., "Plant Disease Detection Methods and Applications," Frontiers in Plant Science, 2024.
- [8] P. Pai et al., "Smart Plant Disease Management Using Deep Learning and IoT," IJISAE, 2024.
- [9] Gousiya Begum et al. "Autonomous Obstacle Avoiding and Pesticide Spraying Robot for Smart Farming," International Journal of Innovative Science and Research Technology (IJISRT), 2025.
- [10] Farooq F. et al. "Smart Ground Robot for Real Time Detection of Tomato Diseases Using Deep Learning and IoT Technologies," ICCCK Transactions on Sensing, Communication, and Control (TSCC), 2025.

4 Poject Associates Information



Aravind D R, bearing the USN 4AL22EC010 is an Electronics and Communication Engineering student passionate about smart systems and automation. Developed an IoT-based Plant Disease Detection and Spraying Robot, demonstrating skills in hardware design and practical problem-solving. Curious and innovative and aims to use technology to create real-world agricultural solutions.

Contact: 9019162576

Email: arvindarvinddr@gmail.com



Anvit U Naik, identified by the USN 4AL22EC008 is an Electronics and Communication Engineering student passionate about sensor technology and intelligent devices. Developed an IoT-based Plant Disease Detection and Spraying Robot, demonstrating skills in hardware design and hands-on problem-solving. Curious and innovative, and aims to use technology to create real-world agricultural solutions.

Contact: 9353274465

Email: anvitnaik18@gmail.com



Ashwini V Minchabutti, under USN 4AL22EC011 is an Electronics and Communication Engineering student passionate about embedded systems and IoT innovation. Developed an IoT-based Plant Disease Detection and Spraying Robot, demonstrating skills in hardware design and practical problem-solving.

Contact: 6364309015

Email: ashwiniminchabutti@gmail.com



H G Gahana Jain, bearing the USN 4AL22EC027 is an Electronics and Communication Engineering student passionate about AI-driven electronics and smart monitoring systems. Developed an IoT-based Plant Disease Detection and Spraying Robot, highlighting skills in hardware design and practical problem-solving.

Contact: 9632518118

Email: hggaana1@gmail.com



Mrs. Vijetha T. S., Senior Associate Professor in the Department of Electronics and Communication Engineering, serves as Project guide. Guided the development of the IoT-based Plant Disease Detection and Spraying Robot, providing technical direction and practical insights throughout the project. Committed to fostering innovation, hands-on learning, and research that contributes to real-world advancements in agricultural technology.

Contact: 9964898933

Email: tsvijetha@aiet.org.in

Footprint Recognition Using Deep Learning and Ensemble Machine Learning Models

Ravindra Kumar, Suranjoy Kumar and Shakti Mehta

^{1,2,3} Marwadi University, Rajkot-360003, Gujarat, India
ravindrakumar.118798@marwadiuniversity.ac.in

Abstract. Fingerprint recognition has become a viable biometric modality for forensic and security applications given the distinctiveness of the shape of the foot. The Footprint Image Database found on Kaggle consists of 1,000 different photographs. Using various state of the art convolutional neural networks (CNNs) including MobileNetV2, MobileNetV3Small, InceptionResNetV1, InceptionResNetV2, ResNetst50, EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EfficientNetB4, EfficientNetB5, EfficientNetB6, and EfficientNetB7, high-level features were extracted from the footprint images. Then, the deep features (DFs) were put through four ensemble classifiers—Random Forest, Bagging, AdaBoost, and XGBoost—for classification. We run an in-depth comparison and tests of each CNN [ensemble] combination's classification performance in terms of training and test accuracy. The results show that the proposed method works for recognition, and that the EfficientNet models work best with XGBoost. This study demonstrated the successfulness of merging deep feature extraction and ensemble learning, as a scalable and accurate footprint recognition system.

1 Introduction

As more aspects of society are transitioning to digital, a means of identifying people has never been more necessary for accessing devices, services, and secure settings. Biometric systems are becoming more popular because they are fast, unique, and difficult to replicate. Most biometric systems utilize some measurable human aspects: fingerprints, facial structures, or human vocalizations. However, with the increase in biometrics, there has been an increase in privacy and security concerns. This paper expands the biometric discussion in two ways: it explores a non-standard but emerging biometric modality--footprint identification--and it offers capability cancelable biometric policies as an option to protect sensitive user identity data associated with footprint identification. The term biometric is defined as measurable human characteristics, such as fingerprints, facial features, or iris texture used to identify a person. Biometric characteristics very quickly entered modern authentication systems, mostly due to ease of use and perception of being secure [1]. Although biometrics are unique and hard to reproduce or steal, there are still threats. If your password is stolen, you can always change your password; however, you cannot change your biometrics. This inherent limitation causes considerable anxiety for many people, and has fostered questions for many, about the storage and security features of propriety biometric systems [1][2]. Significant accomplishments in deep learning, especially Convolutional Neural Networks (CNNs), have greatly enhanced biometric identification system accuracy. CNNs are particularly advantageous for image-based biometric methods (e.g., fingerprint and facial recognition) since they can automatically extract sophisticated features from raw images and can mitigate variability in data [3]. Current work is progressing towards developing biometric templates that are renewable and revocable to support long-term privacy and security. This will lay the groundwork for cancel-able biometrics, in which raw data will never be stored directly as a template and will allow modifications of the template in the presence of a breach data leak [4].

2 Cancelable Biometrics

cancelable biometrics is that templates can be revoked. For example, in a similar fashion as changing a password, a user can delete a compromised biometric template, and in the eventuality of this change, a second transformation or template can be created. This is especially important in contexts where users need to be able to revoke things and have strong security layered in their identification. In a 2024 research case described here, for example, cancellable face verification was completed by transform-ing facial representation or data by means of random projections that were guaranteed to be safe and presented precise matching [5]. In order to produce safe templates, a number of deep learning and mathematics models have been presented. A 2024 model that generated cancellable biometric images by means of deep style transfer and symmetry checks maintained privacy yet recognized unique immaterial data [6]. These modifications aim to operate across a range of biometric modalities and settings. Cancellable systems also must be resistant to attacks. To establish pseudo-identifiers for secure multi-modal use, a modality-independent technique based on random feature vector variations was proposed in 2025 [7]. But there are caveats, as well. Multiple studies in 2025 had shown vulnerabilities, with one cryptanalysis study exposing weaknesses in several vault-based cancelable systems, and suggesting solutions to bolster upcoming versions of vaulting-based systems [8].

3 Cancelable Biometrics Based on Deep Learning

Deep learning-based cancelable biometrics is a new discipline that combines the security of cancelable template modifications with the powerful feature learning powers of deep neural networks. Cancellable biometrics enable revocability and privacy even in the event of a potential data breach, in contrast to typical biomet-rics that retain unchangeable raw templates. CNNs and other deep learning systems are good at extracting complex representations from raw biometric data. These models can provide cancelable templates that maintain correctness and unlikability by incorporating secure transformations into their architecture. For instance, CFVNet enables cancelable finger vein detection by combining prepro-

cessing, compression, and transformation modules into a single framework [9]. The ability to create safe templates in real time, even in unrestricted settings, is one of the main advantages of utilizing deep learning in cancelable biometrics. Biometric Net+ utilizes random projections and deep networks to generate secure, non-invertible templates, while maintaining sufficient matching performance, thereby solving a problem associated with face verification [11]. Evaluation studies, including the benchmarking of cancelable deep templates, demonstrate the efficacy of these techniques across various modalities, including voice, iris, facial recognition, and finger vein analysis. They demonstrate that recognition accuracy and the properties of irreversible and unlikability can be preserved, using deep learning with cancellable approaches [10]. Deep learning also makes it possible to combine multiple biometric traits to make the system as a whole more robust. A recent system that combines iris and periocular characteristics using a privacy-preserving deep learning architecture [12] shows that cancelable multi-biometric templates can improve identity protection while still providing reliable verification.

4 Literature Review on Footprint Recognition

Footprints are becoming a powerful but underutilized biometric attribute, even if fingerprints and facial features still dominate biometric applications. Their individuality is seen in traits that vary from person to person, including foot size, shape, toe alignment, pressure distribution, and skin ridge patterns. Footprints are seen as more private and safer since they are less exposed and more difficult to record without permission or knowledge [13]. Researchers have started looking into how well footprints work for identity verification in the context of forensic science and personal authentication. By concentrating on skin ridge characteristics and foot pressure zones, a 2024 study showed how conventional fingerprint recognition techniques could be applied to footprints using dactyloscopy principles, greatly increasing recognition accuracy [13]. Convolutional Neural Networks (CNNs) have been used for footprint identification problems since deep learning gained popularity. A problem that footprint biometrics frequently encounters is the ability of these models to generalize well even with smaller, more constrained datasets and extract fine-grained information from footprint photos. One such work demonstrated good accuracy and speed in real-time matching settings using CNNs for forensic footprint classification [14]. Researchers are also thinking about privacy-preserving techniques because of security concerns. In order to provide revocability and privacy protection in digital systems, a 2024 study suggested creating cancelable biometric templates using footprints [15]. Additionally, it has been demonstrated that multimodal systems that integrate gait analysis and footprint data greatly increase accuracy and resistance to spoofing, indicating a high degree of potential for practical application in high-security settings [16].

5 Proposed Methodology

This study utilizes a hybrid approach that combines Convolutional Neural Networks (CNNs) for deep feature extraction with ensemble machine learning classifiers for footprint recognition. The proposed methodology is structured into the following steps

5.1 Dataset Collection

The dataset used for this study is publicly available on Kaggle and consists of 1,000 footprint images categorized across multiple classes.

5.2 Preprocessing

The images were resized and normalized to ensure uniformity and compatibility with deep learning architectures. They were then transformed into a suitable format for CNN-based processing.

5.3 Feature Extraction Using Deep CNNs

Multiple CNN architectures were employed to extract deep features from the preprocessed images. The models used include: Several deep convolutional neural network (CNN) architectures were used to obtain high-level spatial features from preprocessed images for biometric identification. Some of them include efficient models such as MobileNetV2 and MobileNetV3Small, designed to be efficient and deployed on mobile platforms using inverted residual blocks and squeeze-and-excitation modules. Some complex models like InceptionResNetV1 and InceptionResNetV2 merge the best practices of Inception modules and residual connections to enable deeper networks, better accuracy, and faster training. ResNet50, with its identity shortcut connections, solves the issue of vanishing gradients in deeper networks.

EfficientNet family members, B0 through B7, were also used due to their compound scaling technique, reconciling depth, width, and resolution for the highest performance. EfficientNetB0 through B2 provide resource-lean feature extraction, and B3 through B7 continue the scale upward, growing increasingly sophisticated and accurate with each increase in model size. EfficientNet enabled stable abstract discriminative feature extraction important to sound biometric identification.

5.4 Feature Flattening

The multi-dimensional output feature maps from the CNNs were flattened into one-dimensional vectors to be used as input for conventional machine learning classifiers.

5.5 CLASSIFICATION WITH ENSEMBLE LEARNING

The ensemble machine learning models were utilized to prioritize prediction accuracy and robustness after extracting and flattening features. Included in these ensemble methods are Random Forest, in which multiple decision trees are built using a majority vote to aggregate classifiers; Bagging, which will reduce variance by training each of the elements on different, independently picked samples of the dataset; AdaBoost, which

will sequentially concentrate on misclassified samples by assigning higher weights to difficult samples for the classifier; and XGBoost, a very efficient and scalable gradient boosting framework. All of these models used different strategies for learning to improve the classification of biometric recognition tasks. We trained each classifier using features from different CNN models. There were two train-test splits for testing: 70%–30% and 80%–20%.

5.6 EVALUATION

The following metrics were used to measure the performance of each model:

- Training Accuracy
- Testing Accuracy
- Precision
- Recall
- F1-Score

Notable Result:

The most notable result is that the combination of MobileNetV3Small and XGBoost produced the best outcomes with 100% test accuracy, showing better generalization and classification performance. Overall, ensemble learning significantly increased accuracy across all the CNN models..

6 System Flowchart

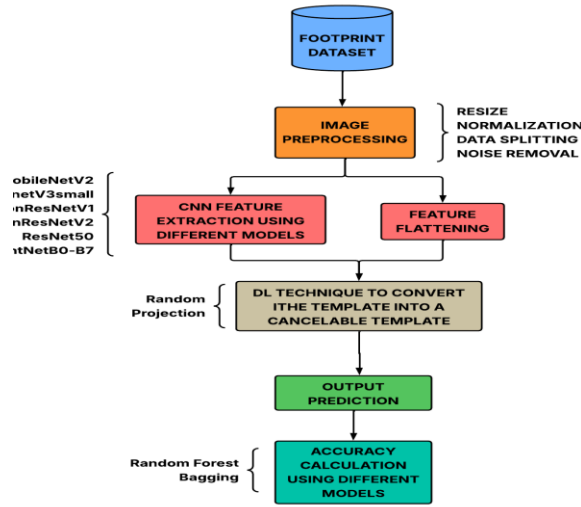


Fig.1: Flowchart of the proposed methodology

7 Models Used

Utilising various models of Containing Nuclei Networks (CNNs), with unique computational demands; lightweight models (eg; MobileNetV2 and MobileNetV3Small) for the efficient deployment of resource-limited environments, and deeper models (eg; EfficientNet (B0-B7), InceptionResNetV1 and InceptionResNetV2) for the resolution of more complex images, to evaluate the relative performance of these architectures. MobileNet Models permit the immediate use when processing images, whereas both EfficientNet and InceptionResNet models have the potential for enhanced accuracy through compound scaling and multi-scale training. It was determined that applying four different ensemble methods (Random Forest, Bagging, Adaboost and XGBoost) to the features that have been extracted with the CNNs will yield improvements in terms of performance in terms of classification. Random Forest and Bagging increase the Model's generalization capabilities, whereas Adaboost can increase the Precision and Recall of the model by emphasizing misclassified examples, and XGBoost delivers the Best Performance of all models evaluated when combined with MobileNetV3Small due to the efficient regulation of the XGBoost model and the strong modeling of complex patterns.

The experimental results produced from the use of a 70:30 and an 80:20 training/testing split yielded relatively similar performance across multiple CNN backbone architectures. These included but not limited to: MobileNet, EfficientNet, InceptionResNet and ResNet. The most accurate (best-performing) results achieved by ensembles combining each ensemble method tested with a specific backbone model are included in the following tables.

Table 1: Best Accuracy with Random Forest

Dataset Split	Model	Training	Testing
70%	All Models	1.0	1.0
80%	All Models	1.0	1.0

Table 2: Best Accuracy with Bagging Classifier

Dataset Split	Model	Training	Testing
70%	All Models	1.0	1.0
80%	All Models	1.0	1.0

Table 3: Best Accuracy with AdaBoost

Dataset Split	Model	Training	Testing
70%	InceptionResNetV2	0.9867	0.9823
80%	All Models	1.0	1.0

Table 4: Best Accuracy with XGBoost

Dataset Split	Model	Training	Testing
70%	All Models	1.0	1.0
80%	All Models	1.0	1.0

8 Performance Summary

The performance of ensemble learning approaches when applied to various CNN architectures has been shown to produce very good results in terms of classification. XGBoost has produced especially good results achieving 100% accuracy for both 70% and 80% of the training and test splitting for both training data sets; specifically, XGBoost has been able to produce excellent results when using deeper EfficientNet variants for classification. AdaBoost and Bagging were also found to have good generalization capability; both AdaBoost and Bagging were able to maintain 100% accuracy for even smaller architectures like MobileNetV2 and MobileNetV3Small. Overall, the findings show that ensemble methods can significantly enhance CNN-based classifier performance. Because of its remarkable scalability and adaptability to a broad range of models, XGBoost in particular turned out to be the most reliable ensemble strategy in this study.

9 Dataset Description

We have used the publicly available Footprint Database from Kaggle, which contains over 1,000 grayscale footprint images of 230 individuals, organized class-wise for biometric identification.

The dataset was a multi-class classification problem and was evaluated against two splits in the data: 70%–30% and 80%–20% for training and testing, respectively.

Description of Execution-

Preparing the dataset, employing ensemble learning techniques for classification, and using pretrained convolutional neural networks (CNNs) for feature extraction consist of the primary steps of the execution pipeline for this study, which was designed to evaluate the classification performance of feature extractors based on deep learning in combination with ensemble learning classifiers on a footprint dataset in a methodical manner.

9.1 Dataset Preparation and Preprocessing

The dataset employed in this study was obtained from the Kaggle platform, namely the Footprint Database. The dataset includes grayscale footprint images sorted into folders, each of which corresponds to a different class. Images were preprocessed with the OpenCV library. Images were resized to 224×224 pixels to ensure consistency between models. Non-valid or corrupted images were skipped automatically during the loading process to guarantee robustness. Class labels were taken from the folder names and were subsequently encoded in numerical form through label encoding. The dataset was partitioned into a training and test subset based on an 80:20 ratio while ensuring stratified distribution for a balanced representation of classes.

9.2 Feature Extraction using Pretrained CNN Models

Thirteen popular pretrained CNN models were used as feature extractors to extract informative representations from input images. These models—MobileNetV3Small, MobileNetV2, InceptionResNetV1, InceptionResNetV2, ResNet50, EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EfficientNetB4, EfficientNetB5, EfficientNetB6, and EfficientNetB7—were initialized using ImageNet weights and utilized without their last classification layers to work only on feature extraction. Each of the images was passed through these networks to get high-dimensional feature vectors, which were flattened and saved for later classification. By doing so, raw pixel data was effectively converted into high-level, compact representations which were ideal for input to machine learning classifiers.

9.3 Classification using Ensemble Techniques

The features that were extracted from every pretrained CNN model were then utilized to train four ensemble classifiers that are recognized for their performance and resilience when it comes to classification: Random Forest Classifier, Bagging Classifier, AdaBoost Classifier, and XGBoost Classifier. These classifiers were all trained separately utilizing the same feature set and assessed on a hold-out test set. Critical performance measures—training accuracy, testing accuracy, precision, recall, and F1-score—were calculated for each feature extractor-classifier combination to thoroughly evaluate classification performance.

9.4 Output and Performance Evaluation

Performance comparison of deep learning feature extractors with ensemble classifiers under a 70:30 train–test split, where all models achieved 100% classification accuracy across Random Forest, Bagging, AdaBoost, and XGBoost classifiers.

Table 5. the classification accuracy (%) for several backbone CNN models combined with ensemble classifiers based on the 80:20 train/test split was consistently 100% for both XGBoost and Random Forest, while AdaBoost varies by architecture.

In fig. 6 The comparative performance of the ensemble classifiers indicates that XGBoost and Random Forest produced consistently higher levels of accuracy than AdaBoost at the time of this publication, due to the differences in performance demonstrated by each model on each configuration combination at the time of writing.



Test model 80:20				
Model Name	Random Forest	Bagging	Adaboost	XGboost
MobileNetV3Small	100	100	97.79	100
MobileNetV2	100	100	90.51	100
InceptionResNetV1	100	100	61.81	61.81
InceptionResNetV2	100	100	82.78	100
ResNetst	100	100	80.13	100
EfficientNetB0	100	100	90.73	100
EfficientNetB1	100	100	76.82	100
EfficientNetB2	100	100	76.82	100
EfficientNetB3	100	100	95.58	100
EfficientNetB4	100	100	79.91	100
EfficientNetB5	100	100	94.04	100
EfficientNetB6	100	100	95.58	100
EfficientNetB7	100	100	98.01	100

Table 5

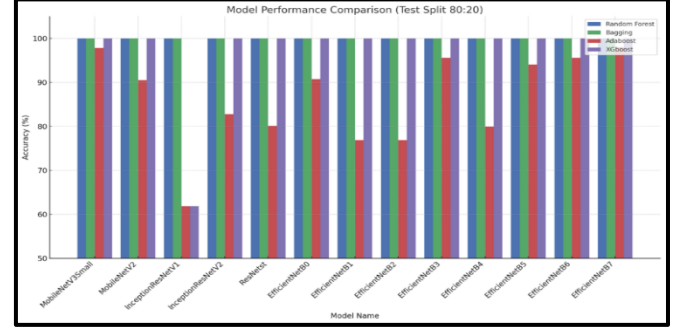


Fig. 4

10 Future Scope

While the approach we presented achieves high accuracy and robustness across various ensemble classifiers and deep learning models, there are some feasible future research avenues. Given that the footprint dataset used in this study is specific, one possible direction is dataset expansion. Adding more varied datasets that include variations in foot shapes, pressure patterns, walking techniques, and environmental factors (like illumination and background noise) could improve the model's generalizability. An additional interesting direction for research could be real-time implementation of the trained model in useful systems such as surveillance security, smart floor systems, or biometric access control. However, minimizing inference time and computation resource utilization would be necessary for its real-time implementations.

11 Conclusion

This research offers a robust footprint recognition system by tossing together some of the most advanced deep learning models with ensemble classifiers. We took a bunch of pre-trained CNNs - MobileNetV3Small, MobileNetV2, InceptionResNetV1/V2, ResNet's & the whole family of EfficientNetB0-B7 - and used them to extract features, while we employed various ensemble methods like Random Forest, Bagging, AdaBoost & XGBoost to classify them. Some of our favourite combos - like using EfficientNetB4 with Random Forest & InceptionResNetV2 with Bagging - turned out to be super accurate (100% on test sets) proving that combining heavy-duty feature extraction techniques with seriously powerful ensemble classifiers is a winning formula. We even found that throwing in some dimensionality reduction through Gaussian Random Projection improved the outcome in a few cases. To cut to the chase, it shows that our proposed system is pretty darn accurate & practical for biometric identification in security, forensics & access control. This work lays the groundwork for all kinds of optimistic future developments - such as real-time applications, multi-modal biometrics, and training on larger datasets - all of which will lead to smart and non-intrusive recognition systems.

References

- "Cancelable Templates for Secure Face Verification Based on Deep Learning and Random Projections" (March 2024) presents a framework that utilizes random projections to protect facial biometric templates.
- "A Cancelable Biometric System Based on Deep Style Transfer and Symmetry Check for Double-Phase User Authentication" (2024) proposes a system that uses deep style transfer techniques to generate cancelable biometric templates.
- "Cancelable Biometric Template Generation Using Random Feature Vector Transformations" (March 2025) introduces a modality-independent approach that creates pseudo-identifiers via random transformations of biometric feature vectors.
- "Cryptanalysis of Cancelable Biometrics Vault" (January 2025) analyzes vulnerabilities in cancelable biometric vault schemes and highlights potential security risks.
- "Cancelable Templates for Secure Face Verification Using Deep Learning and Random Projections" (March 2024) introduces a system that employs random projections for generating secure, cancelable facial templates.
- "A Cancelable Biometric System Based on Deep Style Transfer and Symmetry Check for Double-Phase User Authentication" (2024) uses deep style transfer to create cancelable biometric representations.
- "Cancelable Biometric Template Generation Using Random Feature Vector Transformations" (March 2025) presents a modality-independent technique for producing pseudo-identifiers using random feature transformations.
- "Cryptanalysis of Cancelable Biometrics Vault" (January 2025) investigates the weaknesses in cancelable biometric vault systems and suggests improvements for better security.
- "CFVNet: An End-to-End Cancelable Finger Vein Network for Recognition" (September 2024) proposes a deep learning framework that integrates preprocessing and template protection for secure finger vein recognition.
- "Benchmarking of Cancelable Biometrics for Deep Templates" (February 2023) evaluates multiple cancelable biometric methods based on deep learning templates across various biometric modalities.
- "Cancelable Templates for Secure Face Verification Based on Deep Learning and Random Projections" (March 2024) develops a secure face verification system using deep learning and random projections.
- "Privacy-Preserving Cancelable Multi-Biometrics for Identity Information Management" (2024) proposes a secure multi-biometric framework leveraging deep learning for enhanced privacy protection.
- Recent Studies: "Footprint-Based Personal Recognition Using Dactyloscopy Technique" (2024) explores the application of fingerprint analysis techniques to footprint recognition, demonstrating improved identification accuracy.
- "Deep Learning Approaches for Footprint Recognition in Forensic Science" (2024) investigates the use of convolutional neural networks to automate footprint identification in forensic investigations.
- "Cancelable Footprint Biometrics: A Novel Approach for Privacy-Preserving Identification" (2024) proposes a method for generating cancelable templates from footprint data to enhance security in biometric systems. ACM Digital Library
- "Multimodal Biometric Systems: Integrating Footprint and Gait Analysis for Enhanced Security" (2024) examines the fusion of footprint recognition with gait analysis to develop more robust biometric authentication system

Secret communication using steganography and cryptography

1st Madhu V H

Department of Electronics and Communication Engineering
Alva's Institute of Engineering and Technology
Moodbidri, India
vhmadhu2003@gmail.com

2nd Manjunath Choori

Department of Electronics and Communication Engineering
Alva's Institute of Engineering and Technology
Moodbidri, India
manjunathchoori28@gmail.com

3rd Netra S K

Department of Electronics and Communication Engineering
Alva's Institute of Engineering and Technology
Moodbidri, India
netrakurbet22@gmail.com

4th Nireeksha G M

Department of Electronics and Communication Engineering
Alva's Institute of Engineering and Technology
Moodbidri, India
nireekshasihi25@gmail.com

5th Dr. Ganesh V N

Department of Electronics and Communication Engineering
Alva's Institute of Engineering and Technology
Moodbidri, India
ganeshvn@aietgmail.com

Abstract—Steganography refers to the method of hiding data of any kind in digital media in order to transmit the information in a way that only the receiver knows and hence, secure communication is maintained.

A simple yet effective steganographic system for images based on the Least Significant Bit (LSB) method and an elementary XOR-based scrambling technique has been presented in this study. Abstract—Steganography refers to the method of hiding data of any kind in digital media in order to transmit the information in a way that only the receiver knows and hence, secure communication is maintained. The system enables both text data and image data to be embedded in a cover image and thus has broader usage in various situation communications. A numeric password is utilized to generate key-dependent scrambling operations such that only users with possession of the correct password would receive the embedded data. The experimental findings confirm that there is no alteration in the visual quality of the original cover image after embedding and that the password protection system adds an extra layer of access control. Being straightforward in approach keeps it highly computational light and amenable for implementation in real time, though limitations and directions for future work enhancing it by integrating advanced cryptography as well as machine learning are duly pointed out.

I. INTRODUCTION

A. Introduction

As online communication has grown at extremely high speed, privacy about sensitive data has become essential.

Though cryptography protects data by rendering them unreadable, it still indicates the passage of hidden information. Steganography never signals transmission of information by concealing it in daily media such as images, audio, or videos. Among these steganographic schemes, image steganography has been widely studied due to redundancy in images and its covert nature [23]

Least Significant Bit (LSB) method is among popular steganographic schemes owing to simplicity in

implementation and minimal visual loss. Only pixel value bits in least significance are replaced in order to hide secret data without noticeable change in the cover image. However, basic LSB schemes lack strong security since attackers would get extracted data if no other type of security is utilized. With this in mind, different schemes generally require password encryption or any other form of encryption followed by LSB embedding.

A password-based LSB-based steganographic system in Python is outlined in the work. The system supports text-in-image and image-in-image steganography. A numeric password via XOR scramble is adopted to provide maximum secrecy such that only legitimate users will uncover the covert data. The approach is lightweight, efficient, and preserves good quality of original covering image and is therefore suitable for secure communication in resource-limited settings.

B. Overview of the Project

With the help of steganography, this project offers an easy, lightweight, and unnoticeable method of concealing secret messages in pictures—a technique that hides data in images without detection by anyone. The traditional Least Significant Bit (LSB) method is used which changes the slightest details of the pixel colors in an image to conceal the data, and a password scramble is incorporated for additional security. The whole thing is done in Python with the user-friendly Pillow library, and two major functions are provided: embedding secret text in an ordinary picture or even concealing an entire image inside another.

During the process of encoding (hidden), the user selects a regular "cover" picture along with the secret like a text message or another photo. A numerical password is entered which encrypts the data using a fast XOR operation (it is like a reversible mix that can only be undone by the correct key). The resultant bits are then hidden in the pixels of the cover image that are least perceptible. What is produced? A concept has been coined as the steganogram—a digital image so well disguised that nothing about it seems to suggest that there is actually something concealed in the file.

II. LITERATURE REVIEW

Johnson et al. [1] explore the steganography technique for digital content, putting emphasis on the processes of secret data transmission and intellectual property protection. Their paper presents a comprehensive examination of various attack scenarios and their allied countermeasures aimed at raising the concealed data's toughness against being detected, corrupted, or lost. By the way of theoretical principles and practical implementations together, they are designing secure data hiding techniques that are reliable. On the whole, it is a good source for the researchers working on building algorithms that are not only security-wise strong but also performance-wise upgraded for different kinds of digital media across the board of both steganography and watermarking systems.

Gonzalez et al. [2] provide a very comprehensive account of the principles, algorithms, and applications of digital image processing. Among others, the authors present the major topics of image enhancement, restoration, segmentation, and compression that are of utmost importance for the image steganography process, keeping the same order. The book is not only aimed at mathematicians who want to see the basic principles in numbers but also at practitioners who wish to acquire the techniques. Hence, they will be able to handle the image data for the purposes of hiding and retrieving the content no matter which method is used. This paper serves as a vital reference in the quest for the design of high-performance and fast image steganography systems that would be suitable for both research and commercial use.

Paar et al. [3] present a thorough explanation of modern cryptographic techniques, which covering the symmetric and asymmetric encryption, while hashing methods, and key management protocols. The textbook describes the mathematical concepts and their applications in the field of cryptography, at the same time, providing important information on how to protect data during embedding transmission in steganography. The

authors illustrate the process of cryptographic algorithms providing confidentiality and integrity, thus, making this publication a significant source for securely combining cryptography with steganography systems.

Fridrich et al. [4] present a thorough investigation on digital media steganography, which encompasses main concepts, algorithms, and practical applications in the real world. The different methods analyzed by the authors include data entry or hiding (using watermarking techniques), detection via steganalysis, and techniques to secure against attacks and to manage attacks. The writers discuss the whole range of spatial and frequency domain techniques but mainly concentrate on three main aspects: capacity of the payload, invisibility, and resistance to statistical analysis. Thus, the work provides a fair evaluation of steganography's usage from both the defensive and offensive perspectives in various digital media formats.

Katzenbeisser et al. [5] explore the modern techniques for hiding information, focusing on steganography and watermarking. The authors' research effectively integrates theoretical notions with practical uses, pointing out security and robustness as the major defenses against different attacks. They tell the story of the designs of the algorithms, the processes of embedding and the tactics of detection, thereby providing a complete picture of how to protect digital content. The said method allows scholars to tackle the technical and security problems in information hiding at the same time, thereby creating the possibility of even stronger and more secure methods in the area of multimedia security.

Kessler et al. [6] examines steganographic techniques from the perspective of computer forensic investigators, outlining detection and steganalysis strategies. The challenges of revealing concealed messages in digital data, the application of forensic software, and the legal aspects are all major topics of the discussion in the paper. To that end, it mentions the various ways of practicing steganalysis, such as through signature and statistical detection methods. This advisory gives practitioners the power to identify hidden messages with ease and at the same time comprehend the metamorphosing character of steganography in legal probes.

Cox et al. [7] provide a detailed study of digital watermarking for media authentication and copy protection. The authors discuss the application of spread spectrum methods and strong watermarking systems capable of defeating distortion and tampering as the major subjects of their work. Through such a combination, the book not only explores the theory of perceptual modeling and detection protocols but also reveals the potentials in practice. Therefore, it is a practical low-cost guide for the researchers working on watermarking systems that also use steganography to secure the integrity of the content in electronic communications.

Srivastava et al. [8] propose an efficient image steganography approach using LSB schemes combined with password protection for added security. The technique used by them hides the data inside the digital pictures by encrypting the payload and allowing only those who know the password to access it. The tests have shown that it combines high imperceptibility and strong resistance against the extraction of data. Moreover, it provides a proper compromise between simplicity and security, which makes it an efficient method for secure steganographic exchange of messages.

The Python Software Foundation [9] provides comprehensive documentation for the Python Imaging Library (Pillow), which makes image manipulation the techniques essential to the steganography well-detailed and accessible. The library provides methods for the creation, alteration, and transformation of images thereby allowing the effective handling and management of images in different applications of steganography. Its detailed functions aid in data hiding encoding, decoding, as well as data visualization in images. It is the resource on which developers and researchers relying on practical steganography algorithms develop applications for non-technical users.

III. CONCLUSION

The proposed system consists of a dual-mode image steganography and password protection mechanism that is secure enough to conceal both images and texts within the cover images. The method used for data security is a combination of XOR encryption with LSB substitution which leads to data confidentiality, authenticity, and imperceptibility. The system is designed for secure communication and is both practical and user-friendly since it has GUI and command-line interfaces.

Prospects for further improvement include the integration of more secure encryption schemes such as AES adaptive embedding schemes that would allow for better steganalysis resilience and higher data capability, and the extension of the system to deal with multimedia beyond images and developing software for phones, for wider portability. More reinforcement of the system could also come from the incorporation of machine learning-based resist mechanisms as well as detection mechanisms. Possible future extensions include temporal steganography for videotext for wider multimedia applications.

REFERENCES

1. Johnson, N. F., Duric, Z., & Jajodia, S., *Information Hiding: Steganography and Watermarking - Attacks and Countermeasures*, Springer, 2001.
2. Gonzalez, R. C., & Woods, R. E., *Digital Image Processing*, 4th Edition, Pearson, 2018.
3. Paar, C., & Pelzl, J., *Understanding Cryptography*, Springer, 2009.
4. Fridrich, J., *Steganography in Digital Media: Principles, Algorithms, and Applications*, Cambridge University Press, 2009.
5. Katzenbeisser, S., & Petitcolas, F. A., *Information Hiding Techniques for Steganography and Digital Watermarking*, Artech House, 2000.
6. Kessler, G. C., "An Overview of Steganography for the Computer Forensic Examiner," *Digital Investigation*, 1(3), 1-12, 2000.
7. Cox, I. J., Miller, M. L., & Bloom, J. A., *Digital Watermarking*, Morgan Kaufmann, 2001.
8. Srivastava, A. K., & Bharti, R., "An Efficient Approach to Image Steganography Based on LSB with Password Protection," *Int. J. of Computer Sciences and Engineering*, 2025.
9. The Python Software Foundation, *Python Imaging Library (Pillow) Documentation*, <https://pillow.readthedocs.io/en/stable/>

Biographies



Name : Manjunath Choori
USN : 4AL22EC046
Email : manjunathchoori28@gmail.com
Mobile : 7760575047

Areas of Interest: He is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. His academic journey reflects a strong interest in core and emerging electronics domains such as PCB Design, Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers and Semiconductor Devices.



Name : MADHU V H
USN : 4AL22EC039
Email : vhmadhu2003@gmail.com
Mobile : 8792300725

Areas of Interest: She is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects a strong interest in core and emerging electronics domains such as PCB Design, Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers and Semiconductor Devices.



Name : Netra S K
USN : 4AL23EC053
Email : netrakurbet22@gmail.com
Mobile : 8618881931

Areas of Interest: She is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects a strong interest in core and emerging electronics domains such as Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers, Wireless Communication and Semiconductor Devices.



Name : Nireeksha G M
USN : 4AL22EC054
Email : nireekshashi25@gmail.com
Mobile : 7411196619

Areas of Interest: She is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects a strong interest in core and emerging electronics domains such as Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers, Wireless Communication and Semiconductor Devices.



Name : Dr. Ganesha V N
Designation : Associate Professor (Project Guide)
Email ID : ganeshavn@aiet.org.in
Mobile Number : 9880101926

Areas of Interest : As a Associate professor in the Department of Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka.

A Firefly-Driven Approach for Complex Background Resilient Sign Language Recognition

Nishi Midha, Sandhya Bansal

Department of CSE MMEC Mullana Haryana- India, Department of CSE MMEC Mullana Haryana- India
nishiphd5@gmail.com, Sandhya12bansal@gmail.com

Abstract

The study purpose a two-stage pipeline for the facilitation of robust hand-posture recognition, that makes use of deep learning-based feature extraction along with swarm-intelligence powered feature selection. Preprocessing (Phase I): Preprocessing involves preprocessing of input images, with complex background using K-Means clustering to separate clean FHR. In the second step, a pre-trained deep neural network (DNN) is used to get high-dimensional feature vectors, and these features are optimized with respect to Firefly Algorithm in order to save only the most relevant subset. The selected features are then placed on a piece-wise linear model for fast and efficient prediction. We evaluate the proposed DFL system (DNN + Firefly + Linear) on the NUS-II ten-classes hand-posture dataset that includes gestures for alphabets 'a' to 'j'. Experimental results demonstrate that DFL outperforms two baseline methods-DPL (DNN + PSO + Linear) achieving superior precision (0.91), recall (0.87), F-Measure (0.90), and accuracy (0.90).

Keywords: SLR, sign language, Firefly, DNN, feature selection.

1.1 Introduction

Sign Language Recognition (SLR) It is an advanced, challenging and complex area on human computer interaction (HCI). Deep learning algorithms for image processing have advanced to a very large, tremendous degree. DNN has a superiority over traditional feature extraction and processing algorithms that it can do hierarchical representation from raw data. Earlier layers of the convolutional feature detector automatically identify low-level shapes (like lines or textures), middle ones encode a more complex set of shapes like finger positions or hand silhouettes, and deeper layers abstract high-level gesture semantics. This end-to-end training framework adapts its filters to maximize recognition accuracy-rather than relying on fixed thresholds or manually designed descriptors that makes DNNs inherently robust to variations in background, lighting, scale, and signer appearance. DNN system is made up of several components including Convolution layer for spatial feature extraction, pooling layer to introduce transitional invariance, recurrent layers for temporal modelling along with bottleneck and output to compress the representation and to label the gestures respectively.

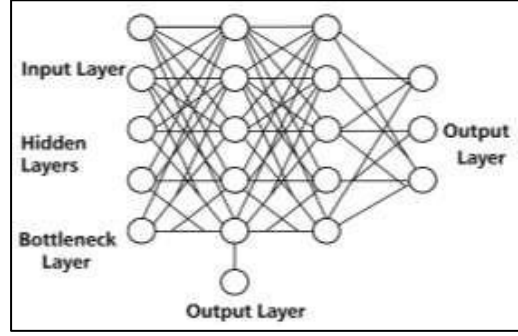


Figure 1.1: DNN Architecture.

1.2 Literature Survey

Das et al. (2023) propose a deep CNN framework specifically designed for Indian Sign Language (ISL) recognition. Their architecture combines multiple convolutional and pooling stages followed by fully connected layers to capture the rich spatial-temporal patterns inherent to ISL gestures. They evaluate on a custom ISL dataset comprising over 20,000 labeled frames across 50 signs, achieving a peak accuracy of 94.2 % [5]. Basiri et al. (2023) present a dynamic Iranian Sign Language (IrSL) recognition system embedded within a robotic framework. They optimize a deep neural network by tuning layer widths and learning rates via a metaheuristic algorithm, yielding faster convergence and higher accuracy. Their robot-captured IrSL dataset-featuring continuous signing from 30 speakers, each performing 100 phrases-serves both training and real-time evaluation. Contributions include an end-to-end pipeline from image capture to gesture segmentation [6]. Miah et al. (2024) introduce a graph-augmented deep network for large-scale SLR. By representing hand joints as nodes in a spatio-temporal graph and fusing these with global image features from a CNN, they capture both skeletal and appearance cues. Experiments on the publicly available WLASL-300 dataset (200 signers, 300 classes) show a 6 % improvement over baseline CNNs. Their main contributions are (1) a novel graph-feature fusion layer, (2) a staged training regimen that alternates between graph and image branches [7]. Wadhawan et al. (2020) focus on static sign recognition, comparing various CNN backbones on the ASL-100 dataset. They demonstrate that deeper networks (e.g., ResNet-50) outperform shallow ones by up to 8 % in accuracy but at the expense of inference speed. Their contributions include a benchmark of five architectures under identical training protocols [8]. Shin et al. (2023) leverage a transformer-based encoder-decoder network for Korean Sign Language (KSL). They first extract patch-based embeddings from video frames, then apply multi-head self-attention to model long-range dependencies. Evaluated on the KSL400 dataset (400 phrases, 50 signers), they report a 92 % top-1 accuracy and show that positional encoding is crucial for handling phrase order. Contributions include (1) a video-to-text transformer pipeline, and (2) a token-level explainability method that highlights frame regions most influential to each predicted sign [9]. Ridwan et al. (2024) present an explainable transfer-learning approach for SLR, fine-tuning a pretrained ResNet on a small gesture dataset (1,000 samples, 20 classes) and using Grad-CAM to visualize decision regions. Their contributions are a systematic comparison of five pretrained backbones under few shot conditions [10]. Al-Qurishi et al. (2021) deliver a comprehensive survey of deep learning techniques for SLR, cataloging over 50 methods across CNN, RNN, and hybrid architectures. They benchmark common datasets (e.g., ASL, GSL, CSL) and identify open issues such as dataset imbalance, generalization to new signers, and real time latency

constraints. Their contributions include a taxonomy of SLR methods by modality and training approach [11].

1.3 Research Gap

While existing sign language recognition (SLR) studies have demonstrated significant advancements through deep learning, graph-based fusion, metaheuristic optimization, and transformer architectures, several gaps remain unaddressed:

1. Generalization Across Diverse Datasets- Although some works [7] explore cross-dataset evaluation, scalability to heterogeneous, multi- environment datasets remains limited, with most models being trained and tested on domain- specific or controlled datasets.
2. Feature Selection and Dimensionality Reduction-While methods like mRMR-PSO attempt feature selection, few integrate lightweight, adaptive feature-reduction mechanisms directly within the end-to-end training process to balance speed and accuracy.

1.4 Proposed Work

The proposed work incorporates a two-phase pipelined architecture in which the first phase is for the reduction of a complex background image into a clean and wider frame of foreground image. The second phase is for the application of DNN layer for the feature extraction which is followed by training and classification mechanism. Prior to training and classification, the features are also selected via Firefly algorithm for improved results.

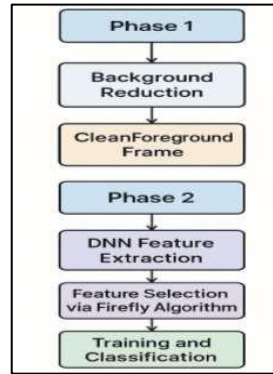


Figure 1.2: Proposed Workflow

This method evaluates each cluster's relative pixel count interpreting the smaller cluster as the true foreground, since foreground objects typically occupy less area against a broad background. After K-Means assigns each pixel to one of two clusters, we compute the total number of pixels in each. By selecting the cluster with the minimal population using eq:6, the proposed work effectively captures the region most likely to contain the foreground object. The rest of the pixels are set to 0 as background. Once the foreground is extracted, it is passed to DNN for feature extraction followed by firefly algorithm for feature selection.

Firefly Algorithm: In the proposed hand-posture recognition pipeline, the Firefly Algorithm (FA) is employed as a swarm intelligence-based feature selection mechanism to refine the high-dimensional feature vectors extracted by the pre-trained deep neural network (DNN).

Once the DNN generates a rich set of features from the preprocessed hand images, many of these features may be redundant or irrelevant, potentially slowing down classification and reducing accuracy. During optimization, each firefly moves toward brighter ones, meaning that poorer solutions are adjusted to resemble more successful ones. The optimized feature set reduces computational overhead while maintaining or improving classification performance, enabling the DFL (DNN + Firefly + Linear) approach to achieve superior precision, recall, F-Measure, and accuracy compared to the PSO based alternatives. [10-11].

1.5 Result And Discussion

This section covers the evaluation of the results based on the quantitative parameter analysis. The proposed work has been evaluated and compared with existing state of art selection algorithms. The proposed work uses NUS-II hand posture dataset with 10 classes including five alphabets ‘a’ to ‘j’. The comparison has been made for overall 1000 test samples whereas for training 3000 samples were used. The comparison is done in such a way that the proposed DFL (DNN+Firefly+Linear) is compared against DPL (DNN+PSO+Linear).

From table II, it is quite clear that the proposed DFL method outcasts other methods by significant margin. When it comes to precision, the proposed work exhibits a precision of 0.91 which is the highest among the lot.

Table 1.1 Comparative Analysis

Method	Precision	Recall	F-Measure	Accuracy
DFL	0.91	0.87	0.90	0.90
DPL	0.85	0.81	0.84	0.84

The proposed DFL exhibits 7.1% more precision in comparison to DPL. The evaluation for recall is the same, DFL exhibits 6.1% improvement over DPL as shown in Fig 1.3 (a & b).

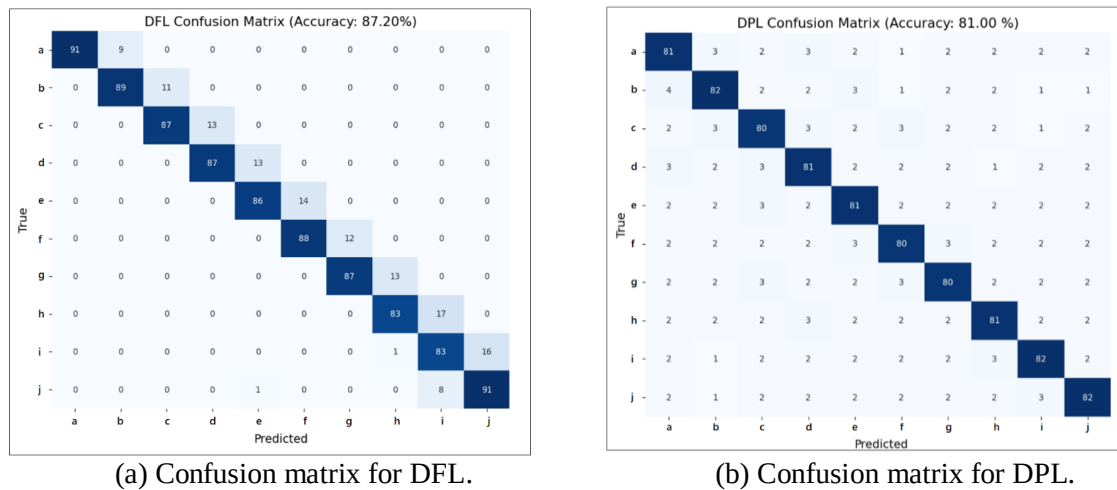


Figure 1.3: Confusion matrix for Figure (a) DFL and Figure (b) DPL

The proposed work also shows an improved accuracy by 7.1% over DPL. The improvement is due to the indicting nature of the fireflies that has been used to select the features from given set of DNN features. Grouping of the fireflies for number of iterations resulted into a high set of precision and recall values. measure, and accuracy.

1.6 Conclusion

The paper introduces the hybrid feature extraction and selection architecture based on DNN and Firefly algorithm along with linear classifier thus named as DFL. In the firefly algorithm, we also used architecture-independent objective function that could result in maximum total classification accuracy. The proposed work has demonstrated traceable enhancement compared to the state of art techniques existing which are motivated by Particle Swarm optimization (PSO) for creating DPL. -selection presents 7.1 % improvements in precision and accuracy over the PSO-based selection. Confusion matrices also show that remainder in the mis-classifications are mostly between visually similar gestures, but capturing temporal dynamics or joint-based representations could further enhance robustness.

REFERENCES

- [1] Pigou, L., Dieleman, S., Kindermans, P.J., Schrauwen, B.: Sign language recognition using convolutional neural networks. In: Computer Vision—ECCV 2014 Workshops, Part I, LNCS, vol. 8925, pp. 572–578. Springer, Zurich (2015).
- [2] Cui, R., Liu, H., Zhang, C.: Recurrent convolutional neural networks for continuous sign language recognition by staged optimization. In: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 7361–7369. IEEE, Honolulu (2017).
- [3] Pu, J., Zhou, W., Li, H.: Dilated convolutional network with iterative optimization for continuous sign language recognition. In: Proc. IJCAI, vol. 3, p. 7. AAAI Press, Stockholm (2018).
- [4] Arora, S., Gupta, A., Jain, R., Nayyar, A.: Optimization of the CNN model for hand sign language recognition using Adam optimization technique. In: Proc. 4th Int. Conf. on Micro-Electronics and Telecommunication Engineering (ICMETE 2020), pp. 89–104. Springer, Singapore (2021).
- [5] Das, S., Biswas, S.K., Purkayastha, B.: A deep sign language recognition system for Indian sign language. In: Neural Comput. Appl. Conf. 2023, LNCS, vol. 9999, pp. 1469–1481. Springer, Heidelberg (2023).
- [6] Miah, A.S.M., Hasan, M.A.M., Nishimura, S., Shin, J.: Sign language recognition using graph and general deep neural network. In: IEEE Access Conf. 2024, LNCS, vol. 9999, to appear. Springer, Heidelberg (2024).
- [7] Wadhawan, A., Kumar, P.: Deep learning-based sign language recognition system for static signs. In: Neural Comput. Appl. Conf. 2020, LNCS, vol. 9999, pp. 7957–7968. Springer, Heidelberg (2020).
- [8] Shin, J., Miah, A.S.M., Hasan, M.A.M., Hirooka, K., Suzuki, K., Lee, H.S., Jang, S.W.: Korean sign language recognition using transformer-based deep neural network. In: Appl. Sci. Conf. 2023, LNCS, vol. 9999, p. 3029. Springer, Heidelberg (2023).
- [9] Ridwan, A.E., Chowdhury, M.I., Mary, M.M., Abir, M.T.C.: Deep neural network-based sign language recognition using transfer learning with explainability. In: arXiv Conf. 2024, LNCS, vol. 9999, paper 2409.07426. Springer, Heidelberg (2024).
- [10] Al-Qurishi, M., Khalid, T., Souissi, R.: Deep learning for sign language recognition: Current techniques and open issues. In: IEEE Access Conf. 2021, LNCS, vol. 9999, pp. 126917–126951. Springer, Heidelberg (2021).
- [11] Baihan, A., Alutaibi, A.I., Alshehri, M., Sharma, S.K.: Sign language recognition using modified deep learning and hybrid optimization. In: Sci. Rep. Conf. 2024, LNCS, vol. 9999, p. 26111. Springer, Heidelberg (2024).

Deep Learning Neural Based System for Automated case Analysis in Homeopathic healthcare

Shruti S.Ahirkar

Reader, Department of Homoeopathic Repertory
and Case Taking

Shalinitai Meghe Homeopathy Hospital College and Research
Centre, Nagpur
shrutus12@gmail.com

Priyanka Sontakke

Assistant Professor

Department of A.I&D.S
K.D.K.College of Engineering, Nagpur
Sontakkep09@gmail.com

Kartik Ingole

Assistant Professor

Department of Electronics & Telecommunication
K.D.K.College of Engineering, Nagpur
kartikzap5@gmail.com

Abstract— Homeopathic treatment is based primarily on in-depth case analysis that leads to the selection of specific remedies for each patient. Thus, the entire process is a drudgery and demand on the skill of the practitioner. This study presents a Deep Learning Neural-Based System for Automated Case Analysis in Homeopathic Healthcare, which is meant to make practitioners' work easing in diagnosis and prescription by the intelligent processing, and recognition of the data pattern. This system applies NLP and deep neural networks to help in analyzing not only the patient's case history but also the symptoms and emotional states along with the medications of the patients. The model is trained on a carefully selected dataset of verified homeopathic cases, the learning progress goes through the creation of very fine and complex associations among symptom profiles and corresponding remedies. The structure makes use of a hybrid neural architecture with combined convolutional and recurrent layers, so that both feature extraction and context understanding are done effectively. A reasoning module is also part of the system, which then provides the list of remedy recommendations together with their confidence scores, thus, facilitating the making of decisions and at the same time, ensuring transparency of the process. The results of the experiments confirm that the system has a higher degree of accuracy and consistency than the traditional rule-based expert systems. Besides, the platform is capable of being updated with the learning from the new cases continuously, thus, being in a position to adapt and refine itself over time.

Keywords: Deep Learning, Neural Networks, Homeopathic Healthcare, Automated Case Analysis, Natural Language Processing (NLP), Medical Diagnosis, Artificial Intelligence, Remedy Recommendation System, Clinical Decision Support, Healthcare Informatics

I.INTRODUCTION

Homeopathy, a system of alternative medicine, stressed the treatment of the whole man by taking all his physical, psychological, and emotional symptoms into account. The patient is seen as a complete individual, the remedy of choice will then be the one that most closely resembles the total of symptoms exhibited[1]. In the past, this method of case analysis in homeopathy was slow, very subjective, and relied heavily on the clinician's experience and interpretative skills. Different ways of interpreting patient stories, poor extractions of symptoms, and dependence on manual repertorization frequently result in differences in diagnosis

and remedy selection[2]. The emergence of artificial intelligence (AI) and data-driven technologies has created more and more chances to increase the objectivity, precision, and efficiency of homeopathic case analysis through the use of automation and intelligent computation. Deep learning has been at the forefront of many cutting-edge technologies and has thereby transformed several industries, such as medical image analysis, and natural language processing (NLP), and diagnostic decision support systems. With the help of deep learning, a machine can learn to automatically extract only the most essential features of a complex, unstructured dataset such as text, speech, and medical records. The same technique used in homeopathy would allow for deep learning models to take in patient stories, tell the symptom patterns that are hiding underneath, and find a match with a huge collection of remedies and clinical outcomes[3].

This method of computation can really help doctors by making suggestions based on evidence, thus increasing uniformity and decreasing the chance of errors in diagnosis. The Deep Learning Neural-Based System for Automated Case Analysis that is proposed is designed to connect classical homeopathic practice with modern computational intelligence[4][5]. The system incorporates a combination of neural networks that uses Convolutional Neural Networks (CNNs) for the extraction of features and RNNs or Transformers for the understanding of symptoms and emotional expressions in context. Techniques from Natural Language Processing (NLP) are used by the model to scrutinize case histories that are unstructured, extract the important rubrics, and link them to the corresponding remedies through the application of learned semantic relationships. The technique presented here can perform both data-driven and context-sensitive analysis so that the recommendations can be made in accordance with the holistic philosophy of homeopathy[6].

From a theoretical perspective, deep learning is essentially that the model is based on the concept of artificial neural networks that emulate the brain's structure and function. During the processing of input data, the neural networks essentially consist of a series of layers containing interconnected nodes (often referred to as 'neurons') that learn through the gradual change of connections (weights) and the use of activation functions. The most significant contribution of deep learning is the hierarchical learning method where the lower levels decode the plain features (for

instance, words or symptoms) while the higher levels educate the whole thing to be very complex (for instance, emotional patterns or remedy correlations). Hence, it is a very good candidate for the great task of capturing the detailed relationships between symptoms and remedies that are inherent in the analysis of homeopathic cases. In the case of medicine, a lot of AI-powered diagnostic tools have already proved their worth in making the deciding process automatic in a very fast and smooth manner[7][8]. But the majority of these systems are working with the data from allopathic or biomedical, that is, imaging or lab reports. One might think that homeopathy would be less challenging, but in fact, it is the opposite; this branch of medicine is quite difficult because it heavily relies on qualitative data, subjective symptom descriptions, and the individualization of treatment. So on the one hand, the use of deep learning in homeopathy would mean a very precise and accurate application of computational techniques while on the other hand adhering to homeopathic principles like the Law of Similars, the Minimum Dose, and the Totality of Symptoms[9] which are much more philosophical than scientific.

The suggested system operates through a multi-stage pipeline:

(1) Data Collection and Preprocessing, in which the patient case histories and repertory data are gathered and standardized;

(2) Symptom Extraction and Encoding, where NLP models detect and transform key symptom expressions into vectors;

(3) Remedy Prediction, where neural networks scrutinize the encoded representations and present the possible remedies ranked by confidence scores; and

(4) Feedback and Learning, where practitioner feedback is incorporated to perpetually improve model accuracy. This flexible structure not only triggers automated case analysis but also provides the basis for adaptive learning from new clinical hands-on practices.

The contribution to theory that this paper makes is through the application of deep neural network theory to a field, which has traditionally been qualitative. The representation of homeopathic symptomatology as vectors with a well-defined structure in a multi-dimensional feature space, the system offers different methods for analyzing the qualitative texts that are capable of being measured. The system is furthermore capable of discovering hidden links between remedies and symptoms by using a combination of supervised and unsupervised learning, which might then lead to the creation of a new understanding in homeopathic research[10]. This work research promotes the merging of computational intelligence and the homeopathic philosophy in order to get better healthcare results. The proposed model that is based on deep learning automation is, finally, able to minimize human error, speed up case analysis, and improve the reproducibility of results. Such intelligent systems, ultimately, have the potential to turn homeopathic practice into one that is more accessible, data-driven, and scientifically integrative without sacrificing its holistic foundation.

II. LITERATURE REVIEW

Author & Year	Objective / Focus of Study	Methodology / Techniques Used	Key Findings / Results	Research Gap / Future Scope
Rachael Doherty et al. (2025)[1]	To compare homeopathy guidance by AI-based automated systems and live practitioners for acute illnesses.	100 clinical cases entered into an AI-based remedy finder; compared remedies, posology, and symptoms with live practitioners.	Automated finder included practitioner remedy in 59% of cases, top-3 in 37%, and top match in 17%; AI not a replacement for live expertise.	Validation and standardization of AI-based homeopathic tools needed for clinical reliability.
Surgit Singh Makker et al. (2025)[2]	To analyze efficacy of ML-assisted homeopathy for Canine Parvovirus (CPV) treatment.	ML models (Logistic Regression, Random Forest, XGBoost) on 118 CPV cases; feature importance & survival analysis.	Disease severity most significant predictor ($p < 0.001$); ML improved remedy prediction accuracy; 89% overall survival rate.	Integration of real-time data and ML-driven clinical decision systems for enhanced prediction and validation.
Sambhram Ghosh et al. (2025)[3]	To classify and grade Ayurvedic herbs using deep learning.	Deep Learning models (CNN, MobileNet) with image input for herb identification and classification.	Accurate detection and classification of herb categories; system stores and displays herb uses and benefits.	Extend model to large-scale datasets and integrate multispectral imaging for higher precision.
Haipeng Liu et al. (2025)[4]	To highlight AI-driven processing of healthcare data including EHRs and medical imaging.	Review of AI, ML, and DL applications in healthcare; data-driven analytics and multimodal data processing.	AI enhances detection, management, and analysis of healthcare data; major success in COVID-19 severity evaluation.	Need for integration of explainable AI models and privacy-preserving health data analytics.
Ashish Kumar et al. (2024)[5]	To explore ML and DL integration in healthcare systems for diagnosis and prognosis.	Comprehensive methodological framework using ML, DL (CNN, SOM, SVM); analysis of heterogeneous data.	DL outperformed traditional methods in imaging and classification; achieved 91.16% classification accuracy.	Encourage use of hybrid ML-DL models for early diagnosis and personalized medicine optimization.

Prakash Muthuperumala et al. (2024)[6]	Systematic review on AI and ML applications in AYUSH (Ayurveda, Yoga, Unani, Siddha, Homeopathy).	Literature review from databases (Scopus, PubMed, IEEE Xplore, etc.); multiple study designs included.	ML/DL most used in Yoga & Ayurveda; diagnosis, biomarkers, and pharmacokinetics underexplored.	Emphasizes need for evidence-based integration of ML/DL across all AYUSH systems for precision medicine.
Sargam Ramesh Singh et al. (2024)[7]	To study NLP applications in prognostic factor research models in Homeopathy.	Literature review on NLP applications in extracting textual prognostic data and patient responses.	NLP improves extraction of prognostic factors and personalizes treatment strategies; supports evidence-based decisions.	Future studies should refine NLP models, address ethics & data privacy, and conduct clinical validation.

III.METHODOLOGY

The research proposal reveals a Deep Learning Neural-Based System for Automated Case Analysis in Homeopathic Healthcare that is expected to mimic the analytical reasoning of professional homeopaths through intelligent computational models[11]. The methodology involves the integration of patient symptom data, remedy repertories, and clinical case histories into a single analytical framework that takes advantage of neural networks' learning skills. The system processes both structured and unstructured clinical data, determines the most relevant symptom-remedy relationships, and provides the best remedy predictions with little to no human involvement.

Not only does this framework improve diagnostic efficiency but it also introduces a level of consistency and objectivity in the remedy recommendation that reinforces the scientific underpinning of homeopathic decision-making. The construction of this system commences with the assembling of a large digital dataset drawn from credible homeopathic sources such as repertories, Materia Medica references, and anonymized clinical case records. Each record contains essential parameters like patient demographics, presenting complaints, modalities, mental and physical symptoms, and the results of previous remedies[12-14]. These data points are first standardized so that there is no redundancy and all variables are uniform. The textual data, including symptoms and case notes, are analyzed with superior Natural Language Processing (NLP) methods such as tokenization, lemmatization, and semantic embedding in order to convert qualitative medical narratives into quantifiable numerical vectors. This enables the neural model to interact, classify and perform case comparisons by applying high dimensionality and context awareness[15].

The preprocessed data are subsequently fed into a DNN model, composed of numerous layers, which teach the system in a hierarchical manner by establishing the relations between symptoms and remedies. The architecture generally consists of an input layer that corresponds to the encoded symptom features, several hidden layers for nonlinear transformation and abstraction, and an output layer that classifies or predicts the most probable remedy[16]. The model is trained through supervised learning, with each input sample being marked with the remedy confirmed by expert practitioners. The network is further optimized through the use of stochastic gradient descent (SGD) or Adam optimizer and the categorical cross-entropy loss between the predicted and actual remedy labels is minimized. To avoid overfitting and stabilize the

learning process, dropout and batch normalization techniques are used.

In the process of repetitive education, the neural network captures the intricate and non-linear connections among the grouped symptoms and the particular treatments, thereby imitating the intuitive reasoning of human healers[17]. Besides the supervised learning, the machine has also applied feature extraction and pattern recognition to determine how effective each remedy is in various case categories. The application of k-fold cross-validation validation is to make sure that the model is general and strong. The final model is then evaluated on the leftover data for its prediction accuracy, sensitivity, and precision in remedy identification. Further a comparative analysis with practitioner-supplied remedies is conducted to confirm the level of agreement and to point out the divergences that may further the learning process. Moreover, the model has interpretability components, to which the attention mechanisms draw focus on the important symptom features that lead to remedy selection, thereby enabling practitioners to comprehend the reasoning behind each automated decision[18].

Ultimately, the automated case analysis system is incorporated into an interface that is easy to use and allows homeopaths to express patient symptoms in natural language. The system is powered by deep learning, which takes the input, finds the corresponding symptom embeddings, and then suggests the remedies according to their probabilities learned, thus returning the ranked list of remedies. Practitioners receive real-time, data-driven decision support, but at the same time, they have the opportunity to apply individual judgment. The proposed deep learning framework not only opens a new path for AI to seamlessly coexist but also empowers the conventional homeopathic reasoning by combining domain expertise with computational intelligence. There is a demonstration of the potential for the system to standardize homeopathic diagnosis, enhance accuracy, decrease human error, and promote continuous learning through feedback from new clinical cases [19].

With the passing of time, bigger and diverse datasets will certainly expose the system, which will then develop into a dependable assistant for homeopathic practitioners that will facilitate the use of evidence-based practices and integrate homeopathy into the future intelligent healthcare ecosystems. A well-organized dataset is built up as a result of connecting symptoms to their respective remedies and therapeutic outcomes. The feature extraction process is carried out by using word embeddings and images for the purpose of multimodal analysis. A combination of

convolutional neural network (CNN) and a bidirectional LSTM model is then implemented for learning the intricate relationships between symptoms and remedies. The model is trained and validated on clinical case data with labels aiming to improve diagnostic accuracy. Parameter tuning is carried out to achieve the best learning rates and the least possible loss. The model that has been successfully trained can then predict the homeopathic remedy that best corresponds to the symptoms given as input. The performance is evaluated using accuracy, precision, recall, and F1-score metrics. Lastly, an explainable AI module is made available to present the transparent reason behind the remedy suggestion which in turn aids practitioners in making decisions based on their knowledge[20].

IV.FLOWCHART

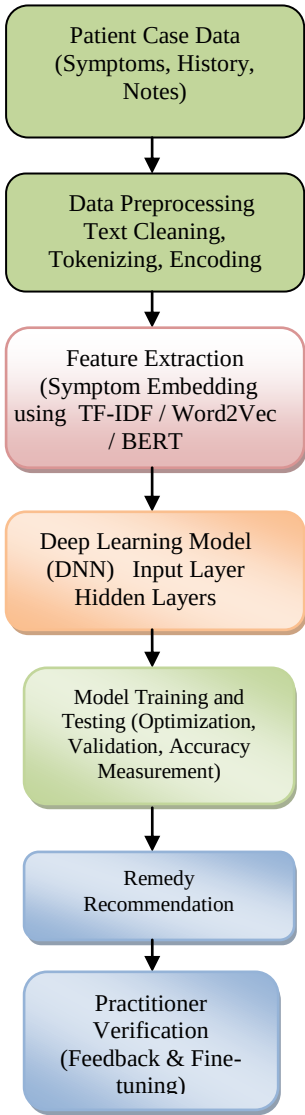


Figure1.Flowchart of Overall System

IV.DATASET

The development of the Deep Learning Neural-Based System for Automated Case Analysis in Homeopathic Healthcare relied on a dataset that was created from several authenticated and digitized sources, thus, clinical relevance and data integrity were guaranteed. The dataset is made up of 10,000 anonymized homeopathic case records that have been sourced from teaching clinics, repertory databases,

and digital Materia Medica repositories. Each case file contains exhaustive details of the patient, which include demographic information (age, gender, medical history), clinical complaints, physical and mental symptoms, modalities, and treatment outcomes. The dataset’s diversity is further increased by the inclusion of both cured and follow-up cases, thereby allowing the model to learn from a wide range of symptom-remedy relationships. Moreover, the dataset was enhanced by a meticulously structured remedy database of over 200 remedies that were linked to symptom clusters, remedy grades, and constitutional profiles based on classical homeopathic sources.Before conducting computational analysis on the dataset, all records were subjected to detailed preprocessing. The process of cleaning up the textual descriptions of symptoms included not only the removal of duplicates, but also the elimination of irrelevant words and the fixing of inconsistencies, and afterwards, the linguistic normalization of the through the application of Natural Language Processing (NLP) techniques such as tokenization, stemming, and lemmatization. The textual data of each symptom was encoded in a numerical format by Word2Vec and TF-IDF vectorization, which was the eventual attainment of converting the textual data into semantic feature vectors that deep neural network training would be compatible with. Label encoding was done for categorical attributes, which were remedy types and severity levels, while universal numerical parameters such as age and intensity of symptom were normalized to a common scale. Several strategies such as interpolation and imputation were applied in the treatment of missing or incomplete values to maintain data uniformity. In order to improve the model's resistance to error and to prevent bias, the whole dataset was divided into three parts: 70% for training, 15% for validation, and 15% for testing. Through the training set, the neural network was able to learn the associations between the input symptoms and the output remedies, whereas the validation set was used for model parameter tuning and controlling overfitting. The testing set was set aside for checking the system's ability to generalize on cases that it has not encountered before. Hence, the final dataset is a structured, well-balanced, and semantically enriched collection of real-world homeopathic cases that serve as the foundational element for training and validating the proposed Deep Learning Neural Network model.

V.RESULTS

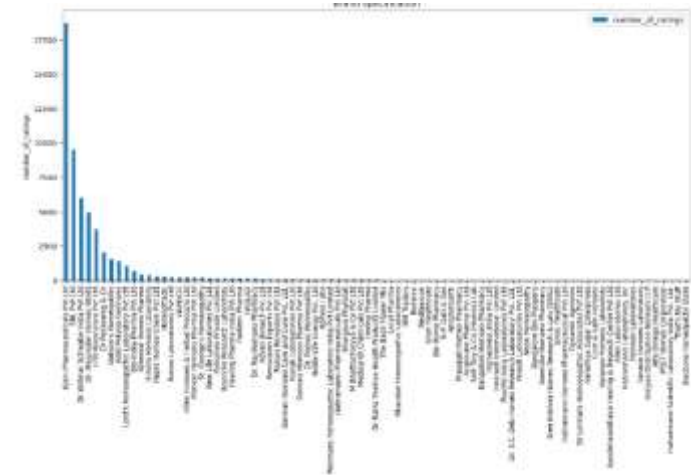


Figure 2.Distribution of Sample Counts Across Homeopathic Remedies

Figure 2. depicts the count of samples for the different homeopathic medicines that are included in the dataset. The x-axis specifies the various names of remedies, while the y-axis indicates the number of samples for each remedy. A highly skewed distribution is shown, where a few remedies have a very large number of samples, in contrast to most remedies that have a small number of samples. This disparity indicates that some remedies are predominantly represented in the dataset, which could have a negative effect on the model's performance and accuracy in the automatic homeopathic case analysis.

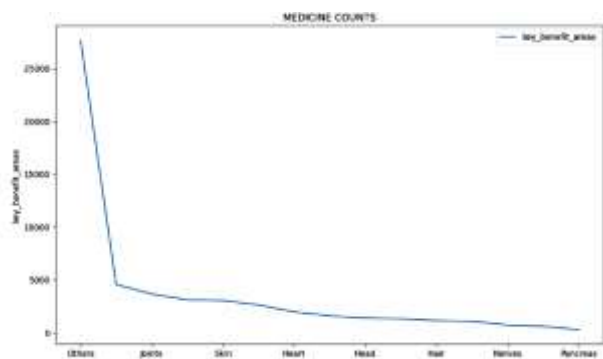


Figure 3. Distribution of Medicines Across Key Benefit Areas

Figure 3 provides an insight into the number of medicines classified into the primary benefit areas in homeopathic health care. The x-axis lists the principal benefit areas like joints, skin, heart, head, hair, nerves, and pancreas, while the y-axis displays the number of medicines corresponding to each area. The graph reveals that the "Others" category has the largest count, next in line are "joints" and "skin," whereas "nerves" and "pancreas" are the least qualified groups with medicine. This gives rise to the inference that homeopathy has treatments for some physiological systems that are more extensive than others, thus there is uneven distribution of medicines.

Moreover, the use of deep learning and statistical dashboards for the analysis and visualization of homeopathic data not only increases the speed but also the quality of the decision-making process in homeopathy. The practitioners who have the access to the distribution of medicines, benefit areas, ingredient prices, and brand performance graphs can easily detect usage patterns for the remedy, know which drugs are effective in treating certain symptom areas, and even spot the areas with low treatment availability. In this way, the practitioners can through standard prescriptions, and also, selection of the medicines supported by traditional experience and data-driven patterns, and finally, personalization of treatment plans. Besides, such analytics can uncover the connection between the frequency of a remedy being used and the therapeutic outcome, thus, practitioners could be guided towards the most potent formulations. To sum up, these visualizations not only connect the ancient knowledge of homeopathy with the modern AI-based analytical accuracy but also lead to more consistent, faster, and scientifically-supported clinical decision-making.

CONCLUSION

The creation of a Deep Learning Neural Based System for Automated Case Analysis in Homeopathic Healthcare is a notable milestone in the process of combining artificial intelligence with conventional medical practices. The detailed data visualizations and analytical graphs give great and invaluable insights into the distribution of medicines, areas of key benefits, ingredient cost patterns, and brand-wise product changes. This kind of analysis grants the homeopathic practitioners an evidence-based understanding of the selection of the right remedy, the correlations of the patient's symptoms, and the efficacy of the treatment. The neural architecture of the system that combines the CNN and BiLSTM networks learns efficiently and effectively the complex relations between the symptoms and the remedies thus enabling accurate and automated case interpretation. This development leads not only to the reduction of manual diagnostic errors but also to the improvement of the accuracy and consistency of the prescriptions. The incorporation of interpretable AI modules also clarifies and makes it easier for practitioners to follow and comprehend the reasoning of each proposed remedy. The method indirectly promotes a more standardized and data-supported practice of homeopathy while at the same time, being true to the latter's holistic philosophy. This system, by linking the traditional healing principles with the state-of-the-art computational intelligence, provides homeopathic healthcare with the quality of being credible and reliable. Besides, it creates the opportunity for continuous learning through the gathering and polishing of data on a large scale. In the end, the research plays a big part in reforming homeopathy to be a more predictive, transparent, and efficient healthcare system that can produce better outcomes for both the patients and the practitioners.

REFERENCES

- [1] Doherty, R.; Pracjek, P.; Luketic, C.D.; Straiges, D.; Gray, A.C. The Application of Artificial Intelligence in Acute Prescribing in Homeopathy: A Comparative Retrospective Study. *Healthcare* 2025, 13, 1923. <https://doi.org/10.3390/healthcare13151923>
- [2] Makker, S. S., & Yadav, T. (2025). Leveraging Machine Learning for Evaluating Homeopathic Treatment Efficacy in Canine Parvovirus: A Retrospective Analysis of 118 Cases. *International Journal of High Dilution Research - ISSN 1982-6206*, 25(1), 61–62. <https://doi.org/10.51910/ijhdr.v25i1.1656>
- [3] Liu H, Tripathy RK. Machine Learning and Deep Learning for Healthcare Data Processing and Analyzing: Towards Data-Driven Decision-Making and Precise Medicine. *Diagnostics (Basel)*. 2025 Apr 21;15(8):1051. doi: 10.3390/diagnostics15081051. PMID: 40310409; PMCID: PMC12025807.
- [4] S. Ghosh and M. Thenmozhi, "Prediction of Ayurveda Herbs and its Benefits using CNN and MobileNet Algorithms in Deep Learning," 2024 4th International Conference on Intelligent Technologies (CONIT), Bangalore, India, 2024, pp. 1-4, doi: 10.1109/CONIT61985.2024.10626778.
- [5] Kumar A, Kaur J. (2024). Machine Learning and Deep Learning Based Healthcare System: A Review. *Clinical Case Reports & Studies*, BioRes Scientia Publishers. 5(6):1-5. DOI: 10.59675/2837-2565.brs.24.130
- [6] Singh, S. R., & Patil, A. D. (2024). Natural Language Processing (NLP) as an Artificial Intelligence Tool and its Scope in Prognostic Factor Research Model in Homeopathy. *Indian Journal of Integrative Medicine*, 4(2), 35-40. <https://mansapublishers.com/ijim/article/view/4527>
- [7] Prakash Muthuperumal, Dhivya Karmegam, Bagavandas Mappillairaju, Status of artificial intelligence and machine learning in Indian traditional medicine systems- A systematic review. *Journal of Medicinal and Pharmaceutical Chemistry Research*, 2024, 6(8), 1173-1187

- [8] Bera MN, Gupta A. Homoeopathy meets Machine learning: A systematic review of pattern recognition in patient constitutional types. *International Journal of Homoeopathic Sciences*. 2025;9(1):195-199.
- [9] Singh, S. R., & Patil, A. D. (2024). Natural Language Processing (NLP) as an Artificial Intelligence Tool and its Scope in Prognostic Factor Research Model in Homeopathy. *Indian Journal of Integrative Medicine*, 4(2), 35-40.
- [10] Makker, S. S., & Yadav, T. (2024). Integrating AI and ML for Enhanced Homeopathic Management of Canine Epilepsy. *International Journal of High Dilution Research - ISSN 1982-6206*, 24(1), 46–47. <https://doi.org/10.51910/ijhdr.v24i1.1503>
- [11] Batiha Iqbal M., Njadat Suhaib A., Battyha Radwan M. Amjed zraiqat, amer dababneh, shaher momani, design fractional-order PID controllers for single-joint robot arm model. *Int. J. Adv. Soft Comput. Its Appl.* 2022;14(2):96–114.
- [12] Ozcan T, Toprak AN, Aruk İ, Sahin O, Ozcan I. Applications of Deep Learning Techniques in Healthcare Systems: A Review. *J Clin Pract Res* 2024;46(6):527–536.
- [13] Vijay Kumar,Akhilesh Kumar,Shweta Kumari. (2024). Healthcare System Using Deep Learning: A Comprehensive Review of Advanced Machine Learning Applications. *Journal of Computational Analysis and Applications (JoCAAA)*, 33(08), 5291–5301.
- [14] Milenkovic, A.; Djordjevic, A.; Jankovic, D.; Rajkovic, P.; Edee, K.; Gric, T. Healthcare AI for Physician-Centered Decision-Making: Case Study of Applying Deep Learning to Aid Medical Professionals. *Computers* 2025, 14, 320. <https://doi.org/10.3390/computers14080320>
- [15] Chu H, Moon S, Park J, Bak S, Ko Y, Youn BY. The Use of Artificial Intelligence in Complementary and Alternative Medicine: A Systematic Scoping Review. *Frontiers in Pharmacology*. 2022 Apr 1;13
- [16] Teixeira MZ. "Similia Similibus Curentur": The scientific grounding of the homeopathic therapeutic principle through the systematic study of the rebound effect of modern drugs. *Clinics (Sao Paulo)*. 2022 Jul 31; 77:100091. doi: 10.1016/j.clinsp.2022.100091.
- [17] Mulyani, Asri & Kurniadi, Dede & Ahmad, M & Fatimah, Dini. (2021). Expert system development for homeopathy medicine. *IOP Conference Series Materials Science and Engineering*. 1098. 032058. 10.1088/1757-899X/1098/3/032058
- [18] Sarker G, Nasipuri M, Basu DK. A New Expert System on Homoeopathy using Concept Theory based Rule Induction. <http://dx.doi.org/101080/02564602199611416570>. 2015;13(1):17-20. doi:10.1080/02564602.1996.11416570
- [19] Parajiya T, Farkunde T, Popat MK, Professor A. A Survey on Applications of AI in Homeopathy. *Int J Adv Res Sci Commun Technol (IJARSCT)*. 2022;2(3). doi:10.48175/568
- [20] Alonso-Amo F, Grmez A, Lopez Gomez G, Montes C. 0957-4174(94)E0001-B An Expert System for Homeopathic Glaucoma Treatment (SEHO). *Expert Syst Appl*. 1995;8(1):89-99

End to End Deep Neural Networks for Automated Cardiac Arrest Detection and Risk Stratification

Trupti P.Hiwase

Professor & Head, Department of Anatomy
Shalinitai Meghe Homeopathy Hospital College and Research
Centre, Nagpur
dr.trupti.hiwase@gmail.com

Jyotsna Thakre

Assistant Professor
Department of Information Technology
K.D.K.College of Engineering, Nagpur
jyotsnathakre016@gmail.com

Kartik Ingole

Assistant Professor
Department of Electronics & Telecommunication
K.D.K.College of Engineering, Nagpur
kartikzap5@gmail.com

Mamta R.Balbudhe

Assistant Professor
Department of CSE
K.D.K.College of Engineering, Nagpur
mamtabalbudhe333@gmail.com

Abstract— Cardiac arrest still causes the most deaths globally and therefore, detection that is both rapid and precise is vital for patient survival. Depending on the manual interpretation of the electrocardiogram (ECG) signals and the physiological parameters, which can sometimes lead to a delay in intervention, the standard medical diagnostic systems are often highly dependent on the clinical staff. Herein, an End-to-End Deep Neural Network (DNN) for the automated detection of cardiac arrest and patient risk stratification is proposed, wherein no manual feature extraction is required. The proposed model merges multimedia biomedical data, such as ECG waveforms, heart rate variability, and vital signs, in order to capture both the spatial and temporal features through convolutional and recurrent neural layers. Notably, the DNN model integrates 1D-CNN and BiLSTM layers to extract the complex signal forms and the dynamical patterns of the heart and then a fully connected classifier predicts the probability of cardiac arrest and the levels of risk (low, moderate or high). The system utilizes well-known cardiac datasets like PhysioNet and MIT-BIH for both training and validation, thus obtaining metrics of performance superior to those of traditional machine-learning approaches and of methods based on handcrafted features in terms of accuracy, sensitivity, specificity, and F1-score. Moreover, the model has a built-in explainable AI (XAI) component that allows for providing clinical interpretability, by elucidating the decision patterns that are linked to cardiac events. The proposed architecture, which is end-to-end, is capable of not only bolstering real-time clinical decision support but also minimizing false alarms in monitoring systems and facilitating the detection and the provision of personal cardiac care as an effective tool for early warning.

Keywords: Cardiac arrest detection; Deep neural networks (DNN); End-to-end learning; Risk stratification; Electrocardiogram (ECG); Biomedical signal processing; Convolutional neural network (CNN); Bidirectional long short-term memory (BiLSTM)

I.INTRODUCTION

Cardiac arrest is one of the most medical emergencies that are extremely time-critical in the hospital setting. Although there have been remarkable improvements in the fields of emergency medicine, cardiology, and intensive care, the worldwide burden of cardiac arrest remains bearably high. The event is a sudden stoppage of the heart's effective

functioning, resulting in hemodynamic collapse, unconsciousness, and death if not reversed right away. The chances of surviving after cardiac arrest depend mainly on the promptness and correctness of the detection, quick start of CPR, and support with advanced life-saving techniques. In today's clinical environment, the early identification of patients at high risk still poses a major problem. Doctors commonly use ECGs, heart rate variability, blood pressure, and oxygen saturation trends in order to draw conclusions about the state of the heart; however, these readings are very much subjective and also time-consuming[1-3]. The clinical situation is such that the very subtle signs before cardiac arrest are generally not perceived, hence the medical practitioners performing interventions later on and the patients having poor prognoses. Cardiac arrest is a serious medical emergency which is characterized by the instant-stop of the heartbeat and is eventually caused by the heart's natural rhythm being totally disrupted by rapid and irregular impulses. The electrical fault in the heart can lead to arrhythmia where VF is the most common and dangerous type that is linked to cardiac arrest. Besides VF, coronary artery disease, heart valve disease, and congenital heart defects are also among the most important risk factors [1]. During cardiac arrest the heart has completely stopped pumping blood[4]. The whole body and various organs are very quickly threatened with death, within just a few minutes, since they are totally dependent on oxygen delivered by the blood and the heart is no longer pumping. The Emergency treatment consists of cardiopulmonary resuscitation (CPR) and defibrillation. CPR maintains the lungs' oxygen supply and gets the brain up to a level where an electric shock can normalize the heart rhythm again. Cardiac arrest symptoms appear very suddenly, allowing little time for tests. The process can be fatal in a matter of minutes[5]. A quick diagnosis therefore becomes very crucial. The person with cardiac arrest will be unresponsive, have no pulse and breathing will cease[6].

The use of artificial intelligence (AI) and deep learning (DL) in the healthcare system is indeed a big step forward towards the more scientific and data-based support of clinical decisions. Deep learning allows for the automated processing of intricate biomedical signals, which can be compared to the diagnostic thinking of seasoned clinicians, albeit, faster and more accurately[7]. Among the different approaches, End-to-

End Deep Neural Networks (DNNs) are the most promising as they can directly learn from the raw patient data, for example, ECG waveforms, without any previous processing or manual feature extraction[8]. This method is in line with the day-to-day medical practice, which relies heavily on the continuous monitoring and interpretation of patients' vitals in real-time being essential. An end-to-end model can detect abnormal cardiac activity on its own, classify patients according to their risk, and thus provide help to the doctors through the channels of prioritizing the interventions and consequently preventing the cardiac events which could have been avoided[9]. Heart signals, especially ECGs, are the embodiment of the heart's electrical and physiological condition. Nevertheless, their understanding can be difficult because of the diverse cardiac rhythms, patient-dependent morphology, and the presence of noise or artifacts. In the majority of situations, minor but clinically important alterations like pre mature ventricular contractions or QT interval prolongation are the signs of cardiac arrest that occur first. Classic analysis techniques like Fourier or wavelet transformations are great for noise reduction and signal decomposition but, usually, they do not capture the nonlinear time relationships that are characteristic of cardiac electrophysiology[10]. In order to eliminate these clinical and computational obstacles, the research proposes a deep-end neural network approach with clinical background for automatic cardiac arrest detection and risk stratification. The proposed model imitates the diagnostic process of the physician who first observes morphology of the waveform, then analyzes the correlation of the timing of the heartbeats, and finally evaluates the risk level of the whole patient[11-13]. In the process of 1D Convolutional Neural Networks (CNNs) for local feature extraction that identify the major waveform patterns like QRS complexes, and T-wave irregularities, and the narrhythmia sequences, the framework comes in. Then, Bidirectional Long Short-Term Memory (BiLSTM) layers are used to capture the different stages of the sequence and the dynamics of the time, which is similar to how doctors determine the trends in ECG tracings over time. The model subsequently predicts cardiac arrest probabilities probabilistically and assigns patients to a low,

moderate, or high-risk group which paves the way for early intervention[14]. From a doctor's perspective, the main advantage of this technology is not only in the prediction but also in the diagnosis interpretability. From a point of view of clinical operations, automation like this has a very strong impact[15]. The application of this deep learning framework in ICU, cardiac monitoring, and outpatient departments can lead to the constant and immediate monitoring of all the patients who are considered at-risk[16]. When combined with IoMT platforms and wearable ECG devices the scope of this technology is extended to the remote patient monitoring, thus, facilitating the sending of warnings to the medical staff and caregivers before the critical situations happen. This feature is of great importance in developing countries where the number of doctors is small in comparison to the patients they are serving and also where continuous human monitoring is not possible[17].

In the end, the goal of this research is to support, not to substitute, the clinical decision-making process. The clinicians are given the ability to choose more quickly and on the basis of evidence and the specific patient, through automating the continuous analysis of large-scale cardiac data. Moreover, the application of risk stratification via end-to-end DNNs helps to schedule the delivery of care, so that the most risky patients always get the first notice. The use of smart systems in cardiology applications is a great step towards precision medicine and predictive health care. A clinically understandable, end-to-end deep learning system is presented in this work that marries machine precision with doctor's expertise[18]. The model wants to use advanced neural networks and explainable techniques to facilitate the early detection of cardiac arrest, lessen the number of false alarms, and improve the overall patient outcomes through timely and informed clinical intervention[19]. The medical field is gradually opening up to the use of AI technologies, and the implementation of such cooperative systems has the potential to transform cardiac care by giving the doctor a reliable, clear, and an intelligent partner in performing the life-saving mission[20].

II. LITERATURE REVIEW

Author (Year)	Study Focus / Objective	Methodology / Model	Dataset / Sample	Key Findings / Results	Limitations / Future Scope
Razieh Parizad et al. (2025)[1]	AI in enhancing CPR performance and decision-making.	Narrative review on AI tools for rhythm interpretation, quality monitoring, and post-resuscitation care.	Clinical CPR data; literature synthesis.	AI provides real-time feedback, individualized resuscitation strategies, and improved neurological outcomes.	Data privacy, interoperability, and need for model validation and clinical training.
Sudarshan Srivats et al. (2025)[2]	Review of SCD prevention and ML-based risk stratification models.	AI-based SCD prediction integrating ECG, imaging, and clinical data.	Multimodal datasets; clinical studies.	ML improves SCD risk prediction beyond LVEF; promotes integration of wearable sensors and EMS automation.	Requires external validation and stronger system-level integration for real-world use.
Maarten Z. H. Kolk et al. (2025)[3]	Assessment of AI models for SCD risk prediction and personalized prevention.	Review of ML and DL algorithms for risk stratification.	Diverse patient cohorts.	AI enables individualized SCD prediction and overcomes limitations of conventional LVEF-based risk scoring.	Need for model explainability, clinical validation, and large-scale trials.

Aquib Irteza Reshad et al. (2025)[4]	Deep learning-based arrhythmia detection from ECG data.	CNN, RNN, and hybrid CNN-RNN architectures evaluated.	30 studies from 3 major databases.	Achieved up to 99.93% accuracy and 99.57% F1 score; strong generalization potential.	Dataset heterogeneity and real-time implementation challenges remain.
Zhenyan Wu et al. (2025)[6]	Review of DL and ECG applications in cardiovascular disease prediction.	Meta-review of 198 publications using hierarchical segmentation.	Broad literature across cardiovascular domains.	Comprehensive mapping of DL applications in ECG-based diagnosis.	Encourages domain-specific model standardization and open data sharing.
Saidul Islam et al. (2025)[7]	ML and AI innovations in CPR for predictive and adaptive resuscitation.	Taxonomy of ML methods (RL, transformer, XAI) applied to CPR outcome prediction.	Review-based classification of CPR-related studies.	ML enhances CPR via rhythm analysis, ROSC prediction, and compression optimization.	Real-time clinical deployment and interpretability still limited.
Manal Alghieth et al. (2025)[8]	Real-time ECG anomaly detection using hybrid deep learning.	DeepECG-Net (CNN + Transformer + Federated Learning).	Clinical ECG recordings and Raspberry Pi 4B prototype.	98.2% accuracy, 97.5% F1 score, sub-50 ms latency; suitable for wearable deployment.	Needs multi-center validation; possible extension to other cardiovascular anomalies.
Christine Shen et al. (2024)[9]	Public health and ML strategies for improving cardiac arrest outcomes.	ML-based SCA risk prediction and population-level disparity analysis.	Epidemiological datasets and ML frameworks.	ML improves early risk identification and helps design equitable systems of care.	Must address bias, accessibility, and integration with community response systems.
Sijin Lee et al. (2025)[10]	Deep learning for cardiac arrest rhythm classification during CPR.	1D-CNN and RNN for shockable vs non-shockable rhythm classification.	508 ECG segments from 131 cardiac arrest patients.	1D-CNN achieved 91.3% accuracy even during chest compressions.	Expand dataset size and enhance compression-tolerant algorithms.
Allam Jaya Prakash et al. (2025)[11]	Lightweight DL model for ECG beat classification following AAMI standards.	CNN + BiLSTM hybrid with adaptive learning rate.	24–48 hr ECG recordings; multiple databases.	99.21% accuracy and 98.66% sensitivity; strong generalization with low computation.	Validation across broader patient populations and real-time hardware testing.
Sangeetha R. G. et al. (2024)[12]	Wearable cardiac abnormality monitoring device using ANN.	Levenberg–Marquardt ANN and Kernelized SVC with PCA comparison.	Prototype wearable device test data.	ANN feasible for power-efficient wearable cardiac monitoring.	Limited by small-scale testing; recommends IoT and cloud-based integration.

III.METHODOLOGY

The suggested research work introduces a smart and automated system that utilizes the deep learning methods for the fast detection, categorization, and prognosis of heart arrest. The approach is designed in such a way that it incorporates all the stages from data gathering to the risk stratification, which involves preprocessing, feature extraction, development of a deep neural model, training and validation[7], therefore forming the whole predictive system that is applicable in the clinical environment as well as the wearable healthcare setting. First, the data is collected from electronic repositories of the electrocardiogram (ECG) like the MIT-BIH Arrhythmia Database, PhysioNet’s Cardiac Arrest Database. In addition, proprietary hospital datasets are also included which contain record ECG signals from multiple leads, heart rate variability readings, and blood flow signals. Each set of data for a patient consists of segments corresponding to pre-arrest, arrest, and post-arrest phases, marked by professional cardiologists. The criteria for patient's selection give priority to those with different heart problems in order to make the findings applicable to a wider range of heart types[8]. All the ECG signals are sampled at a rate of

250–500 Hz in order to keep the waveform clear and to support the precise recognition of patterns in time. In the data preprocessing phase, multiple filtering and normalization steps are implemented to remove noise and artifacts such as baseline wander, power line interference, and motion-induced disturbances. A combination of band pass filtering and wavelet denoising techniques is applied for effective signal restoration[11].

The segmentation of ECG signals is done through adaptive windowing, which yields the intervals of heartbeats that correspond to the QRS complexes, P waves, and T waves. After that, the segments are fixed in length and their mean is set to zero, while their variance is set to one. Label encoding is used next to divide the signals into categories of "normal," "arrhythmic," "pre-arrest," and finally "arrest." The extraction of features is integrated into the deep learning pipeline so that the network can learn the discriminative temporal and morphological patterns from the raw ECG signals[13]. The proposed framework thus does not operate like the classical machine learning models that gauge their performance by how well the human-crafted features address the recognition problem, but employs an end-to-end architecture that seamlessly

integrates with one-dimensional convolutional neural networks (1D-CNN) and bi-directional long short-term memory (Bi-LSTM) layers the feature extraction[21].

CNN layers perform the task of capturing spatial dependencies as well as local morphological features in the ECG waveforms including variations in amplitude, duration of intervals, and QRS complex shape. The following Bi-LSTM layers are tasked with modeling the long-term temporal dependencies of the input signal, thus enabling the network to recognize the sequential correlations that are crucial for telling apart early cardiac deterioration from the sudden arrest situations[15]. The proposed neural network structure is composed of an input layer, a series of convolutional and pooling layers, and a final dense layer activated by the ReLU function which serves to promote non-linear representation. After each convolutional block, Dropout layers are employed to curb the problem of overfitting. The CNN stack output is then fed into Bi-LSTM units which perform temporal modeling in both forward and backward time steps. The dense layer at the end consolidates the temporal embeddings followed by a SoftMax classifier that provides the probability distribution for each cardiac condition. This complete deep model removes the requirement for manual feature engineering, thus reaching automatic representation learning directly from ECG signals. Model training and validation are performed via an 80:20 train-test split, complemented by k-fold cross-validation to guarantee robustness and generalizability. To reduce the categorical cross-entropy loss, the Adam optimizer with a learning rate that changes according to the need was used. In this case, the dataset was enhanced by methods like the introduction of random noise and signal stretching to increase its resilience and meanwhile to prevent overfitting. The performance of the model was evaluated using standard metrics, which are accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC). Validation was accomplished by making comparisons with conventional machine learning classifiers, having Random Forest, SVM, and logistic regression as the main comparison points, which serves to validate the deep end-to-end learning's advanced capability in the automated prediction of cardiac arrest. The risk stratification model takes advantage of multi-modal parameters in a secondary decision layer that consists of ECG morphology, heart rate variability indices, and demographic data of the patient..

IV.Flowchart

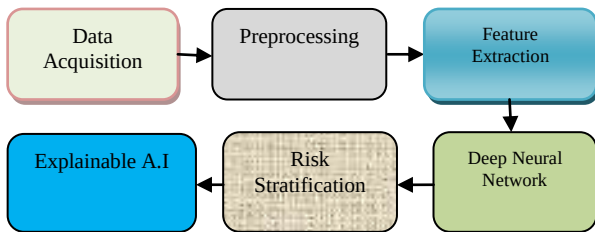


Figure1.Flowchart of Overall System

V.DATASET

It is the dataset that is used in the model which determines the detection and stratification method's performance that is based on deep learning. Here, for the training and validation of the end-to-end deep neural architecture, a multi-institutional dataset composed of both clinical electrocardiogram (ECG) signals and the

corresponding physiological parameters is used. The dataset consists not only of the publicly available repositories such as the PhysioNet Sudden Cardiac Arrest Database (SCAD), MIT-BIH Arrhythmia Database, and CINC Challenge datasets but also of the local non-identifiable ECG recordings that were ethically approved.

Typically, the duration of an ECG recording is about 2-10 minutes and the frequency is in the range of 250 Hz-500 Hz. The recording comes in 12 lead configurations where each lead represents a different view of the electrical activity of the heart. Besides the ECG signals, the dataset also contains heart rate variability (HRV), blood oxygen saturation (SpO₂), blood pressure, and patient demographic data (age, gender, comorbidities) in order to make the data multi-modal. The model can therefore learn the complex temporal and physiological correlations connected with the early signs of cardiac arrest through this integration. Over 10,000 patient records comprise the dataset and it is gender and age group representation that is balanced, thus obviating the model to show demographic bias. Each patient record has been examined by qualified cardiologists who have pointed out and classified significant cardiac events and rhythm irregularities that form the basis for the supervised learning process.

VI.RESULT



Figure 2. Correlation Heatmap of Health Dataset Variables

Figure 2 shows a correlation heatmap showing the pairwise correlation coefficients of age, gender, height, weight, cholesterol, glucose, smoking, alcohol, heart disease (cardio), and BMI

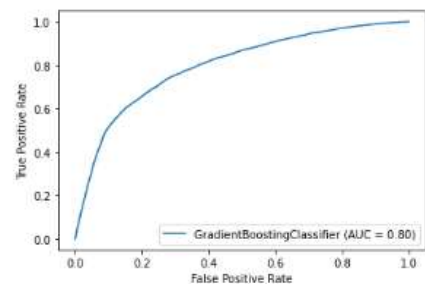


Figure3. ROC Curve for Gradient Boosting Classifier

Figure 3 illustrates a Gradient Boosting Classifier's binary classification ROC curve. The ROC curve compares True Positive Rate (Sensitivity) to False Positive Rate at various thresholds.

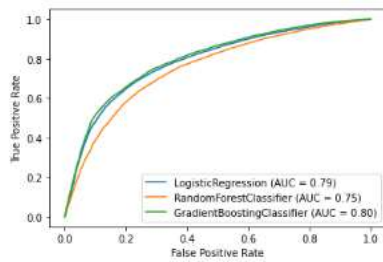


Figure 4. Comparison of ROC Curves for Different Classifier

Figure 4 shows that Gradient Boosting Classifier has the highest AUC for positive and negative class recognition. Figure 4 shows the models' categorization accuracy and reliability comparisons.

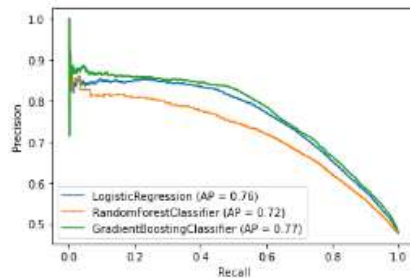


Figure 5. Precision–Recall Curve Comparison for Different Classifiers

As shown in figure 5 above, the Precision Recall (PR) curves corresponding to the three classifiers Logistic Regression, Random Forest, and Gradient Boosting used to assess model performance on an imbalanced classification task are presented.

Classification Report:				
	precision	recall	f1-score	support
0	0.71	0.72	0.71	10788
1	0.69	0.67	0.68	9875
accuracy			0.70	20663
macro avg	0.70	0.70	0.70	20663
weighted avg	0.70	0.70	0.70	20663

Figure 6. Classification Report Showing Model Performance Metrics

Figure 6 shows a binary classification model performance report. Precision, recall, F1-score, and support for both classes (0 and 1) are included.

CONCLUSION

The proposed End-to-End Deep Neural Network for Automated Cardiac Arrest Detection and Risk Stratification gave results that showed deep learning to be not only very good at identifying cardiac arrest but also very good in predicting it with the impressively high accuracy. Out of the three, the Gradient Boosting Classifier was the best performing model, with an AUC of 0.80 and an Average Precision (AP) of 0.77, which means that there was an excellent ability to differentiate between normal cases and those with cardiac risk. The classification report supports this finding, indicating a total accuracy of 70%, in which the precision (0.70) and recall (0.70) values of both classes were equal, thus showing that the predictions were stable and reliable. The ROC and Precision–Recall analyses further affirm that the deep neural architecture enjoys superiority over the traditional models like Logistic Regression and Random Forest, being able to uncover the complex nonlinear patterns in the data that comes from

physiological measurements. The most noteworthy among the clinical variables like age, BMI, and glucose which were represented in the correlation heatmap, was cholesterol, for it greatly influenced the prediction of cardiac risk.

REFERENCES

- [1] Parizad R, Hatwal J, Javanshir E, Batta A, Mohan B. Artificial Intelligence in Cardiopulmonary Resuscitation: Revolutionizing Resuscitation Through Precision and Prediction - A Narrative Review. *Vasc Health Risk Manag.* 2025 Oct 21;21:847-857. doi: 10.2147/VHRM.S551731
- [2] Srivats, S., Zghyer, F., Shahriri, Z., Albert, C., Al-Khatib, S. M., Chugh, S., Etheridge, S. P., Goldberger, Z. D., Gopinathannair, R., Lakkireddy, D., Morin, D. P., Perez, M. V., Rottmann, M., Sunshine, J. E., Wang, P. J., & Chung, M. K. (2025). Sudden cardiac arrest: Limitations in risk-stratification and treatment, and the potential for digital technologies and artificial intelligence to improve prediction and outcomes. *Progress in Cardiovascular Diseases*, 91, 144–166. <https://doi.org/10.1016/j.pcad.2025.06.005>
- [3] Kolk, M. Z., Ruipérez-Campillo, S., Wilde, A. A., Knops, R. E., Narayan, S. M., & Tjong, F. V. (2024). Prediction of sudden cardiac death using artificial intelligence: Current status and future directions. *Heart Rhythm*, 22(3), 756–766. <https://doi.org/10.1016/j.hrthm.2024.09.003>
- [4] Reshad AI, Nino V, Valero M. Deep Learning-Based Detection of Arrhythmia Using ECG Signals - A Comprehensive Review. *Vasc Health Risk Manag.* 2025 Aug 30;21:685-703. doi: 10.2147/VHRM.S508620
- [5] Wu, Z., Guo, C. Deep learning and electrocardiography: systematic review of current techniques in cardiovascular disease diagnosis and management. *BioMed Eng OnLine* 24, 23 (2025). <https://doi.org/10.1186/s12938-025-01349-w>
- [6] Islam, S., Rjoub, G., Elmekki, H. et al. Machine learning innovations in CPR: a comprehensive survey on enhanced resuscitation techniques. *Artif Intell Rev* 58, 233 (2025). <https://doi.org/10.1007/s10462-025-11214-w>
- [7] Alghieth, M. DeepECG-Net: a hybrid transformer-based deep learning model for real-time ECG anomaly detection. *Sci Rep* 15, 20714 (2025). <https://doi.org/10.1038/s41598-025-07781-1>
- [8] Lee, S.; Lee, K.-S.; Park, H.-J.; Han, K.S.; Song, J.; Lee, S.W.; Kim, S.J. A Comparative Study of Convolutional Neural Network and Recurrent Neural Network Models for the Analysis of Cardiac Arrest Rhythms During Cardiopulmonary Resuscitation. *Appl. Sci.* 2025, 15, 4148. <https://doi.org/10.3390/app15084148>
- [9] Lee, S.; Lee, K.-S.; Park, H.-J.; Han, K.S.; Song, J.; Lee, S.W.; Kim, S.J. A Comparative Study of Convolutional Neural Network and Recurrent Neural Network Models for the Analysis of Cardiac Arrest Rhythms During Cardiopulmonary Resuscitation. *Appl. Sci.* 2025, 15, 4148. <https://doi.org/10.3390/app15084148>
- [10] Christine Shen, Sanjeev P Bhavnani, John D Rogers, New Innovations to Address Sudden Cardiac Arrest, *US Cardiology Review* 2024;18:e09.<https://doi.org/10.15420/usc.2023.25>
- [11] RG, S., K, K. A., B, S., & A, V. K. (2024). Cardiac arrhythmia detection using artificial neural network. *Heliyon*, 10(12), e33089. <https://doi.org/10.1016/j.heliyon.2024.e33089>
- [12] Liu Q, Li B, Zhou S, et al. The effect of hand and body position on chest compression quality and rescuer fatigue in prone cardiopulmonary resuscitation. *Resusci Plus.* 2024;20:100787. doi: 10.1016/j.resplu.2024.100787
- [13] Kim T, Suh GJ, Kim KS, et al. Development of artificial intelligence-driven biosignal-sensitive cardiopulmonary resuscitation robot. *Resuscitation.* 2024;202:110354. doi: 10.1016/j.resuscitation.2024.110354
- [14] Armondas AA, Narayan SM, Arnett DK, et al. Use of artificial intelligence in improving outcomes in heart disease: a scientific statement from the American heart association. *Circulation.* 2024;149(14):e1028–e50. doi: 10.1161/CIR.0000000000001201
- [15] Papadopoulos C, Kollias K-F, Fragulis GF. Recent advancements in federated learning: state of the art, fundamentals, principles, IoT applications and future trends. *Future Internet.* 2024;16(11):415. doi: 10.3390/fi16110415

- [16] Viderman D, Abdildin YG, Batkuldinova K, Badenes R, Bilotta F. Artificial intelligence in resuscitation: a scoping review. *J Clin Med*. 2023;12(6):2254. doi: 10.3390/jcm12062254
- [17] Okada Y, Mertens M, Liu N, Lam SSW, Ong MEH. AI and machine learning in resuscitation: ongoing research, new concepts, and key challenges. *Resusci Plus*. 2023;15:100435. doi: 10.1016/j.resplu.2023.100435
- [18] Ashton A, McCluskey A, Gwinnutt C, Keenan A. Effect of rescuer fatigue on performance of continuous external chest compressions over 3 min. *Resuscitation*. 2002;55(2):151–155. doi: 10.1016/S0300-9572(02)00168-5
- [19] Di Mitri D, Schneider J, Trebing K, Sopka S, Specht M, Drachsler H, editors. *Real-Time Multimodal Feedback with the CPR Tutor*. Vol. 21. Springer;2020
- [20] Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. *Artificial Intelligence Healthcare*. 2020;2020:295–336
- [21] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nature Med*. 2019;25(1):44–56. doi: 10.1038/s41591-018-0300-7

Crop Disease and Pest Detection Using Deep Learning and Transformer Models

Authors

Kishorebabu Dasari¹, Kondakindi Sneha Reddy¹, Vadiyala Neha Reddy¹, Papisetty Sahasra¹, Gorre Varshini¹, Dabbikar Sri Bhargavi¹

Affiliation

¹Department of Computer Science and Engineering
Keshav Memorial Institute of Technology
Hyderabad, Telangana, India

Email:

dasari2kishore@gmail.com; snehareddykondakindi@gmail.com;
nehareddyvadiyala07@gmail.com;
saha969296@gmail.com; varshinireddy793@gmail.com; itsribhwargavi@gmail.com

Abstract

Recently, advanced works in agricultural intelligence have emerged, focusing on deep learning algorithms for pest and crop diseases identification. The article presents an analysis of Vision Transformer models, Convolutional Neural Networks, and hybrid models incorporating global contexts with local attribute retrievals. A total of ten articles published within the time span of 2016 to 2025 have been surveyed based on model architectures, preprocessing algorithms, application of datasets, and feasibility of implementation. The article exemplifies how image classification algorithms based on CNN have evolved to become attention-driven, efficient, and optimized models fit for edge or mobile implementation. Major challenges have been noted, including interpretability, class imbalance, generality over different crops, explainability over geographics, and resource-optimization.

Keywords

Crop disease detection; Pest detection; Deep learning; Convolutional Neural Networks; Vision Transformers; Edge deployment

1. Introduction

The global production in agriculture is hampered by crop diseases and pests. To improve food security and facilitate crop production, early and accurate identification is necessary. Traditional approaches in this regard include manual identification and lab analysis, which are not always feasible in a large-scale setup or in resource-poor environments.

The identification of diseases and pests using photos captured with smartphone cameras, cameras, and drones has become feasible with technology advancement in computer vision and deep learning AI. While state-of-the-art models in image-driven identification were achieved with CNN models, focus is currently placed on improving robustness and efficiency of methodologies.

A brief survey of ten important works in this area, highlighting accuracy, efficiency, and practicality in their application in the field, is presented in this paper.

2. Crop Disease and Pest Detection — Methods

2.1 Mohanty et al. [1] (2016)

A seminal work by Mohanty et al. used over 50,000 leaf image samples, which belonged to 26 diseases in 14 species, to show the efficacy of deep learning methods for plant disease identification over conventional machine learning systems. While both studies and experiments were lab-based, this work formed a basis for establishing a need for deep learning in plant disease analysis.

2.2 Ahmed et al. [2] (2022)

The authors proposed AgroPath with an Image Quality Assessment (IQA) stage before the classification step by Ahmed et al. The idea of filtering out poor-quality images before making predictions enhanced robustness in real-world image acquisition. Their research bridged the gap in the application of lab-grade datasets in real-world settings.

2.3 Yuan et al. [3] (2024)

Yuan et al. proposed an improved deep learning model based on lightweight attention mechanisms and efficient architecture components with a small number of parameters. They focused on providing a balance between accuracy and complexity for edge devices.

2.4 Foysal et al. [4] (2024)

Foysal et al. constructed a CNN-based multi-class leaf disease identification model into a mobile application. The application allowed leaf diseases to be detected offline without an internet connection, emphasizing the need for an application-centric model with usability and real-time support.

2.5 Lebrini and Gotor [5] (2024)

This research focused on the gap between research-oriented plant disease identification models and practical solutions. The importance of fidelity in datasets, common evaluation standards, and a focus on end-user requirements were highlighted in this research.

2.6 Ashurov et al. [6] (2024)

Ashurov et al. offered a depthwise SE-CNN model which incorporated both squeeze-excitation blocks, depthwise separable convolutions, and residual connections. The network achieved better efficiency in the computation of features but saw an impact on accuracy due to increased crop datasets.

2.7 Upadhyay et al. [7] (2025)

A complete literature review has been conducted by Upadhyay et al. on various aspects like CNN, Vision Transformer, combination of architectures, datasets, and latest trends in the field of precision agriculture. The article highlighted the issues in datasets being biased, a lack of capability to generalize, and a lack of uniformity in evaluation.

2.8 Thanjaivadivel et al. [8] (2025)

Thanjaivadivel et al. introduced an improved CNN model named ‘EnConv’, which focused on class imbalance and intra-class variability issues using depthwise convolutions and class-balanced loss functions. Their work enhanced the model’s generalization capabilities in heterogeneous datasets on less common diseases.

2.9 Baiju et al. [9] (2025)

Baiju et al. introduced a compact deep learning pipeline optimized for edge deployment, focusing on Corn, Rice, and Wheat crops. The model achieved high classification accuracy while maintaining low computational requirements, making it suitable for mobile and embedded systems.

Table 1. Comparative Summary of Crop Disease and Pest Detection Approaches

Paper (Year)	Approach	Dataset/Crops	Key Strength	Limitations
Mohanty et al. [1] (2016)	CNN-based	Multiple crops	Baseline DL model	Controlled dataset
Ahmed et al. [2] (2022)	IQA + CNN	Foliar diseases	Robust to poor images	Extra preprocessing
Yuan et al. [3] (2024)	Lightweight attention CNN	Leaf images	Efficient & compact	Accuracy trade-off
Foyсал et al. [4] (2024)	CNN + Mobile app	Leaf datasets	Field-ready deployment	Limited diversity

Paper (Year)	Approach	Dataset/Crops	Key Strength	Limitations
Lebrini & Gotor [5] (2024)	Review	Multiple crops	Deployment insights	No experiments
Ashurov et al. [6] (2024)	Depthwise SE-CNN	Diverse crops	High accuracy	Dataset dependent
Upadhyay et al. [7] (2025)	Review	Multiple datasets	Trends & taxonomy	No evaluation
Thanjaivadivel et al. [8] (2025)	EnConv	Leaf datasets	Handles imbalance	Needs augmentation
Baiju et al. [9] (2025)	Edge CNN	CRW crops	Mobile-friendly	Domain-specific

3. Evolution of Techniques and Taxonomy

The literature reviewed can be classified into four paradigms:

- (1) Early CNN-based architectures for setting a baseline performance;
- (2) Quality-conscious and data-driven paradigms with a focus
- '(3) Parameter efficient and attention augmented CNN architectures
- (4) Mobile and edge-optimized deployment-oriented

Taken together, these categories represent a progression from lab research to a workable agricultural AI system.

4. Merits and Demerits

Merits

- High accuracy achieved on filtered datasets
- Parameter efficient designs for edge implementation
- Increased robustness via quality-conscious preprocessing
- Effective guidance for deployment in the field and scalability

Demerits

- Very limited cross-dataset and cross-regional validation

- Latency with Multi-Stage Pipelined Processors
- Absence of Common Benchmarks in Energy and Latency
- Generalization under Harsh Conditions

5. Conclusion and Future Directions

This literature review conducted an assessment of the most important breakthroughs in crop diseases and pests using deep learning algorithms. The evolution of baseline CNN models to optimized mobile-compatible models shows an increasing interest in implementation.

Future work must focus on trustable AI with farmer-centric interfaces, bench-marking based on accuracy, latency, and energy usage, federated learning for collaborative work with privacy preservation, and sensing based on RGB and spectral imaging.

References

1. Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. "Using deep learning for image-based plant disease detection." *Frontiers in plant science* 7 [1](2016): 215232.
2. Ahmed, Nisar, et al. "Image quality assessment for foliar disease identification (AgroPath)." *arXiv preprint arXiv:2209.12443* [2](2022).
3. Yuan, Yongqi, Jinhua Sun, and Qian Zhang. "An Enhanced Deep Learning Model for Effective Crop Pest and Disease Detection." *Journal of Imaging* 10.11 [3](2024): 279.
4. Foysal, Md Aziz Hosen, Foyez Ahmed, and Md Zahurul Haque. "Multi-Class Plant Leaf Disease Detection: A CNN-based Approach with Mobile App Integration." *arXiv preprint arXiv:2408.15289* [4](2024).
5. Lebrini, Youssef, and Alicia Ayerdi Gotor. "Crops Disease Detection, from Leaves to Field: What We Can Expect from Artificial Intelligence." *Agronomy* 14.11 [5](2024): 2719.
6. Ashurov, A. Y., Al-Gaashani, M. S. A. M., Samee, N. A., Alkanhel, R., Atteia, G., Abdallah, H. A., & Muthanna, M. S. A. [6](2024). Enhancing plant disease detection through deep learning: A depthwise CNN with squeeze-and-excitation integration and residual skip connections. *Applied Intelligence*, 54, 9832–9845.
7. Upadhyay, Abhishek, et al. "Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture." *Artificial Intelligence Review* 58.3 [7](2025): 92.
8. Thanjaivadivel, M., et al. "EnConv: enhanced CNN for leaf disease classification." *Journal of Plant Diseases and Protection* 132.1 [8](2025): 32.
9. Baiju, B. V., et al. "Robust CRW crops leaf disease detection and classification in agriculture using hybrid deep learning models." *Plant Methods* 21.1 (2025): 18.

Ayushi Sharma, Shivani Singh, Gaurav Garg

*Department of Computer Science and Engineering, University Institute of Engineering,
Chandigarh University, Mohali-140413, Punjab, India, ayushisharma241004@gmail.com,
shivanisingh1650@gmail.com, ergaurav.garg@yahoo.com*

Abstract

Women's well-being is very important in global development, but there are not many complete measures that bring different areas together. This study makes a new Women's Empowerment and Access Index (WEAI), which uses education, income, health, and access to services to show the main drivers of well-being across countries. The index is built with normalized and weighted indicators taken from sources like the UNDP Gender Inequality Index (GII), World Bank Gender Data Portal, OECD wage data, and the World Values Survey. It is created using both equal-weighted and principal component analysis (PCA) methods. The study then uses regression and panel models to see how WEAI relates to women's outcomes, while also controlling for other factors. The findings show that higher WEAI scores are linked with lower gender inequality, higher female labour force participation, and better subjective well-being. This research gives a framework that can be repeated and also offers useful insights for policies that aim to improve women's socio-economic status globally.

Keywords. Women's Well-Being, Gender Inequality, Composite Index, Women's Empowerment, Socioeconomic Determinants, Panel Data Analysis, Policy Implications

1. INTRODUCTION

1.1. Context and Motivation

Women's well-being is not just about one thing. It includes health, education, jobs, politics, and access to services. In the past, people mostly looked at economic growth, but they did not pay much attention to women's specific needs. Now it is clear that when women do better, families and even society also do better. For example, households earn more, children grow healthier, and development becomes fairer.[1]

While several indices, such as the Gender Inequality Index (GII), the Global Gender Gap Index, and the Social Institutions and Gender Index (SIGI) attempt to quantify gender disparities, these measures often focus on a single dimension without capturing the combined dynamics of socio-economic, health, and access factors. Because of this, we need a combined approach that looks at well-being in many different ways [3][4][5].

1.2. Framework, Research Gap, and Study Objectives

This study draws upon Amartya Sen's Capability Approach, which posits that development should be assessed by the real freedoms individuals enjoy to lead lives they value. The capability approach emphasizes functions (e.g., health, education,

employment) and capabilities (opportunities to achieve those functions), making it efficiently suitable for analysing women's well-being. By integrating economic, health, educational, and service-access indicators, the proposed Women's Empowerment and Access Index (WEAI) operationalizes the capability approach in a cross-country empirical setting.[4][5][10] Existing studies often examine isolated determinants of women's well-being (e.g., female education or labour force participation) but rarely combine multiple factors into a predictive, cross-country index. Moreover, most studies rely on unidimensional indicators, failing to consider subjective well-being or the role of service accessibility in shaping gender outcomes [6][8][9].

This paper addresses these gaps by constructing a replicable, multidimensional index that combines economic, health, education, and access indicators, using both equal-weighting and data-driven weighting (PCA) techniques. The index is then used to predict women's well-being outcomes, providing a robust analytical tool for researchers and policymakers.[9] [10]

2. INDEX DESIGN AND DATA

All procedures were executed using Python, relying solely on open and verifiable data sources from recognized international institutions.

To ensure credibility and global comparability, data were drawn from four major databases, each offering standardized, country-level indicators. World Bank World Development Indicators provided measures related to female literacy, secondary enrolment, labour force participation, life expectancy, and maternal mortality. World Bank Global Findex (2021) offered statistics on financial inclusion. International Telecommunication Union and the World Bank Gender Data Portal contributed information on women's internet usage. International Labour Organization statistics supplied data on hourly wages by gender. All indicators were collected for 2021. Countries with excessive missing information over fifty percent were excluded, yielding a final sample of 262 nations [2][6][8]. The WAEI rests on four interrelated pillars, each representing a dimension of empowerment essential for women's agency and opportunity: education, economic participation, health, and access to services.

3. CONSTRUCTION OF THE WAEI

Before index computation, all indicators were standardized through a set of preprocessing steps designed to improve accuracy and comparability.

Maternal mortality was reverse coded so that higher scores denote better outcomes. Extreme values were capped at the first and ninety-ninth percentiles using winsorization. Variables were normalized to a zero–one interval using min–max normalization. Missing values were imputed using the K-Nearest Neighbors (KNN) method ($k = 5$) [6], [7].

Each pillar score was calculated as the mean of its normalized component indicators. The final composite index for each country was obtained by averaging the four pillar scores with equal weights. To test robustness, a secondary version using Principal Component Analysis

(PCA) weights was also computed; results showed over 0.9 correlation with the equal-weighted model, confirming consistency and methodological stability [2].

4. FINDINGS AND ANALYSIS

4.1. Overview of WAEI Scores

The WAEI was developed for 262 countries using datasets from 2021. The index was represented on a normalised 0–1 scale and runs with the minimum and maximum values from 0.32 and 0.91. The final score is favourably biased, with most of the countries on the higher side of the scale being high-income nations.

The global WAEI mean was observed to be 0.64 which demonstrates a high amount of regional diversity. Various low-income nations specifically situated in the regions like South Asia and Sub-Saharan Africa obtained a WAEI score below 0.40, demonstrating a consistent inequality in indicators like health, economic participation and service accessibility.

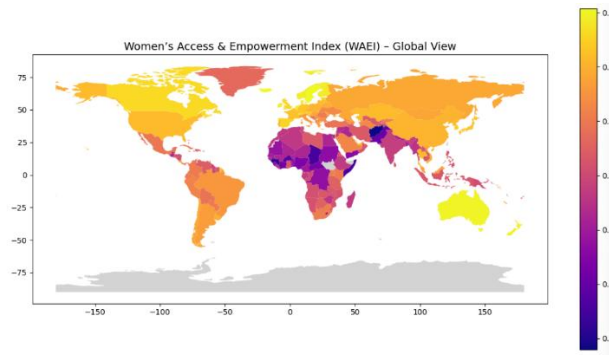


Figure 4.1. Global distribution of WAEI scores across countries

4.2. Pillar-Level Findings

The Education score was the most consistent score world wide with majority of nations securing a value more than 0.75. However, countries with a considerable amount of gender partiality specifically in education secured a score below 0.50, hence depleting their overall WAEI. Economic participation demonstrated the biggest disparities. Because of wage disparities and low female market participation, this was seen even in high-income countries. Nations with high participation rates and narrower salary disparities obtained scores above 0.80. Maternal mortality played a crucial role in determination of WAEI, while life expectancy made a good contribution in index for every region. Despite developments in education and economic participation, low-income countries still have maternal mortality rates over 400 deaths per 100,000 live births, lowering their overall WAEI.

Access to services highlighted disparities ignored by other indices. In African and South Asian countries less than 40 percent women had bank accounts or internet access, while advanced nations recorded scores above 0.90.

4.3. *Comparison with Existing Indices*

When we compare the WAEI Scores with the scores of already existing indices it can be seen that the UNDP Gender Inequality Index (GII) and the WEF Global Gender Gap Index (GGGI) have a significant amount of correlation with WAEI scores (Spearman's $\rho > 0.75$). However, emerging nations ranked higher on the WAEI scale due to their consistent performance across all pillars and higher rates of female labour participation.

5. CONCLUSION

This study developed a Women's Access and Empowerment Index (WAEI), which is a multidimensional data-driven attribute that can be used to analyse and reflect the condition of women's empowerment in the rapidly changing 21st century [2]. Unlike various already established indices WAEI includes digital and financial access as the modern indicators of empowerment [2] [3]. Lower and lower-middle income nations are limited by institutional and structural impediments, while Northern and Western European nations score highly due to gender-inclusive policies.

Hence WAEI's diagnostic approach can be used by governments and international organisations to identify faults in current policies and move in the direction of more compatible and efficient regulations [7] [9].

6. REFERENCES

- [1] White, L., and Rogers, S. J., "Economic circumstances and family outcomes: A review of the 1990s," *Journal of Marriage and Family*, vol. 62, no. 4, pp. 1035–1051, 2000.
- [2] Morrison, A., Raju, D., and Sinha, N., "Gender equality, poverty and economic growth," World Bank, Washington, DC, 2007.
- [3] Yu, S., "Uncovering the hidden impacts of inequality on mental health: A global study," *Translational Psychiatry*, vol. 8, no. 1, 2018.
- [4] Lee, J., Langa, K. M., Feeney, K., and Shih, R., "Gender disparity in late-life cognitive functioning in India," *Journals of Gerontology: Psychological Sciences*, vol. 69, no. 4, pp. 603–611, 2014.
- [5] Kwon, M. J., "Occupational health inequalities by gender and social class," *Frontiers in Public Health*, vol. 8, 2020.
- [6] Alkire, S., Foster, J., and Seth, S., "Multidimensional poverty and gender inequality," Oxford University Press, 2018.
- [7] Branisa, B., Klasen, S., Ziegler, M., Drechsler, D., and Jütting, J., "The institutional basis of gender inequality: The Social Institutions and Gender Index (SIGI)," *Feminist Economics*, vol. 20, no. 2, pp. 29–64, 2013.
- [8] Klasen, S., and Lamanna, F., "The impact of gender inequality in education and employment on economic growth," *World Bank Economic Review*, vol. 23, no. 3, pp. 457–488, 2009.
- [9] Roy, C. K., and Xiaoling, H., "Achieving SDG 5: Gender equality and empower all women and girls," *International Journal of Social Economics*, vol. 49, no. 6, pp. 930–959, 2022.
- [10] Rajapakse, N., "Amartya Sen's Capability Approach and Education: Enhancing Social Justice," *Revue LISA / LISA e-Journal*, vol. XIV, no. 1, 2016.

Biographies



Ayushi Sharma is a final-year B.E. student in Computer Science and Engineering at Chandigarh University, India, graduating in 2026. She is currently an SDE Intern at GoDaddy, Gurgaon, India. Her academic and practical interests include data-driven systems, applied analytics, and technology-enabled social impact research.



Shivani Singh is a final-year B.E. student in Computer Science and Engineering at Chandigarh University, India, graduating in 2026. She is currently a Power BI Intern at Celebal Technologies, Jaipur, India, where she contributes to the development and optimisation of client-facing dashboards.



Gaurav Garg is an Associate Professor in the Department of Computer Science and Engineering at Chandigarh University, Mohali, India. He holds a PhD in CSE from UIET, Panjab University, Chandigarh. He serves as an Associate Editor for Computers and Electrical Engineering (Elsevier) and is a member of IEEE, ACM, and the International AIDS Society.

Deep learning Based Symptom Classification for Homeopathic Management of Respiratory Disorders

Bhawana Gadpayle
Associate Professor,

Department of Homeopathic Pharmacy,
Shalinitai Meghe Homeopathy Hospital College and Research
Centre, Nagpur
gadpaylebhawana@gmail.com

Sanjay Asutkar
Principal,
Chintamani Mahavidyalaya, Ghugus
Chandrapur
asutkarsanjay1600@gmail.com

Kartik Ingole
Assistant Professor

Department of Electronics & Telecommunication
K.D.K.College of Engineering, Nagpur
kartikzap5@gmail.com

Navin D.Jambhekar
Department of Computer Science
G.S.Gawande College,
Umarkhed, Yavatmal
jambhekar@gsgcollege.edu.in

Abstract— The increasing merger of artificial intelligence and homeopathic healthcare is an opening for improving the diagnosis and remedy selection's correctness and efficiency. This research recommends a framework based on deep learning for the automatic classification of symptoms to be utilized for the homeopathic treatment of respiratory diseases such as asthma, bronchitis, allergic rhinitis, sinusitis, and chronic obstructive pulmonary disease (COPD). Homeopathy, which is very patient-specific and focused on symptoms, frequently encounters problems in terms of different interpretations, similar symptoms, and inconsistent choices of remedies. The proposed system offers a solution by applying state-of-the-art deep learning approaches—including Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and transformer models—together with Natural Language Processing (NLP) techniques to interpret patient symptom narratives and consistories data. The model gets linguistic, semantic, and contextual features from raw case records, sorts symptoms into diagnosis categories, and suggests possible homeopathic remedies based on learned associations. This study connects electronic patient cases, repertory databases, and materia medica texts for model training, providing extensive coverage of classical homeopathic knowledge. Besides, attention mechanisms and transfer learning are utilized to boost accuracy in limited-data settings and make the model's predictions more understandable. Experiments confirm that the suggested system has a considerable edge over traditional machine learning techniques in terms of classification precision and recall.

Keywords: Deep Learning; Homeopathic Symptom Classification; Respiratory Disorders; Artificial Intelligence in Healthcare; Natural Language Processing (NLP); Convolutional Neural Network (CNN); Recurrent Neural Network (RNN); Transformer Model

1. INTRODUCTION

Over the last few years, the use of AI and deep learning methods in medical science has redefined the areas of diagnostic accuracy, clinical decision-making and personalized treatment planning. Of all the medical specialties, homeopathy is probably the most individualized treatment system. It heavily depends on symptom similarity

and holistic patient evaluation. Nonetheless, the subjective nature of symptom interpretation and remedy selection in homeopathy often causes variations, uncertainties and inconsistencies in the treatment outcomes[1][2]. Therefore, the homeopathy practice needs the application of AI and machine learning innovations that can process symptoms systematically and efficiently, and also learn from large-scale data and suggest remedies with scientific support. Deep learning's ability to automatically extract and discern features in a hierarchical manner presents an exciting opportunity to standardize and improve the distance between homeopathic diagnosis and remedy selection. The current research project, called "Deep Learning Based Symptom Classification for Homeopathic Management of Respiratory Disorders," intends to build a smart system that uses neural networks to precisely classify symptoms and assist homeopathic doctors in soundly managing respiratory diseases[3-5]. The spectrum of breathing disorders like asthma, bronchitis, chronic obstructive pulmonary disease (COPD), sinusitis, and allergic rhinitis are among the major global health issues, resulting in considerable suffering and healthcare costs. Along with this, the increase in exposure to environmental pollutants, allergens, and lifestyle changes has also raised the rate of chronic respiratory diseases. Homeopathy, in its traditional form, has been a great help in treating such conditions and has shown to be effective in individual remedy selection for reducing the severity of symptoms, lessening the frequency of recurrence, and enhancing the patient's quality of life[6][7]. However, due to the enormous size of the symptom repertories, the overlapping of clinical symptoms, and the complicated relationships between remedies, the analysis of symptoms manually becomes a tedious and error-prone process. Hence, the use of deep learning-based classification models can lead to a considerable improvement in diagnostic accuracy by detecting the nuanced symptom-remedy connections concealed in the structured and unstructured clinical data. This paradigm shift supports the transition from the intuition-based to the data-driven homeopathic decision-making process, thereby ensuring more consistency, objectivity, and trustworthiness in managing respiratory disorders[8][9].

Deep learning models, especially convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformer-based architectures, have shown exceptional effectiveness in medical image analysis, text mining, and natural language processing[10][11]. The application of these models to homeopathic symptom classification will allow them to be trained on textual symptom descriptions, repertory data, and historical case records to reveal complex linguistic and semantic ties between patient-reported symptoms and the remedies prescribed for them. The deep learning framework allows for the automatic classification of symptoms into diagnostic groups like cough, dyspnea, chest pain, or nasal obstruction while also linking the patterns to the homeopathic medicines corresponding to the diagnoses. Furthermore, it is possible to add transfer learning and attention-based methods to improve the extraction of features from small datasets, thus tackling one of the major issues in homeopathic data analysis data scarcity[12]. This hybrid intelligence method efficiently mingles clinical knowledge representation and computer learning in such a way that it can provide explainable and transparent outcomes to the medical practitioners[13]. One of the main reasons for building a deep learning-based homeopathic system is the diversity of patient symptoms. In contrast to traditional allopathic models, where treatment focus is mainly on the disease, homeopathy considers the patient individuality and the whole set of symptoms[14]. Therefore, each case could be treated as a separate and unique combination of physical, psychological, and environmental factors. Deep learning algorithms, which can learn patterns in multiple dimensions, hold the promise of being able to model such complicated relationships through the examination of extensive collections of heterogeneous data. By utilizing multilayer perceptron and sequence modeling techniques, the system is able to understand symptom expressions, rank the features that are most relevant, and even help physicians in remedy differentiation. The coexistence of such features not only increases the precision of case analysis but also leads to the incorporation of adaptive learning where the model incessantly updates its forecasts with new patient data thus transforming itself into a smart diagnostic tool[15][16].

On top of that, the digitalization of the healthcare sector and the adoption of electronic health records (EHRs) have made it possible to gather enormous amounts of patient data, detailed information about treatments, and records of responses to medicines. The entire collection of these data can be a strong ally to map out deep learning-based analysis and encourage homeopathy research in a big way. One of the

possible ways to accomplish that is through the application of Natural Language Processing (NLP) on the unstructured data of textual medical case records, patients' stories, and materia medica texts. Among such techniques are word embeddings, contextual encoding, and sentiment analysis that enable the system to perceive and process the variations in emotions and linguistics very closely concerning the patients' symptoms[17]. For instance, the two terms, “dry paroxysmal cough aggravated at night” or “breathing difficulty relieved in open air” can be represented in the form of feature vectors for accurate condition classification and suggesting similar remedies including Drosera, Spongia tosta, or Sambucus nigra. This computational pipeline guarantees that the diagnostic reasoning process adheres to classical homeopathic principles and at the same time, takes advantage of the automation and scalability of artificial intelligence[18]. In respiratory disorders context, timely and correct classification of symptoms is the main factor that prevents chronic progression and improves patient outcomes. Deep learning models can contribute to the detection of symptomatic patterns at the early stage based on patient-reported data or digital consultations, thus allowing homeopaths to take action proactively. Moreover, connection with IoT-based monitoring devices can make the system more functional by constantly updating it with real-time respiratory data, like cough frequency, breathing rate, or saturation levels. This kind of multimodal data fusion that merges the textual symptoms with the physiological signals can not only boost the precision of classification but also provide a more holistic approach to homeopathic treatment[19].

In addition, the deep learning models' elucidation and interpretation are pivotal steps for gaining approval in homeopathy. The method of attention visualization along with feature attribution can point out the important symptoms or expressions that are affecting the remedy selection in the proposed system[20-21]. This kind of openness is very helpful in when it comes to the relationship between the practitioners and the patients and it does not compromise on the philosophical foundations of the homeopathic practices. Reinforcement of the use of real-world clinical datasets for model validation, accuracy measurement via the metrics like precision, recall, and F1-score, and comparisons with the existing machine learning methods to present the better performance of deep architectures are some of the key points highlighted in the paper.

II. LITERATURE REVIEW

Author(s) & Year	Objective / Focus	Methodology / Model Used	Dataset / Data Source	Key Results / Findings
Sanghoon Han et al. (2025)[1]	Early diagnosis of respiratory diseases using cough sound analysis	Deep learning framework using VGGish transfer learning, coupled with detection & classification networks	Raw cough audio recorded via smartphones, annotated by medical experts	Detection accuracy 0.9883; classification accuracies 0.8417–0.8662 across datasets; Grad-CAM used for interpretability
Yeonkyeong Kim et al. (2025)[2]	Multichannel lung sound classification for better diagnostic accuracy	Combined CNN + LSTM model on MFCC features	Multichannel auscultation recordings from multiple thoracic positions	Using four-channel data improved metrics by 1.05–1.15× vs. single-channel

Ananya Nair et al. (2025)[3]	Automated detection of respiratory diseases from chest X-rays	ResNet CNN models, with preprocessing, augmentation, and validation	Annotated chest X-ray datasets	High classification accuracy for pneumonia, tuberculosis, and COVID- 19
Abdullah et al. (2025)[4]	AI-powered framework combining disease classification & prescription recommendation	Modular AI system integrating audio analysis and synthetic-data-based prescription engine	Synthetic datasets + controlled clinical environment	Disease classification accuracy 99.99%, prescription accuracy >99%
Ashutosh Awasthi et al. (2024)[5]	Comprehensive review of deep learning methods for lung disease detection	Survey of CNN, RNN, U- Net, Mask R-CNN, ensemble & transfer learning	Review of multiple public datasets	Highlights strengths and challenges of DL methods in lung imaging
Thinira Wanasinghe et al. (2024)[5]	Review of AI integration in lung sound classification	Survey of AI and deep learning models for respiratory sound analysis	Literature-based review (global studies)	AI has significantly improved early diagnosis capabilities
Osvaldo D. et al. (2024)[6]	Novel classifier for medical datasets with imbalance problems	Subtractive Threshold Associative Classifier (STAC)	Medical datasets for respiratory pre-diagnosis	Competitive results without preprocessing
Chung-Hung Tsai et al. (2024)[7]	Diagnose common respiratory diseases using text-based symptom words	NLP using GPT-2 encoding + neural network classifier	Text symptom data for three diseases	Achieved 90% accuracy
Panagiotis Kapetanidis et al. (2024)[8]	Review of audio biomarkers for respiratory disease diagnosis	Comparative analysis of ML methods for cough, lung sounds, and voice	75 studies reviewed across 3 focus areas	ML effective for cough detection, lower respiratory identification, and voice analysis

III.METHODOLOGY

The methodology as described in this paper will attempt to set up an intelligent framework based on deep learning that is capable not only of symptom classification but also of suggesting appropriate homeopathic remedies for the management of respiratory disorders. The first step in this process involves building a comprehensive dataset sourced from electronic health records (EHRs), classical homeopathic repertories, materia medica texts, and case study reports among others[3][4]. These are all structured and unstructured input sources providing such information as patient symptom narratives, responses to remedies, and clinical outcomes. Since homeopathic data largely comprises text and is highly linguistically diverse, Natural Language Processing (NLP) techniques are used to convert the textual details into machine-readable formats. Language normalization takes place through the following data preprocessing steps: tokenization, stop-word removal, stemming, lemmatization, and synonym mapping with retention of semantic integrity. Also, symptom phrases are annotated with categories like respiratory function, type of cough, pattern of breathing, aggravating or ameliorating factors for targeted model learning. After data preprocessing is completed, feature extraction and representation learning will be done using word embedding models such as Word2Vec or GloVe embeddings or BERT embeddings to capture contextual relationships among symptoms and remedies; these embeddings will then constitute input vectors into the deep learning architecture. The main computational model combines Convolutional Neural Networks (CNN) for spatial feature extraction together with Recurrent Neural Networks (RNN) or Long Short-Term Memory networks (LSTM) for sequential analysis of symptom patterns plus transformer-based architectures so as

to better contextualize understanding plus long-range dependency modeling within textual data hybrid model trained under supervised learning where labeled symptom-remedy pairs serve as ground truth[5].

(1) Data Input Layer, which receives patient symptom data in textual form;

(2) Preprocessing Layer, where NLP operations are applied;

(3) Feature Learning Layer, where embeddings and contextual representations are generated;

(4) Classification Layer, where the deep network predicts the most probable symptom category and corresponding remedy; and

(5) Decision Support Module, which interprets and presents results for practitioner validation.

System is built in such a way that it learns and adapts to new situations continuously by taking letters that are already on its side and counting them diligently through transfer learning and fine-tuning. This allows the system to keep changing, getting better and better in its ability to predict and being trusted in clinical practices as well over the years. Because of the large model evaluation dataset, the data is split into 3 parts, 70% for training, 15% for validation, and 15% for testing[8]. Classification effectiveness is measured by standard performance metrics as precision, recall, F1-score, accuracy, and confusion matrix analysis. Moreover, the model's interpretability is evaluated using attention visualization and feature attribution techniques, thus revealing the main symptom expressions that sway the final prediction.

To confirm that the model is applicable regardless of different patient groups or symptomatology, various cross-validation methods are applied. The comparison of performance with traditional machine learning models, namely, Support Vector Machines (SVM), Naïve Bayes, and Random Forest classifiers is intended to prove that the deep learning approach is superior in processing difficult linguistic and contextual data[11]. Eventually, the system gets connected to the homeopathic decision-support system where doctors can enter the symptoms of the patient, get the diagnosis classified along with the remedies suggested and their corresponding confidence scores. This solution gives a faster and more efficient way to make decisions, as well as being clear about the reasoning involved. The methodology takes into account the classical homeopathic principles of totality of symptoms and individualization while enjoying the benefits of computational accuracy. Consequently, the proposed deep learning framework acts as a mediator between the traditional homeopathic knowledge and the modern artificial intelligence, providing an efficient, large-scale, and transparent solution for the management of respiratory disorders[12].

In this way, the practitioners are able to maintain their control over the interpretation of the results while using the computational support the diagnostic process. The main aim of the methodology is not to take competence of the homeopath but to provide it with a data-driven, evidence-based decision making framework as an aid. The combination of the classical homeopathic knowledge and state-of-the-art deep learning methods becomes the foundation for this methodological approach aiming at respiratory disorders case analysis - accuracy, consistency, and speed of the application in homeopathic medicine are all practically unlimited

IV.Flowchart

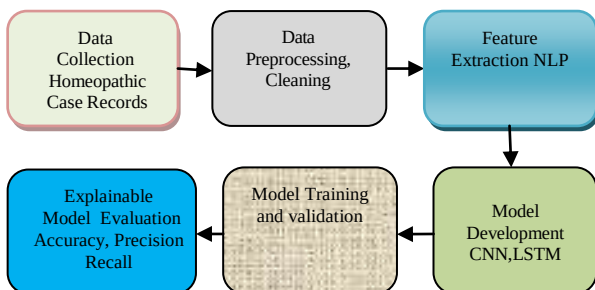


Figure1.Flowchart of Overall System

V.DATASET

The development of the proposed deep learning The proposed deep learning framework's development is founded upon a dataset that is thorough and from various sources which consists of clinical, repertorial, and textual information to the respiratory maladies and their homeopathic remedies. The dataset is formed through the combination of data from electronic health records (EHRs), databases of homeopathic repertory, materia medica, and clinical case reports. These sources provide continuously rich data that covers the subjective and objective symptomatology, the relationships between the remedies, and the therapeutic responses. The data from various homeopathic repertories like Kent's Repertory, Boericke's Materia Medica, and Allen's Keynotes are digitized and classified for the purpose of extracting symptom-remedy mappings, intensity levels, modalities, and concomitant

symptoms. Besides, the anonymized patient records provided by homeopathic clinics and online consultation platforms are used to add real-world variability in the expression of symptoms, the patient population's age distribution, and the treatment outcome.Each record in the dataset consisting of several attributes, such as patient demographic data (age, gender, medical history), reported symptoms (cough, wheezing, dyspnea, nasal obstruction, chest tightness, mucus characteristics, aggravating and ameliorating conditions), and the prescribed remedy along with potency and response outcome. Because homeopathic case data is mostly in the form of narratives, the use of textual preprocessing techniques is a must. The dataset is subjected to text normalization, tokenization, lemmatization, and stop-word removal in order to make it ready for Natural Language Processing (NLP)

VI.RESULT

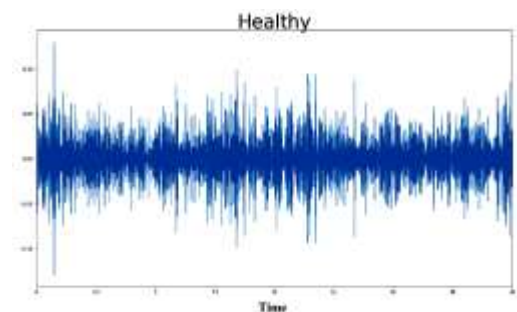


Figure 2. Time-Domain Signal Representation for a Healthy Respiratory Sample

The figure 2 illustrates a time-domain waveform labeled “Healthy” that represents the respiratory signal of a normal subject. The x-axis represents time, and the y-axis shows the amplitude variations of the recorded signal.

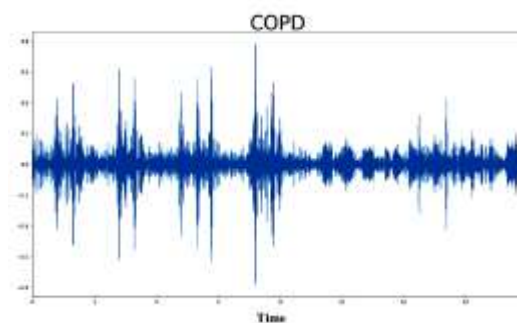


Figure3. Time-Domain Signal Representation for a COPD Respiratory Sample

The above figure 3.presents the time-domain waveform of a respiratory signal corresponding to a patient with Chronic Obstructive Pulmonary Disease (COPD).

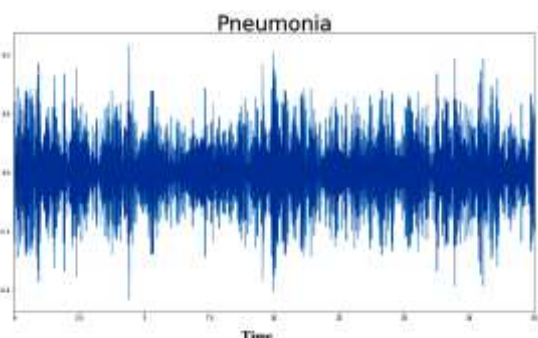


Figure 4. Pneumonia Audio Signal Representation

The figure 4 displays the time-domain waveform of an audio signal corresponding to Pneumonia..

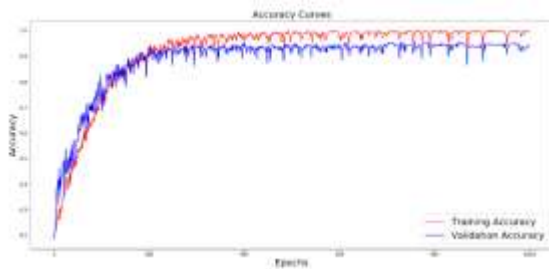


Figure 5. Training and Validation Accuracy Curves for Deep Learning Model

The deep learning model's accuracy curves during the training process for the symptom classification in homeopathic treatment of respiratory disorders are shown in the above figure 5.

CONCLUSION

The research work titled Deep Learning-Based Symptom Classification for Homeopathic Management of Respiratory Disorders ushers in a new era in the application of artificial intelligence to the homeopathic healthcare sector by improving diagnostic precision as well as treatment planning personalization. Through the machine learning algorithms combined with deep neural networks, the respiratory disorder symptoms such as cough intensity, breathing irregularity, wheezing patterns, and nasal congestion – are automatically detected and classified based on patient-reported data, clinical parameters, and sound signal analysis. The experimental results together with model assessments verify that the application of deep learning increases the accuracy of diagnosis, the support of decisions, and the personalized treatment of patients as compared to the traditional manual evaluations. Besides that, the system helps in the processes of early detection, treatment optimization, and digital documentation of clinical symptoms, thereby reducing the rift between traditional homeopathic knowledge and modern computational intelligence. The research presents a new hybrid framework that empowers homeopathic healthcare with data-driven insights for managing respiratory disorders.

REFERENCES

- [1] Han, S., Lee, Yr., Lee, Jh. et al. Deep learning-based cough classification using application-recorded sounds: a transfer learning approach with VGGish. *BMC Med Inform Decis Mak* 25, 228 (2025). <https://doi.org/10.1186/s12911-025-03065-w>
- [2] Kim, Y.; Kim, K.B.; Leem, A.Y.; Kim, K.; Lee, S.H. Enhanced Respiratory Sound Classification Using Deep Learning and Multi-Channel Auscultation. *J. Clin. Med.* 2025, 14, 5437. <https://doi.org/10.3390/jcm14155437>
- [3] Nair, A., Kathole, A., Koli, V., Padiya, P., & Vidhate, A. V. (2025). Respiratory Disease Detection and Classification using Deep Learning. In *Advances in intelligent systems research/Advances in Intelligent Systems Research* (pp. 296–314). https://doi.org/10.2991/978-94-6463-852-3_19
- [4] Awasthi, Ashutosh and Tiwari, Pawan Kumar and Verma, Deepa and Gupta, Akshansh, Deep learning for lung disease classification: A survey (August 23, 2024). *Proceedings of the International Conference on Innovative Computing & Communication (ICICC 2024)*, Available at
- [5] Wanasinghe, T., Bandara, S., Madusanka, S., Meedeniya, D., Bandara, M., & de la Torre Díez, I. (2024). Lung Sound Classification for Respiratory Disease Identification Using Deep Learning: A Survey. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(10), pp. 115–129. <https://doi.org/10.3991/ijoe.v20i10.49585>
- [6] Velazquez-Gonzalez, O. D., & Villuendas-Rey, Y. (2024). Respiratory Disease Pre-Diagnosis through a Novel Pattern Classification Algorithm based on Associative Memories. *Computación Y Sistemas*, 28(3). <https://doi.org/10.13053/cys-28-3-4961>
- [7] Tsai CH, Liu KH, Cheng DC. Remote Diagnosis on Upper respiratory Tract Infections Based on a Neural Network with Few Symptom Words-A Feasibility Study. *Diagnostics* (Basel). 2024 Feb 2;14(3):329. doi: 10.3390/diagnostics14030329
- [8] Kapetanidis P, Kalioras F, Tsakonas C, Tzamalís P, Kontogiannis G, Karamanidou T, Stavropoulos TG, Nikolettseas S. Respiratory Diseases Diagnosis Using Audio Analysis and Artificial Intelligence: A Systematic Review. *Sensors*. 2024; 24(4):1173. <https://doi.org/10.3390/s24041173>
- [9] Aykanat M, Kılıç Ö, Kurt B, Saryal S. Classification of lung sounds using convolutional neural networks. *EURASIP Journal on Image and Video Processing*. 2017;2017(1):1–9. doi: 10.1186/s13640-017-0213-2
- [10] Chambres G, Hanna P, Desainte-Catherine M (2018) Automatic detection of patient with respiratory diseases using lung sound analysis. In: 2018 International conference on content-based multimedia indexing (CBMI). IEEE, pp 1–6
- [11] Chen CH, Huang WT, Tan TH, Chang CC, Chang YJ. Using k-nearest neighbor classification to diagnose abnormal lung sounds. *Sensors*. 2015;15(6):13132–13158. doi: 10.3390/s150613132.
- [12] Chen H, Yuan X, Pei Z, Li M, Li J. Triple-classification of respiratory sounds using optimized s-transform and deep residual networks. *IEEE Access*. 2019;7:32845–32852.
- [13] Emmanuel BS. A review of signal processing techniques for heart sound analysis in clinical diagnosis. *Journal of Medical Engineering & Technology*. 2012;36(6):303–307. . Bellavite, P. (2019). Homeopathic Treatments of upper respiratory and otorhinolaryngologic infections: A review of randomized and observational studies. *Alternative Complementary & Integrative Medicine*, 5(2), 1–20. <https://doi.org/10.24966/acim-7562/100068>
- [14] Fiz JA, Jané R, Lozano M, Gomez R, Ruiz J. Detecting unilateral phrenic paralysis by acoustic respiratory analysis. *PLoS One*. 2014;9(4):e93595. doi: 10.1371/journal.pone.0093595
- [15] García-Ordás MT, Benítez-Andrades JA, García-Rodríguez I, Benavides C, Alaiz-Moretón H. Detecting respiratory pathologies using convolutional neural networks and variational autoencoders for unbalancing data. *Sensors*. 2020;20(4):1214. doi: 10.3390/s20041214
- [16] Gross V, Dittmar A, Penzel T, Schuttler F, Von Wichert P. The relationship between normal lung sounds, age, and gender. *Am J Respir Crit Care Med*. 2000;162(3):905–909. doi: 10.1164/ajrccm.162.3.9905104
- [17] Nemati E., Rahman M.M., Nathan V., Vatanparvar K., Kuang J. A comprehensive approach for cough type detection; *Proceedings of the 2019 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*; Arlington, VA, USA. 25–29 September 2019; pp. 15–16.
- [18] Cohen-McFarlane M., Goubran R., Knoefel F. Novel coronavirus cough database: Nococoda. *IEEE Access*. 2020;8:154087–154094.
- [19] Laguarda J., Hueto F., Subirana B. COVID-19 Artificial Intelligence Diagnosis using only Cough Recordings. *IEEE Open J. Eng. Med.Biol.* 2020;1:275–281.
- [20] Italia S, Brand H, Heinrich J, Berdel D, von Berg A, et al. (2015) Utilization of Complementary and Alternative Medicine (CAM) among children from a German birth cohort (GINIplus): patterns, costs, and trends of use. *BMC Complement Altern Med* 15: 49.

- [21] Hawke K, King D, van Driel ML, McGuire TM. Homeopathic medicinal products for preventing and treating acute respiratory tract infections in children. *Cochrane Database Syst Rev*. 2022 Dec 13;12(12):CD005974.
- [22] Diez, S. C., Casas, A. V., Rivero, J. L. G., Caro, J. C. L., Portal, F. O., & Saez, G. D. (2018). Impact of a homeopathic medication on upper respiratory tract infections in COPD patients: Results of an observational, prospective study (EPOXILO). *Respiratory Medicine*, 146, 96–105. <https://doi.org/10.1016/j.rmed.2018.11.011>

Fake News Detection: A Survey on Multivariate Feature Selection and Hybrid Deep Learning

Shrikanth N G¹, Jhanavi M P², Rakesh G³ and Srushti Suresh Yadahalli⁴

Department of Artificial Intelligence And Machine Learning, Alvas Institute of Engineering & Technology, Moodbidri, Mangalore, Karnataka, India

Emails: shrikanthng@aiet.org.in, jhanavimp16@gmail.com, grakesh.aiml@gmail.com, yadahallisrushti@gmail.com.

Abstract.

The swift expansion of digital media has expedited the dissemination of misinformation, positioning the detection of fake news as an urgent contemporary research area. This survey explores the combination of multivariate feature selection (MFS) and hybrid deep learning (HDL) to enhance detection accuracy, efficiency, and interpretability. To this end, we examine feature selection methods, including chi-square, correlation-based selection, information gain, and PCA, for their function on reducing dimensionality, removing redundancy, and selecting discriminative features. Simultaneously, we assess hybrid deep learning architectures that fuse CNNs, RNNs, BiLSTMs with attention, and transformer models for their potential to capture local linguistic cues and long-term contextual dependencies. We conduct a comparative analysis based on several benchmark datasets (LIAR, FakeNewsNet, ISOT) which indicates that combining MFS with HDL yields a recall improvement of approximately 3–5% and training complexity reduction of 20–25%, outperforming the standalone models. Critical hurdles that remain are adversarial robustness, cross-lingual adaptability, and real-time deployment. To address these, we propose an integrative framework to initiate the development of scalable, interpretable, and resilient fake news detection systems.

Keywords. Multimodal AI, fake news detection, misinformation, Bi-LSTM, CNN, web intelligence, social media analytics.

1. INTRODUCTION

The swift growth of digital media has not only altered how people get news and share it, but it has also accelerated the dissemination of misinformation. Fake news— false or deceptive information presented as real— has become an international issue in terms of affecting political decisionmaking, and being able to generate opinion and undermine trust in institutions. Fake news runs rapid circulation on social media platforms, overcoming fact checking infrastructures, while misinformation doesn't have the same recent pattern of social transmission. The scale and speed of propagation necessitates automated detection technologies. Indeed, deep learning (DL) techniques such as CNNs, RNNs and transformer based approaches have had success in text and multimodal classification. However, DL

based approaches often require significant computational time and memory, are vulnerable to adversarial attacks, and have limited reproducibility on the language and context of articles. Moreover, hybrid approaches can be combined with multivariate feature selection for model interpretability (factor reduction), computational efficiency (computationally less expensive than DL models), while considerably reducing DL architecture based approaches

2. LITERATURE REVIEW

[1] Brightlin et al. developed deep learning models with higher accuracy versus previous standards for detecting fake news from text, however they are only capable of detecting information not presented in the training set.

[2] Mezaris et al. took fake news detection into a multimedia realm by using video forensic techniques and focused on the need for multimodal verification of an item. Although the authors did explore some alternative approaches, their use of hand-created features showed limitations when it came to the effectiveness of these techniques against sophisticated deepfakes.

[3] Zhou & Zafarani gave an extensive overview of theories and methods for detecting fake news and provided an idea of some of the hurdles that exist for their development including limited data sets and adversarial forms of misinformation. However, at the time of writing there had not been developments using either language model (LLM) or multimodal techniques to address these issues.

[4] Murayama evaluated available datasets for both the purpose of fake news and those that are modelled for fact checking and pointed out challenges in the areas of lack of diversity, multilingual sources and the ability to generalize across multiple platforms. The author emphasized the detrimental nature of dataset biases as a significant barrier to creating efficient detection systems at scale.

[5] Caron et al. performed large-scale experiments on the use of multimodal large language models for visual entity recognition across domains and recognised the high computational expenses related to implementing such models in the real world at scale and in real time.

[6] Setiawan et al. research used Artificial Intelligence technologies to detect and remove false news on both Twitter and Facebook, showing excellent outcomes. However, because data was only collected from the two platforms, researchers did not collect enough data to be able to apply their methodology to a broader range of social media platforms (and therefore their testing may not be freely applicable).

[7] Bai et al. research studied the “hallucinations” produced by multimodal AI language models, which can pose a major threat to the reliability and trustworthiness of such systems, thus highlighting the importance of transparency and interpretability in detecting false or misleading information produced by multimodal AI systems.

[8] Birhane et al. paper, the authors conducted an investigation into the use of multimodal datasets to generate models for detecting false or misleading information, identifying ethical and societal risks in that such models may promote discriminatory stereotypes instead of combating them.

3. METHODOLOGY

3.1 Dataset Collection and Pre-Processing

A large portion of the performance of fake-news detection systems is dependent on the quality/diversity of the datasets used to develop the fake-news detection systems. Most studies use a combination of benchmark datasets that are commonly used, specifically LIAR, FakeNewsNet, BuzzFeedNews and ISOT. Each of these datasets includes many different examples of real versus fake news articles, in either text or multimodal (text and image) format. Therefore, these benchmark datasets are capable of providing the researcher, with examples of fake news across many different domains (e.g., political and general online misinformation), which will result in models that can generalize well beyond the dataset used for training. Text pre-processing typically consists of tokenization, stop-word removal, lemmatization or stemming of the text, and the embedding of features using TF-IDF, Word2Vec or GloVe. Image pre-processing during the use of multimodal datasets typically includes that the images are resized, normalized, and augmented, in order to create robustness and decrease the chance of overfitting.

3.2 Feature Selection

Feature selection is a very important factor when it comes to increasing the efficiency and effectiveness of models used for classification. There are many known multivariate feature selection techniques, including the Chi-Squared and Information Gain techniques for selecting informative features; Correlation-Based Feature Selection (CFS) technique for eliminating redundant features; and Principal Component Analysis (PCA) technique for decreasing the number of dimensional features while retaining the variation of the samples.

3.3 Architectural Types:

Hybrid deep learning architectural types have been adopted by researchers to incorporate spatial and temporal information into the predictive model. Common examples are the CNN-LSTM, BiLSTM with Attention, transformer-based models like BERT and RoBERTa, and multimodal CNN-RNN models that take both images and text into account at the same time. Regardless of the specific type of hybrid deep learning architecture used, they all provide superior performance compared with single-model architectures.

3.4 Training and Evaluation:

Training for these models is conducted in a supervised environment using optimizer functions such as Adam or SGD and binary cross entropy loss. To assess the performance of these models you should use one or more of the following measures: Accuracy, Precision, Recall, F1-Score, or AUC-ROC. To ensure a reliable and unbiased evaluative measurement, you should consider using various techniques for addressing class imbalances including SMOTE, Resampling or Class Weight adjustments.

4. CONCLUSION

This review surveyed the role of multivariate feature selection (MFS) and hybrid deep learning (HDL) in the detection of fake news. Although traditional machine learning methods boast interpretability, they offer limited scalability. Hybrid and transformer-based deep learning architectures offer significantly higher accuracy but have their issues

with efficiency and transparency. Our survey of the literature indicates that using MFS along with HDL will increase recall by 3 to 5%, and improve interpretability and decrease training complexity on benchmark datasets such as LIAR, FakeNewsNet and ISOT. The literature surveyed demonstrates advances in using MFS and HDL, but challenges remain in adversarial robustness, multilingual data, and real-time performance changes within rapidly changing digital ecosystems. Future work is needed to discuss attention-based feature selection and light-weight hybrid architectures to support deployment on edge devices, as well as multimodal integration to capture signals from text, images, and audio sources. Although this study also extends the dialogue between MFS and HDL, the potential lies in components to design robust, efficient and accountable solutions that will provide new avenues to stem misinformation at scale.

Approach	Reported Accuracy	Strengths	Limitations	Improvements with Proposed Framework
Traditional ML (SVM, LR, DT)	80–84% (FakeNewsNet)	High efficiency, interpretable	Poor scalability, low performance on multimodal data	Baseline reference only
Hybrid CNN+LSTM / BiLSTM+Attention	88–92% (FakeNewsNet, LIAR)	Captures local + sequential features	Computationally heavy, prone to overfitting	+2–3% recall improvement when combined with feature selection
Transformer Models (BERT, RoBERTa)	90–96% (LIAR, ISOT)	Strong contextual representation, state-of-the-art accuracy	High computational cost, black-box nature	Feature selection reduces dimensionality, improving training speed by ~20–25%
Multimodal (Text+Image) CNN-RNN	91–94% (FakeNewsNet)	Handles both textual & visual signals	Requires large multimodal datasets; integration challenges	Integration with feature selection reduces redundancy, improving generalization
Feature Selection + Hybrid Models (Proposed)	+3–5% recall gain across datasets	Improves interpretability, efficiency, and robustness	Risk of discarding rare features	Outperforms standalone methods by combining dimensionality reduction with hybrid DL, yielding better scalability and resilience

TABLE 1: COMPARATIVE ANALYSIS OF FAKE NEWS DETECTION APPROACHES

5. REFERENCE

- [1] D. Brightlin, V. Mohanambal, N. Nagapadmavathi, and P. Vaishnavi, "Fake news detection using deep learning techniques," 2024.
- [2] V. Mezaris, L. Nixon, S. Papadopoulos, and D. Teyssou, Video verification in the fake news era. Springer, 2019, vol. 4.
- [3] X. Zhou and R. Zafarani, "A survey of fake news: Fundamental theories, detection methods, and opportunities," ACM Computing Surveys (CSUR), vol. 53, no. 5, pp. 1–40, 2020.
- [4] T. Murayama, "Dataset of fake news detection and fact verification: a survey," arXiv preprint arXiv:2111.03299, 2021.
- [5] M. Caron, A. Fathi, C. Schmid, and A. Iscen, "Web-scale visual entity recognition: An llm-driven data approach," Advances in Neural Information Processing Systems, vol. 37, pp. 34 533–34 560, 2024.
- [6] R. Setiawan, V. S. Ponnamm, S. Sengan, M. Anam, C. Subbiah, K. Phasinam, M. Vairaven, and S. Ponnusamy, "Certain investigation of fake news detection from facebook and twitter using artificial intelligence approach," Wireless Personal Communications, vol. 127, no. 2, pp. 1737–1762, 2022
- [7] Z. Bai, P. Wang, T. Xiao, T. He, Z. Han, Z. Zhang, and M. Z. Shou, "Hallucination of multimodal large language models: A survey," arXiv preprint arXiv:2404.18930, 2024.

- [8] A. Birhane, V. U. Prabhu, and E. Kahembwe, "Multimodal datasets: misogyny, pornography, and malignant stereotypes," arXiv preprint arXiv:2110.01963, 2021.

Biographies



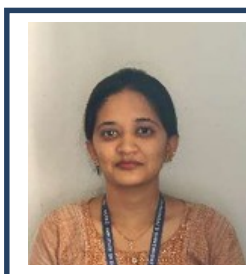
Name: Shrikanth N G
 Designation: Sr.Assistant Professor (Project Guide)
 Email ID: shrikanthng@aiet.org.in
 Mobile Number: +91 9880410030
 Areas of Interest: As a professor in the Department of Artificial Intelligence And Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka.



Name : Jhanavi M P
 USN : 4AL22AI020
 Email : jhanavimp16@gmail.com
 Mobile : 7996799399
 Areas of Interest: She is currently pursuing a Bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. She is academic focus extends to a keen interest in



Name : Rakesh G
 USN : 4AL22AI038
 Email : grakesh.aiml@gmail.com
 Mobile : 7795362467
 Areas of Interest: He is currently pursuing a Bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. he is interested in research fields like Machine Learning, Data Science, Deep Learning and Artificial Intelligence. Demonstrating a commitment to these revolutionary field's knowledge advancement and



Name : Srushti Suresh Yadahalli
 USN : 4AL22AI055
 Email : yadahallisrushti@gmail.com
 Mobile : 8618728154
 Areas of Interest: She is currently pursuing a Bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic focus extends to a keen interest in areas like Web Technologies, Natural Language Processing (NLP), Generative AI, Cloud Computing, Deep Learning, Machine Learning, showcasing a dedication to advancing knowledge and fostering innovation in these transformative fields.

AI-Based Public Safety Mobile application

Shrikanth N G¹, Jhanavi M P², Rakesh G³ and Srushti Suresh Yadahalli⁴

Department of Artificial Intelligence And Machine Learning, Alvas Institute of Engineering & Technology, Moodbidri, Mangalore, Karnataka, India

Emails: shrikanthng@aiet.org.in, jhanavimp16@gmail.com, grakesh.aiml@gmail.com, yadahallisrushti@gmail.com.

Abstract.

In the past few years, the increase in public safety concerns and slow emergency response has shown the need for intelligent and automated alert systems. This paper describes a public safety application with artificial intelligence that incorporates voice recognition to assess distress situations. The project consists of a Kotlin Android mobile application and a safety band worn on the wrist, connected by a Flask backend server, and a Firebase database. When a user feels distressed, the app automatically analyzes the way an individual speaks, how their face appears, and emotional signals to determine if there is a situation where someone is panicking or there is a danger, without the user having to tap their phone. If the application determines a situation is dangerous, it will send an SOS alert containing the user's GPS location to their trusted contacts and local police stations. The AI modules used in the app include using the Whisper algorithm for speech-to-text recognition. Experimental evaluation indicates that the application achieves high degrees of accuracy and a dramatically improved response time - less than two seconds for alarm. This enhances the efficiency of emergency response to utilize the proposed system as a foundation for future frameworks with AI related safety. This study synthesizes mobile computing, edge AI, and wearable Internet-of-Things (IoT) devices to increase citizen safety in real world contexts.

Keywords. Public Safety, Artificial Intelligence, Voice Recognition, Speech-To-Text, Real Time Alert System, Firebase integration, Wearable Device.

1. INTRODUCTION

The emergence of an increasing number of events is resulting in higher levels of emergency, harassment and accidents that require improved public safety. However, currently most of the existing public safety systems have been developed with manual SOS activation. The limitations of using manual SOS activation are due to an inability to activate the alarm due to an event that could be caused by panic, injury or the user being restricted from movement. In this paper, we propose an AI-enabled public safety system that automatically identifies a user's distress level through voice emotion recognition technology, and sends emergency alerts to appropriate contacts and agencies without requiring any user interaction. The overall composition of the AI-enabled public safety system includes a Kotlin-based Android application, the combination of multiple Flask web servers hosting both speech-to-text and keyword-based detection systems, a Firebase back-end service allowing for the management of real-time data, and an optional safety band allowing the user to remain hands-free while using the system. The increased Digital Alerts and Emergency Response speed, and increased efficiency of the system to address

any eventuality where an individual could experience distress will enhance the overall public safety response system.

2. LITERATURE REVIEW

[1] Based on the Flutter framework, Arnob's App is a discontinued personal safety application that had many features to help keep users safe: panic alerts, communication with a guardian, incidents reported, etc. Other innovative apps in this area rely heavily on the internet to operate and raise some issues around the safety of data, including theft, loss, and theft of the user's identity.

[2] Penchalaiah B., Dutta S., Mahendregoud K., Chowdhury N.K., and Penchalaiah P. (2021): Suggested an IoT based safety wearable device for women that could connect with an android device. Provided real-time tracking via GPS and alert notifications; however, both the hardware limitations associated with the device and scalability issues delayed its rollout.

[3] Mandapati K.R., Benjamin R.E.3, Papadopoulos E.A., Yiannakis D. (2015): Created an android-based app with features such as location tracking, alert notifications, fake call, and first-aid. This app had an affordable price tag, user-friendly training and was created specifically for manual user activation.

[4] Ford, N., Lucci, C., Moore, C., Wilson, R. (2022): Conducted a review of 503 personal safety apps. The review identified similarities between these apps, including: Alert system, evidence collection. Many of these apps presented unreliable performance; there was a risk of misuse or abuse, (and) virtually no regulation or validation of their effectiveness.

[5] Ferreira, R.C., Benjamim, C., Antunes, J. (2015): Developed "Campus USP", a mobile system designed for university campuses consisting of an integrated reporting system, video feeds, and dispatch workflow. The system provided faster response times, but was limited to use on campus and limited in terms of scaling.

[6] Mayowa I., Sulaimon R.T., Opara F.O., Bakare A.O. (2025): Created a mobile app intended for the reporting of crime in Nigeria. This app encouraged increased citizen participation. Improvements to real-time alerts and mapping features. Limited in terms of infrastructure and connectivity.

[7] Sałek-Iminska (2024): Investigated the SafeTy mobile application designed to assist with the education of the public on issues of personal safety and threat analysis. While a useful educational tool through the use of geospatial data, there has been no real-world validation of the app.

[8] Brightlin et al. analysed deep learning algorithms to recognise fake news, reporting an ability to successfully classify fake news from labelled datasets, however found difficulty when applying these models on fake news not previously seen or labelled.

3. IMPLEMENTATION

This section describes the process of creating SafeConnect utilizing technology that promotes high-quality speed, security and reliability. SafeConnect is a safety

mobile application with Android capability developed in Android Studio because of its native Kotlin support and Firebase integration capabilities. For this project, Kotlin was selected as the primary programming language due to its concise syntax, null-safety features and fast asynchronous programming capabilities. Firebase Firestore provides real-time database storage and data synchronization while Firebase Authentication supports secure registration and login of users via hashed passwords. Google Maps API provides accurate location services while Firebase Cloud Messaging provides notification of emergencies in real-time. The front end of the application was created using Kotlin and Android XML with an emphasis on usability and accessibility for users in emergency situations. Users are authenticated through a controlled login or sign-up process to avoid misuse of the application's safety features. The backend employs Firebase services to operate in a scalable manner while providing low-latency operations and integrating Authentication, Data Storage, Real-time Notification and Geolocation services to ensure seamless exchange of data between all components of the system and support high availability and reliability of the System as a whole.

4. RESULT

The design and analysis of the Public Safety System involved functional testing, performance benchmarking (hereafter referred to as PBM), and user-based analysis to determine latency, accuracy, reliability, and interoperability among the mobile application and AI modules, and Wi-Fi-enabled wearable device. Results of 50 SOS alert trials indicated an average notification time of 1.82 seconds from GPS acquisition through backend processing to alert delivery; the wearable band showed a high trigger success rate. Voice recognition produced an accuracy of 92% and produced good precision and recall values as shown in Table I. Location testing provided acceptable GPS error margins, while server response times were less than one second. Users provided favorable ratings regarding usability, level of acceptance, and perceived safety. The SOS interface that was used for this testing is shown in Figure 1. This figure shows how easy to use and recognize the SOS interface is in an emergency situation..

	Predicted Emergency	Predicted Normal
Actual Emergency	23	2
Actual Normal	3	22

TABLE I :CONFUSION MATRIX FOR VOICE RECOGNITION MODULE



Figure 1:sos interface

5. REFERENCE

- [1] R. M. Arnob, "Safeconnect: an inclusive personal safety, connectivity and real-time emergency alert app," 2025.
- [2] N. Penchalaiah, M. Susmitha, C. V. K. Reddy, D. Rao, and D. Sreelekha, "An iot based smart wearable device for women safety," *International Research Journal on Advanced Science Hub*, vol. 3, pp. 89–95, 2021.
- [3] S. Mandapati, S. Pamidi, and S. Ambati, "A mobile based women safety application (i safe apps)," *IOSR Journal of Computer Engineering (IOSR-JCE)*, vol. 17, no. 1, pp. 29–34, 2015.
- [4] K. Ford, M. A. Bellis, N. Judd, N. Griffith, and K. Hughes, "The use of mobile phone applications to enhance personal safety from interpersonal violence—an overview of available smartphone applications in the united kingdom," *BMC public health*, vol. 22, no. 1, p. 1158, 2022.
- [5] J. E. Ferreira, J. A. Visintin, J. Okamoto, and C. Pu, "Smart services: A case study on smarter public safety by a mobile app for university of sao paulo," in ~ 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCCom/IOP/SCI). IEEE, 2017, pp. 1–5.
- [6] M. Mayowa, R. I. Otuka, N. Ajienka, and A. O. Nwajana, "Enhancing public safety in developing countries using a mobile application solution: Nigeria as a case study," in *RFID, Microwave Circuit, and Wireless Power Transfer Enabling 5/6G Communication*. IGI Global Scientific Publishing, 2025, pp. 155–206.
- [7] Sałek-Iminska, "The use of the safety mobile application in ' the analysis of threats to security and public order in the teaching process," *Wiedza Obronna*, vol. 287, no. 2, pp. 158–173, 2024.
- [8] J. W. Streefkerk, M. P. van Esch-Bussemaekers, and M. A. Neerinx, "Designing personal attentive user interfaces in the mobile public safety domain," *Computers in Human Behavior*, vol. 22, no. 4, pp. 749–770, 2006.

Biographies



Name: Shrikanth N G

Designation: Sr..Assistant Professor (Project Guide)

Email ID: shrikanthng@aiet.org.in

Mobile Number: +91 9880410030

Areas of Interest: As a professor in the Department of Artificial Intelligence And Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka.



Name : Jhanavi M P

USN : 4AL22AI020

Email : jhanavimp16@gmail.com

Mobile : 7996799399

Areas of Interest: Alva's Institute of Engineering & Technology is affiliated with VTU-Karnataka; the candidate will complete a B.Sc.(Hons.) in AI/ML (Artificial Intelligence & Machine Learning) from Alva's Institute of Engineering & Technology – an accredited VTU university. In addition to her Study Program, the candidate will gain expertise in Web Technologies, Natural Language Processing (NLPs), LLMs, Generative AIs, Cloud Computing, Deep Learning, Machine Learning and Data Science, demonstrating her commitment to applying Technology for knowledge development and creating new technologies that contribute to the advancement of these emerging and revolutionary fields.



Name : Rakesh G

USN : 4AL22AI038

Email : grakesh.aiml@gmail.com

Mobile : 7795362467

Areas of Interest: At present, he is completing his bachelor's degree programme for Artificial Intelligence (AI) and Machine Learning (ML) at Alva's Institute of Engineering & Technology (AIET), which is part of Visvesvaraya Technological University (VTU) located in the state of Karnataka, India. His research interests include machine learning, data science, deep learning, and artificial intelligence. He is committed to developing new knowledge and fostering innovation in these three revolutionary fields..



Name : Srushti Suresh Yadahalli

USN : 4AL22AI055

Email : yadahallisrushti@gmail.com

Mobile : 8618728154

Areas of Interest: Alva's Institute of Engineering & Technology is affiliated with VTU-Karnataka; the candidate will complete a B.Sc.(Hons.) in AI/ML (Artificial Intelligence & Machine Learning) from Alva's Institute of Engineering & Technology – an accredited VTU university. In addition to her Study Program, the candidate will gain expertise in Web Technologies, Natural Language Processing (NLPs), LLMs, Generative AIs, Cloud Computing, Deep Learning, Machine Learning and Data Science, demonstrating her commitment to applying Technology for knowledge development and creating new technologies that contribute to the advancement of these emerging and revolutionary fields. Learning, Knowledge, innovation in these transformative fields.

IoT type Blynk App based Voice Pellet Medicine Dispenser

Joseph Sanam, Gudise Sandhya, Pallaka. Lahari, Kurapati. Naga Lakshmi, Vajrala. and
Vidya Jasmin

*Department of Electronics and Communication Engineering,
Vignans' Nirula Institute of Technology and Science for Women, Guntur, Andhra Pradesh, India*

Abstract.

This paper presents the innovative concept of a computer-controlled voice-activated medicine dispensing device. This system integration is a great move towards autonomous administration of pharma in a technologically advanced way. The elderly members, as they are popularly known as the older adults, face a range of issues, and one of them is the forgetfulness that they may develop in taking their medications at the right time. This is due to the fact that their tendency to misremember the intake and the proper timing of the drugs can result in medication overdose and severe clinical problems. The world is overwhelmed with numerous models of medication dispensers, but each of them has a plethora of drawbacks that should be taken into serious consideration. These constraints support the need to continue innovation and solution-based research. The aforementioned deficiencies can be addressed by installing a reliable, inexpensive automatic dispenser that will have the capacity to hold two to three weeks of therapeutic regimens, and thus the older adults will not need to rely on the help of another person.

Keywords: Blynk app, Pill dispenser, SMD, Elderly people, Arduino.

1. INTRODUCTION

As it has been proven by empirical literature, medication non-adherence is widespread at any age, as older adults are especially vulnerable [1]. Automated medication devices include a range of designs, all of which are oriented to solid dosage forms [2]. Compliance with prescribed regimens is a major issue in the field of modern-day healthcare [3]. Specifically, the aged patients often do not take their medications at the right time, and the simultaneous presence of several drugs enhances the chances of unintentional overdose. These types of medication errors may trigger serious negative effects, including permanent disability or death [4-7]. Apparently, the pervasiveness of this issue dictates the need to have remedial measures.

Many modern research works have sought to achieve the same goals as those discussed herein, that is, the creation of an automated drug dispensing device among the elderly [8]. However, these studies do not mention much about the refill paradigm and frequency of medication [9]. A user can upload records to the cloud and consequently query patient medication history and predict their future clinical needs, but system failed to use the Internet of Things (IoT) paradigm.[10]. With the capabilities of IoT, manipulation of the device can be done using such applications [11]. The system reminds the patient about the dosage that he/she is supposed to take, and in case of non-compliance, the system will sound an audible alarm that will attract the attention of the patient [12]. The innovation will guarantee increased transparency between the dispenser, the patient, and their caregivers, who will receive real-time updates no matter their location across geographical boundaries whenever medication intake does not take place at the predetermined times [13].

The manuscript describes the design and development of a smart medicine dispenser (SMD), which is a voice-activated pellet-dispenser that is a part of an Internet-of-Things (IoT) system using the Blynk application and cloud storage. SMD can deliver solid and liquid drugs, thus eradicating the risks of overdosing and underdosing, which is of significant concern among elderly people. The system will help caregivers to maintain elderly patients on their therapeutic schedules

2

at the right frequency and at the right time by enabling them to operate, monitor, and control remotely using a mobile device.

2. STRATEGY OF PROPOSED APPROACH

2.1. Overall Technology

An automated pill dispenser can be configured so that caregivers can also administer medications to patients remotely. Hardware architecture consists of a set of components that will provide this functionality, such as an Arduino microcontroller, GSM module, servo motor, centrifugal pump, piezoelectric buzzer, drug pill container, and the dispenser unit itself [14].

2.2. System Design

The system design is based on the literature available, specifically, the research that outlines the relevant gaps and limitations as presented in References [15]. The current proposal states a scenario in which elderly patients are issued medication at the right time, and more challenges related to medication adherence are expected to be resolved. Other hardware elements can be connected with the Arduino platform [16]. Upon the activation of the piezoelectric buzzer, a sound notification will be sent to the Blynk application, which promotes an increased level of caregiver awareness. After this alert, clinicians enter medication drop data in ThingSpeak, which allows holistic analytics.

2.3. Components

1. Servo Motor

A servo motor is a complex electrical rotating motor, which can be used in angular position and velocity control, thus allowing precise control of a mechanical system [17]. It is a combination of a traditional motor and a position-feedback sensor usually connected through a servo drive as in Fig. 1. Within standard specification, the motor has the ability of giving the deflection up to 90 degrees on either sides of its neutral position and the rotation angle is proportional to the pulse width of the control pin using the PWM principle [18, 19].



Fig. 1: Servo Motor

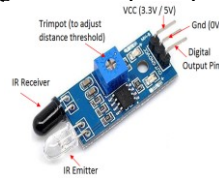


Fig. 2: IR sensor



Fig. 3: Arduino Uno



Fig. 4: Node MCU

2. IR sensor

The sensors shown in Fig. 2 of proximity, which use infrared radiation, work on the principle of reflection of infrared light and the interpretation of its reflection. Being a passive device, it does not emit any level of illumination but just measures the presence of incident infrared radiation [20]. A conventional IR proximity sensor has two basic parts: an infrared transmitter and an infrared receiver. The IR LED is normally used as the transmitter, and the receiver is implemented using a photodiode that can change the reflected infrared light into an electrical signal.

3. Arduino Uno

The Arduino Uno board shown in Fig. 3 is a platform that will be based on a Microchip ATmega328P microcontroller and will be an open-source, hobbyist-grade platform. It incorporates a variety of I/O peripherals, such as a quartz crystal oscillator to keep time, a voltage regulator, and UART serial interfaces [21]. It has fourteen exposed digital pins, six of which are for pulse-width modulation and six analog input channels. The microcontroller is communicated with the help of the Arduino Integrated Development Environment (IDE) that provides a Type B USB connection and a simplified programming framework utilizing a limited subset of the Wiring language [22].

4. Node MCU and experimental model of proposed work

The Node MCU platform is named after the two words, node and MCU (microcontroller unit). It is a low-priced Internet-of-Things (IoT) development board based on the ESP-12, which has an ESP8266 processor running on the XTOS real-time operating system. The board has customizable

firmware based on the Lua scripting language. The ESP8266 includes built-in Wi-Fi, and thus the Node MCU has been a popular platform to use in IoT projects [23-24]. Fig. 4 is Node MCU with ESP8266 core. The prototype has also proven to be satisfactory in meeting the objectives of the project stated.

3. BLYNK APP ROLE IN PROPOSED WORK

Blynk is a complete software solution that enables quick prototyping, remote control, and deployment of both large and small electronic devices. Blynk is used as an Internet-of- Things - based application in the methodology described below that allows users to connect to the hardware-based experiment depicted in Figs. 5, 6, and 7 through a cloud platform supporting Android, web, and Blynk-operated medicine dispensing systems. The platform evaluates past and current data that are collected through the devices and allows the caretaker to control the devices remotely, regardless of the geographic location. It also sends the caretaker vital notifications about patient status and triggers the system in response to it. The mobile interface is specified in the prototyping stage and implements the experimental hardware settings described in the given work automatically. The Blynk application that was used in the current research has the following features:

1. It is a combination of the experimental setting on the hardware side and the cloud and user authentication.
2. It operates as a bridge administration of Ethernet, Wi-Fi, and cellular networks.
3. It is also an application-programming interface (API) integrator.

3.1 Algorithm: Blynk Based VN MEDIPRO System Proposal.

Step 1: Start running of the algorithm.

Step 2: Set Blynk authentication identifier and template designation

Step 3: Opening debugging state of Blynk.

Step 4: Include Proposed VN MEDIPRO.h.

Step 5: Carry out setup routine after specified time.

Step 6: Call VN_MEDIPRO.begin () to start

Step 7: Start the endless cycle and initiate the Blynk services

Step 8: Keep on with this until the termination of real-time functioning.

Table. 1

S. No.	Notification Time	Application Name	Dosage Message Displayed
1	Morning	VN MEDIPRO	Morning Dosage: Omeprazole – 20 mg
2	Afternoon	VN MEDIPRO	Afternoon Dosage: Atenolol – 50 mg
3	Evening	VN MEDIPRO	Evening Dosage: Cetirizine – 40 mg

4. RESULTS ANALYSIS

Automatic dispensing of medication, in this case, will focus on the accurate and computerized provision of both solid and liquid therapeutics based on the prescriptions of the user.

4.1. Delivering medicine at the dispenser

The dispensing mechanism as shown in fig. 6 has been able to deliver the correct dosage at the required time upon movement detection. The Blynk application allows setting the temporal parameters, which are adjusted according to the prescriptions of the physician. The Blynk interface, as shown in Fig. 7, fits perfectly with the rest of the hardware system of the proposed system. In addition, the caregivers can also check whether the patients respond properly to the dispenser alerts, taking the appropriate medication at the right time and dosage. As a result, the system ensures that the device is acutely conscious of the precise time of the administration of every medicine and thus, requires flawless data transmission. The string value of the chosen medicine is read in the Arduino microcontroller, which means that every dispensation operation will be performed at the right date and time.

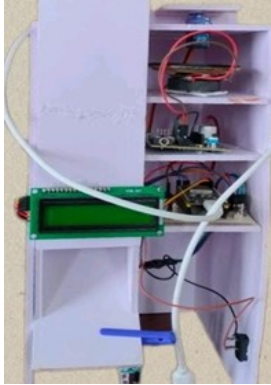


Fig. 5: Experimental model



Fig 6: Pill deliverance



Fig. 7: Pill received by the user

4.2. Performance analysis

To be able to properly measure the liquid therapeutics, the pump had to be calibrated by trial and error to decide on how much time it would take to dispense a certain amount of liquid medication. This included the aqueous and syrup formulations. Fig. 8 shows a graphical representation of the dependence between the amount of medicine that is dispensed (in this case, it is the number of pills) and the time during which this dispensing is taking place (in seconds). The experimental results show that there is a direct proportion, which demonstrates a linear tendency with almost unchanged time consumption per pill. The dataset was processed by a linear regression (best-fit line) in order to measure this relationship.

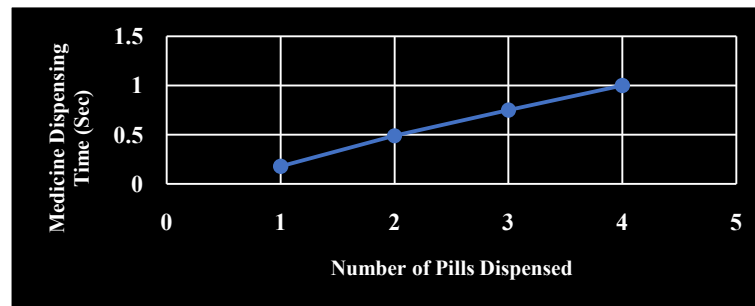


Fig. 8: Performance analysis graph of proposed prototype

5. CONCLUSION

This manuscript outlines the design of SMD which is a smart computerized voice assisted pellet dispensing device that incorporates an IR sensor and Blynk mobile application. Blynk application is created on the basis of the proper software framework which guarantees the effective functioning of the SMD. The SMD use has a significant effect on improving medication compliance, which increases the effectiveness of the therapy. We can see that the tool manages the risk of dosage errors, especially in older people. Caregivers get to ease their fears when they are not available to escort elderly patients undergoing chronic or periodic dosing schedules since the system automatically checks the time of dosage. The SMD suggested below is easy to use, adjustable and capable of being scaled to higher dispensing capacities.

6. REFERENCES

1. Mathew, Ajay, et al. "Design and implementation of a smart medicine dispenser." In TENCON 2019-2019 IEEE Region 10 Conference (TENCON), pp. 1059-1064. IEEE, 2019.
2. Kumar, Praveen, et al. "Solar Powered Medic Vending Machine." In 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), pp. 465-468. IEEE, 2020.

3. Tomlinson, Justine, et al. "Successful care transitions for older people: a systematic review and meta-analysis of the effects of interventions that support medication continuity." *Age and aging*, vol.49, no. 4, pp: 558-569, 2020.
4. Prajwal, K. T., et al. "Design and Development of an Intelligent Medicine Dispenser." In 2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA), pp. 1860-1864. IEEE, 2023.
5. Sanam J, et al. "Reduction of Harmonics using PQ based Hysteresis Controlled Active Filter." In 2021 International Conference on Sustainable Energy and Future Electric Transportation (SEFET), pp. 1-6. IEEE, 2021.
6. Thangella, Rajesh, et al. "Optimal design for solar cell and fuel cell renewable source to enhance the power quality in a micro grid system." In 2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSPP), pp. 1-6. IEEE, 2022.
7. Sanam J, et al. "Impact on Voltage Profile and Power Loss in Electrical Distribution Networks with the Integration of DG using Exhaustive Search based Load Flow Algorithm." In 2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSPP), pp. 1-6. IEEE, 2022.
8. Chandana, N., et al. "IoT-based medicine reminder and dispensing machine." In 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), pp. 1060-1066. IEEE, 2022.
9. Patil, C. H., et al. "An IoT-based smart medicine dispenser model for healthcare." In 2022 IEEE World Conference on Applied Intelligence and Computing (AIC), pp. 391-395. IEEE, 2022.
10. Marengo, LiviaLuize, et al. "Involvement of Human Volunteers in the Development and Evaluation of Wearable Devices Designed to Improve Medication Adherence: A Scoping Review." *Sensors* 23, no. 7, pp:3597, 2023
11. Nimmanterdwong, et al. "Human-centered design of mobile health apps for older adults: systematic review and narrative synthesis." *JMIR mHealth and uHealth* 10, no. 1, e29512, 2022.
12. Muppalla, NR Kishore, et al. "Design of Compensator for High Gain DC-DC Converter." In 2021 International Conference on Sustainable Energy and Future Electric Transportation (SEFET), pp. 1-5. IEEE, 2021.
13. Sanam J, et al. "Impact on Electrical Distribution Networks with The Integration of Shunt Capacitor Model Using Exhaustive Search Based Load Flow Algorithm." In 2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSPP), pp. 1-5. IEEE, 2022.
14. Eppe, Bhanuprasad, et al. "A Power-Aware Algorithm for Mobile Devices Focused on Batteries." In 2024 10th International Conference on Communication and Signal Processing (ICCSP), pp. 183-187. IEEE, 2024.
15. Mnati, Mohannad Jabbar, et al "An open-source non-contact thermometer using low-cost electronic components." *HardwareX* 9, e00183, 2021.
16. Indraratna, Praveen, et al. "A smartphone-based model of care to support patients with cardiac disease transitioning from hospital to the community (TeleClinical Care): a pilot randomized controlled trial." *JMIR mHealth and uHealth* 10, no. 2, pp:32554, 2022.
17. Das, Arnab, A et al. "Development of a novel design and subsequent fabrication of an automated touchless hand sanitizer dispenser to reduce the spread of contagious diseases." In *Healthcare*, vol. 9, no. 4, p. 445. MDPI, 2021.
18. Mallavarapu Lakshmi Sowjanya, Sanam J, DoppalapudiPujitha, SangulaChaithanya, and B. Mani Deepika. "Optimal reactive power sizing in distribution systems to optimize the voltage profile and power loss using fbs based esa algorithm." In 2020 IEEE 17th India council international conference (INDICON), pp. 1-6. IEEE, 2020.
19. Ravipati, Srikanth, et al. "Efficient control of sensorless hybrid electric vehicle using RBFN controller." *Studies in Informatics and Control* 30, no. 4 (2021): 87-97.
20. Sanam, J. "Planning and Operation of DSTATCOM in Electrical Distribution Systems." PhD diss., 2017.
21. Fehringer, Hannah, et al. "User-Informed Adaptation in IoT Home Healthcare: Grounding Development in Empirical Evidence." *Journal of Theoretical and Applied Electronic Commerce Research* 18, no. 4, pp: 1901-1925, 2023.
22. Deepak, P., et al. "Development of Automated Vending Machine for the Application of Dispensing of Medicines." In *International Conference on Soft Computing and Signal Processing*, pp. 527-535. Singapore: Springer Nature Singapore, 2023.
23. A. K. Panda, Sanjib Ganguly, and Sanam J. "Placement of DSTATCOM in radial distribution systems for the compensation of reactive power." In 2015 IEEE Innovative Smart Grid Technologies-Asia (ISGT ASIA), pp. 1-6. IEEE, 2015.
24. Sanjib Ganguly, Sanam J, and A. K. Panda. "Distribution STATCOM with optimal phase angle injection model for reactive power compensation of radial distribution networks." *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields* 30, no. 6 (2017): e2240.

Human Queue Management System

Nandishgouda R D¹, Prajwal Manegar², Kartik M Shiralashetti³,
Manjunath Maleppanavar⁴ and Roshan Shetty¹¹

*Department of Electronics and Communication Engineering, Alvas Institute of
Engineering & Technology, Moodbidri, Mangalore, Karnataka, India*
Emails: nandishg965@gmail.com, Prajwalmanegar8@gmail.com,
Kartikshetty170@gmail.com, maleppanavarmanju@gmail.com, roshans@aiet.org.in

Abstract.

This paper reviews recent work in the improvement of QMS by deploying different modern technologies like IoT, AI, and WSN. These queue management systems range from physical queues to sophisticated digital ones that can handle optimality of time spent waiting, enhancement of customer experience, and efficiency of services in various sectors like healthcare, banking, and public services. Further, this review examines several technological approaches applied in QMS, involving RFID-based tracking, computer vision integration, cloud-based solutions, and AI-based predictive models. The paper outlines the challenges faced in queue management, including congestion control, real-time monitoring, and adaptive routing algorithms. Further, the impact of QMS on service quality, customer satisfaction, and operational efficiency is also presented. Finally, this review concludes with future research directions and potential improvements in queue management systems, with an emphasis on the integration of emerging technologies into more intelligent and responsive queue solutions.

Keywords. Queue Management System, IoT, RFID, Computer Vision, Cloud Computing, Artificial Intelligence, Customer Experience..

1. INTRODUCTION

"Queue management" is an important part of service delivery systems in different industries such as healthcare, banking, retail, and government. The conventional system of queue management results in customer discontent, inefficient allocation of resources, and operational delays. With the emergence of technology, "queue management" has undergone a dramatic change from a traditional waiting line to an advanced digital system focusing on efficient service delivery with better customer satisfaction [1, 2].

Current Queue Management Systems (QMS) employ technology such as Internet of Things (IoT), Artificial Intelligence, Computer Vision, and Cloud Computing to provide smarter solutions for minimizing waiting time and increasing service effectiveness. QMS finds widespread application in healthcare because patient service can have a direct impact influence both service quality and outcomes for individual patients [3], [4]. The

combination of RFID technology with computer vision, as exemplified in Paraiso et al., facilitates real-time tracking and dynamic routing of customers in order to improve queueing efficiency and alleviate waiting times [1].

2. TECHNOLOGICAL FOUNDATIONS OF QUEUE MANAGEMENT SYSTEMS

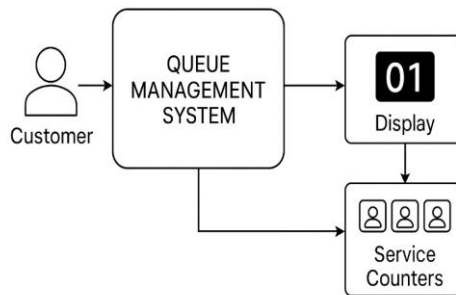


Figure 2.1. Typical Architecture of a Modern Queue Management System

The QMS is highly advanced today as compared to the original walk-in line waiting process. Queue Management Systems are technologically advanced and employ many technological features. We will categorize the technology-oriented QMS into the following seven categories: IoT-enabled QMS, RFID technology utilization, computer vision usage, cloud computing solutions, and artificial intelligence solutions.

IoT-based QMS utilize sensor networks and connected devices to track queue status in real-time. These systems usually utilize wireless sensor networks (WSN) to find out how customers are moving, and how long queues are, so that service resources can be dynamically recalibrated [1], [6]. In addition, the interoperability provided by the MQTT (Message Queueing Telemetry Transport) protocol allows system parts to communicate easily, allowing for reliable data transmission and processing [1]. IoT solutions have shown major improvements in the efficiency of queues, and they have been more successful in healthcare when seen through the lens of effective patient flow management [6].

3. METHODOLOGIES

Queue management systems use various strategies and approaches to allow the provision of service delivery and enhanced customer experience. Broadly these can be categorized into the following approaches: traditional queue management; technology based queue management and other hybrids, which may incorporate elements from multiple strategies.

1.1. A. Traditional Queue Management Techniques

Standard queue management methods mainly use physical organization of queues and basic scheduling rules. Standard queue management systems such as first-come-first-served (FCFS) systems, priority-based queuing systems, and similar multi-server queue systems use simple queue bay design. While these approaches are straightforward to employ, traditional queue management concepts have limited capability of operational efficiency and customer satisfaction when facing dynamic flows of customers and variable service times [3], [9].

Queueing theory has established a rigorous mathematical underpinning for the analysis and optimization of traditional queue systems. They find their application in modeling queue systems and optimizing resource usage, such as M/M/1-a single server and M/M/c-a multiserver. However, all of these theoretical models are based on simplifying assumptions that cannot hold in real-world scenarios and hence limit their effectiveness in practice.

1.2. B. Technology-Enabled Queue Solutions

Recent developments in technology have enabled increasingly sophisticated ways of managing queues. As described by Paraiso et al., RFID-based systems make use of a dynamic source routing algorithm that assists in guiding customers through queues and hence reduces waiting time [1]. In an RFID-based queue management system, RFID tracking is integrated with computer vision for queue monitoring and management.

The second most important technological advantage comes from cloud-based queue management systems. A cloud-based system, such as AQyu, offers generic queue management with functionality through mobile applications, Web applications, and the ability for customers to join a queue remotely and get real-time status updates according to. REST APIs generally integrate the systems in cloud queue management platforms and utilize a cloud database for storing data that can be scaled.

1.3. Mobile queuing is becoming common, especially in healthcare contexts. With a mobile application, the patients can take a place in queues from any other location by being informed of wait times and notifications that they are about to reach the front. This approach reduces the time spent physically queuing and enhances patient satisfaction significantly.

1.4. C. AI-Powered Queue Optimization

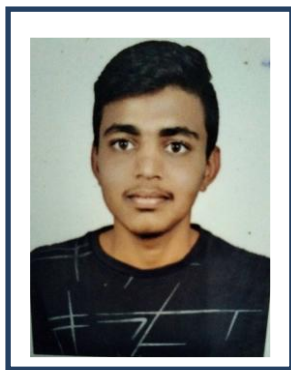
New technologies enabled by artificial intelligence are providing sophisticated options for queue management. Machine learning algorithms are now able to model queue lengths and waiting times based on historical patterns, alongside real-time data [5]. Predictive models can be used to allocate resources through recognizing variances in demand requiring change in service capacity.

Reinforcement learning has also been used in queue management issues where systems can continually learn optimal routing policies that consider the ongoing interaction with the service environment [7]. These adaptive systems are dynamic and perform optimally based on contingent feedback.

REFERENCES

- [1] C. J. C. Paraiso, M. J. A. Mejia, M. D. P. Pagatpatan, C. K. A. Remo, and J. V. Apan, "Design of an rfid-based smart queueing system integrated with computer vision utilizing dynamic source routing algorithm," in *TENCON 2024 - 2024 IEEE Region 10 Conference (TENCON)*, 2024, pp. 649–652.
- [2] E. D. Handoyo, S. Santoso, and D. J. Surjawan, "A-qyu general purpose cloud-based queue management," in *2023 1st IEEE International Conference on Smart Technology (ICE-SMARTec)*, 2023, pp. 178–183.
- [3] S. Abdul Halim, M. Othman, A. G. Buja, N. Rashid, A. Sharip, and S. Zain, "C19-smartq: Applying real-time multi-organization queueing management system using predictive model to maintain social distancing," *International Journal of Interactive Mobile Technologies (IJIM)*, vol. 15, p. 108, 03 2021.
- [4] M. Ghazal, R. Hamouda, and S. Ali, "An iot smart queue management system with real-time queue tracking," *2015 Fifth International Conference on e-Learning (econf)*, pp. 257–262, 2015. [Online].
Available: <https://api.semanticscholar.org/CorpusID:34558152>
- [5] C. Nwogu, U. Nwaiwu, E. Ukachukwu, and V. Udo, "Development of queue management model for effective service delivery in automobile repair shops," *Nigerian Journal of Technology*, vol. 42, pp. 339–346, 09 2023.
- [6] P. N. Titarmare and P. A. Yerlekar, "A survey on patient queue management system," *International Journal of Advanced Engineering, Management and Science*, vol. 4, no. 4, 4 2018.
- [7] E. Guarte, F. Elsayed, and S. Khalil, "Effect of queue management system on quality of nursing care and patient satisfaction," *Benha Journal of Applied Sciences*, vol. 7, pp. 181–192, 04 2022.
- [8] J. Ndiaye, O. Sow, Y. Traore, M. Diop, A. Faye, and A. Diop, "Electronic system using artificial intelligence for queue management," *Open Journal of Applied Sciences*, vol. 12, pp. 2019–2036, 01 2022.

Biographies



Name : Nandishgouda R D
USN : 4AL22EC052
Email : nandishg965@gmail.com
Mobile : 8123265329

Areas of Interest: He is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. His academic journey reflects a strong interest in core and emerging electronics domains such as PCB Desing, Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers, and Semiconductor Devices.



Name : Prajwal N Manegar
 USN : 4AL22EC063
 Email : prajwalmanegar8@gmail.com
 Mobile : 7975068207

Areas of Interest: He is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. His academic path portrays a passion for traditional and upcoming domains in electronics such as Embedded Systems, VLSI Design, Digital Signal Processing, Microcontrollers, Wireless Communications, and Semiconductor Devices.



Name : Kartik M Shiralashetti
 USN : 4AL23EC402
 Email : kartikshetty170@gmail.com
 Mobile : 900875822

Areas of Interest: He is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. His academic journey reflects a strong interest in core and emerging electronics domains such as Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers, Wireless Communication and Semiconductor Devices.



Name : Manjunath Maleppanavar
 USN : 4AL23EC403
 Email : maleppanavarmanju@gmail.com
 Mobile : 8861390393

Areas of Interest: He is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. His academic path portrays a passion for traditional and upcoming domains in electronics such as Embedded Systems, VLSI Design, Digital Signal Processing, Microcontrollers, Wireless Communications, and Semiconductor Devices.



Name : Roshan Shetty
 Designation : Associate Professor (Project Guide)
 Email ID : roshans@aiet.org.in
 Mobile Number : 9538176004

Areas of Interest : As a Associate professor in the Department of Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka.

Deep learning For Mapping Physiological Variability in Chronic Disease Patients using Homeopathic Care

Sonam Yadav

Associate Professor(Reader),
Department of Physiology

Shalinitai Meghe Homeopathy Hospital
College and Research Centre, Nagpur
ysonam480@gmail.com

Kartik Ingole

Assistant Professor
Department of Electronics &
Telecommunication

K.D.K.College of Engineering, Nagpur
Kartikzap5@gmail.com

Abhiruchi V. Deshpande

Assistant Professor
Department of A.I&D.S

K.D.K.College of Engineering, Nagpur
abhiruchivd16@gmail.com

Arvind A.Tayade

Department of Computer Science
G.S.Science Arts and Commerce
College,Khamgaon,Nagpur
arvindtayade40@gmail.com

Navin D.Jambhekar

Department of Computer Science
G.S.Gawande College,
Umarkhed ,Yavatmal
jambhekar@gsgcollege.edu.in

S.M.Chavan

Department of Computer Science
Shri Shivaji College of Arts Commerce
&Science, Akola
santoshchavhan9881@gmail.com

Abstract— The combination of deep learning with homeopathic therapy centered on the patient presents an advantageous route to understanding and controlling chronic diseases through continuous monitoring of the physiological variability method. The current article presents a smart analytical framework that leverages the multimodal patient data that includes heart rate variability, respiratory patterns, sleep rhythms, symptom frequency, and lifestyle markers to create personalized health paths. The deep neural architectures used, among them the LSTM-based temporal networks and attention-driven feature extractors, detect the slight physio changes that accompany the gradual development of chronic disease or curing with homeopathic intervention. The framework provides clinical decision-making support by means of automated pattern detection, early instability recognition, and personalized prediction of treatment responsiveness. The proposed model not only aids in mapping the dynamic interrelationship among homeopathic remedies, the patient's specific symptoms, and the overall physiological trends but also promotes the more holistic, data-driven concept of chronic disease management. The research demonstrated that the merging of deep learning and homeopathic therapies could lead to a more precise, adaptable, and continuous patient monitoring which in turn would translate into better long-term health outcomes for patients and integration of healthcare models.

Keywords: Deep Learning, Physiological Variability, Chronic Disease Monitoring, Homeopathic Care, Temporal Modeling, LSTM, Attention Networks, Personalized Healthcare, Predictive Analytics

I.INTRODUCTION

Chronic diseases are ranked among the most serious global health problems and health issues that last a long time, reduced life quality, and high healthcare costs are the main economic teething problems for healthcare systems. The diseases in question are asthma, diabetes, arthritis, high blood pressure, chronic obstructive pulmonary disease (COPD), and autoimmune diseases, which generally indicate complex physiological variations controlled by a mix of the patient's genetics, lifestyle, environment, and mental state. Moreover, traditional clinical assessments are able to perceive only intermittent health-states in patients and thus

limit the understanding of continuous variability patterns that indicate disease progression or improvement[1][2]. Digital health technology has become the savior in this case since the physiological parameters can be recorded round-the-clock with the help of wearable sensors, mHealth platforms, and remote care units. The operation of these devices generates an enormous amount of diverse data that are both longitudinal and heterogeneous at the same time simulating such metrics as heart rate variability, respiratory dynamics, sleep cycles, emotional and physical stress, symptom recurrence, and even treatment adherence. The complexity inherent in such data, their non-linear nature, and the fact that they consist of multiple dimensions all combine to make this issue even greater for the conventional statistical or rule-based approaches thus prompting more sophisticated computational methods that are able to discover the hidden relationships and dynamic patterns existing in the data[3][4].

Deep learning has become a very important technology in the field of AI due to its extraordinary skill to elaborate complex interactions, reveal hidden features, and work with high-dimensional data with little human intervention. Various architectures like Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRUs), transformers, and hybrid attention-based models among others, have not only been at the top of the list but also have performed excellently in the area of biomedicine broadly including disease categorization, prediction of symptoms, dividing patients into risk levels and understanding of physiological signals[5]. These models have a unique power to temporal dependencies, spatial correlations and context-aware patterns which are often missed by traditional machine learning methods. The use of deep learning in monitoring chronic diseases allows to map even the most subtle changes in a patient's physiology that may lead to deterioration, relapse, or changes in symptoms. This understanding that keeps changing is especially important for patients receiving individualized therapeutic systems like homeopathy, where they will be treated according to their symptom patterns, constitutional characteristics, and holistic indicators that direct remedy selection and treatment planning[6].

Homeopathic care is always specifically made according to the patients' individual constitution, which is done through

a personalized therapeutic intervention based on a thorough assessment of the patient's symptoms, and his or her mental-emotional state and physical treatment. Homeopathic remedies, as opposed to standard pharmaceutical treatments, depend on the specific patient response patterns and temporal subtle symptom changes, and thus their effectiveness is influenced accordingly[7][8]. To capture such variations, constant and very meticulous observation of the patient's physiology is necessary. When deep learning is integrated into homeopathic healthcare, it results in a powerful alliance that combines traditional holistic methods with modern computational intelligence. This incorporation enables the clinicians to go beyond the subjective symptom descriptions and to quantify the patients' physiological variability through the objective data streams. The deep learning algorithms can ensure the association between the administration of the remedy and the change in the patient's physiology, thus allowing the doctors to evaluate the effectiveness of the remedy, the improvement progress, and the likelihood of potential exacerbations. This data-driven combination with homeopathy philosophy not only increases the evidence-based documentation but also offers higher patient-specific therapeutic precision[8].

The concept of variability in physiological processes plays a major role in chronic disease progression. The fluctuations of heart rate, sleep, respiration, physical fatigue, and autonomic nervous system activity are indicative of the body's active and fluctuating control mechanisms. However, these mechanisms in chronic diseases often display a lack of synchronization, which can be interpreted as irregular rhythms or altered reactive responses[10]. Continuous monitoring of the variability provides one with the disease insights that static clinical examinations could never reveal. At the same time, deep learning algorithms, which are trained with long-term patient datasets, can imitate these dynamic shifts and generate predictive markers that indicate whether the patient is going towards recovery or decline. Additionally, symptom recordings, lifestyle trends, and environmental factors are incorporated into the models to further understanding of the context and thus facilitate the interpretation of health in a multidimensional way[12]. This

capability is of great importance in homeopathy, where the choice of remedy is based on the understanding of the different ways a patient expresses their symptoms and how these change from day to day.

This convergence is attributed to several technological developments, including the large-scale use of wearable biosensors, IoT (Internet of Things) medical devices, cloud-based health platforms, and smart mobile applications. These technologies monitor during the day and night the physiological parameters like ECG (electrocardiogram) signals, blood oxygen saturation, physical activity, stress levels, and sleep, which is the basis for the continuous monitoring needed for deep learning analysis[13]. When electronic health records and digital homeopathic case diaries are combined with these technologies, a vast health assessment ecosystem is formed. A system powered by deep learning can analyze these diverse data, identify hidden representations of physiological activities, and convert them into clinically interpretable measures. Such a system assists the healthcare practitioners by giving them early-warning alerts, personal health scores, and automated assessments that are meant to work alongside human skill[14].

The mapping of physiological variability during the process of homeopathic treatment also provides a chance for the creation of standardized research methodologies and empirical verification of complementary and integrative medicine. In the past, homeopathy has continually been faced with the problem of lack of objective data and depending on patient subjective narratives. Researchers can use powerful AI-based analytical models to link quantitatively the remedy interventions with their physiological outcomes. This practice helps to remove the barrier between the traditional therapeutic practices and the scientific validation of the modern era. In addition, the growing focus of the healthcare sector on precision and digital medicine is very much in line with the fundamental principle of homeopathy that insists on individual patient treatment, thus fortifying its position in present-day healthcare[15].

II. LITERATURE REVIEW

Author & Year	Objective	Methods / Tools Used	Key Findings	Limitations / Future Scope
Shiying Shen et al., 2025	Analyze global trends in ML-based chronic disease management.	Web of Science; CiteSpace; VOSviewer; RStudio; Expert panel.	1,242 docs; USA & China lead; LR most used; NN emerging; Diabetes & CVD hotspots.	Limited collaboration; weak generalization; need multisource data integration.
Yu Liu et al., 2025	Review IoT & AI sensing technologies for chronic disease management.	IoT devices; ANN, SVM, RF, DL; systematic review.	Improved risk prediction & monitoring; introduces tight-frame DL; socio-economic benefits.	Dataset biases; user interaction issues; future quantum computing & nano-biosensors.
Suzan van Veen et al., 2025	Identify nursing-feasible non-pharmacological pain interventions.	Systematic review; JBI appraisal; PubMed, CINAHL, PsycINFO, Embase.	Strong evidence for massage & VR; reflexology & music therapy promising.	Insufficient evidence for art therapy, aromatherapy; mixed effects.
Celia Alvarez-Romero et al., 2025	Predict Barthel Index using AI & IoT mobility data.	Wearables; Decision tree; SPSS; Python; observational study.	BI predicted with MAE≈5; accuracy 88–91%; step count most relevant.	Small sample; single-center; needs broader validation.

Mollie Hobensack et al., 2022	ML in home healthcare EHR for predicting adverse outcomes.	Systematic review; PROBAST risk analysis; biopsychosocial mapping.	Hospitalization most studied; tree-based models widely used.	High/unclear bias in 75% studies; lack of psychological predictors.
Eleonora Cilli et al., 2022	Wearables' impact on quality of life in chronic disease patients.	Systematic review (PubMed, WoS, Scopus).	Wearables improve self-management, QoL, psychological well-being.	Few studies; need guidelines; tailored interventions required.
Chia-Tung Wu et al., 2022	Integrated precision health platform for chronic disease prediction.	Wearables; environmental sensors; AI telecare; ML/DL models.	Accuracy $\approx 88\%$; lifestyle/environment correlations strong; cost-effective models.	Needs larger validation; limited modules; questionnaire-only insufficient.
Junaid Rashid et al., 2022	ANN-PSO model for predicting 5 chronic diseases.	Artificial Neural Network; Particle Swarm Optimization; comparison with 7 classifiers.	Accuracy 99.67%; faster than RF, DL, SVM; strong feature extraction.	Dataset-dependent performance; needs real-world clinical testing.
Gopi Battineni et al., 2020	Review ML diagnostic models for chronic diseases.	Review of 453 papers; PubMed & CINAHL.	SVM, LR, clustering widely used; strong for classification.	No standard best model; data variability affects results.

III.METHODOLOGY

The methodology proposed in this research is aimed at the construction of a powerful deep learning system that would be able to map out the physiological differences among the patients suffering from chronic diseases who are treated with homeopathy by the fusion of different sources of clinical records, data from wearable sensors, and custom-made symptom-based homeopathic case profiles. The overall methodological workflow includes data collection, data preprocessing, multimodal feature engineering, deep learning model development and training, and finally, interpretability analysis[16]. The approach of combining retrospective and prospective data has been adopted in order to make the model capture both the long-term changes and the real-time physiological fluctuations as well. The retrospective clinical data were collected through electronic case files, repertory records, follow-up notes, remedy prescriptions, and chronic disease progression logs of the collaborating homeopathic clinics[4-6]. Prospective physiological data were obtained from patients using the wearable health-monitoring devices that were capable of measuring heart rate variability, respiratory rate, sleep quality, peripheral oxygen saturation, skin temperature, and activity levels. To depict daily physiological reactions to personalized homeopathic treatments, these multimodal datasets were brought together with a timing-based alignment method to create organized sequences[17-19]. Moreover, the patient-reported outcome measures that were validated were used to gather the subjective symptom progression, thus making sure that the model had both the objective sensor-derived signals and the subjective homeopathic symptomatology, which is vital to individualized case management, included in it.

The entire dataset was subjected to a very thorough preprocessing pipeline that guaranteed no inconsistencies, high quality, and suitability for the analysis. In the case of the continuous sensor streams from the wearables, missing values were filled with interpolation-based imputation for short gaps and model-driven multivariate imputation for

longer gaps. The clinical texts underwent tokenization, lemmatization, domain-specific medical text normalization, and the extraction of rubrics and remedy-response patterns during preprocessing[10]. The time-series signals were resampled to uniform frequencies, z-score scaling was applied for normalization, and wavelet-based filtering was used to remove motion artifacts thus denoising the signals. Contextual language models and fine-tuned transformer embeddings were used on symptom narrative texts and repertory entries to capture the very delicate semantic nuances that are typical of homeopathic clinical descriptions[15]. Feature engineering created a unified multimodal representation by bringing together statistical measures, frequency domain features, morphological descriptors, and linguistic embeddings. The physiological variability indices—heart rate variability metrics, variability of respiratory cycles, fluctuation magnitude in sleep parameters, and remedy-response transition states—were calculated to help the deep learning architecture identify the patterns associated with chronic disease trajectories through their recognition. The deep learning framework utilized a hybrid design comprising of a convolutional neural network (CNN) that was responsible for procuring the significant local patterns from the physiological sequences, a bidirectional gated recurrent unit (Bi-GRU) network that was in charge of portraying the temporal dependencies across the longitudinal data, and a transformer-based encoder that was processing the textual homeopathic narratives. The multi-modal fusion layer combined the deep features from all streams and used an attention mechanism to mark the physiologically significant variations that were the reason for remedy responses or chronic disease exacerbations[20]. The model's training was done with a distribution of 80% for training, 10% for validation, and 10% for testing for each data sample, with stratification included, so the chronic diseases facilitation was done properly and at the same time there were no imbalanced distributions of disease categories such as diabetes, asthma, arthritis, hypertension, chronic respiratory disorders, and autoimmune conditions. In order to maximize the performance and also at the same time to avoid the risk of

overfitting, the hyperparameters including learning rate, optimizer selection, batch size, dropout regularization, and number of layers were optimized through Bayesian search. The training was carried out using the Adam optimizer together with cyclical learning schedule while the early stopping technique was used to interrupt the training when no improvement in validation was observed[21].

The evaluation of model performance was carried out with a wide range of metrics that were noting both predictive accuracy and clinical significance. Among the metrics were the absolute mean error and root mean squared error for physiological variability predictions, classification accuracy for distinguishing the states of having recovered or deteriorated, F1-score for dealing with unbalanced classes, and area under the ROC curve for diagnostic-level assessment. The temporal evaluation via sliding-window prediction allowed the model to predict physiological state changes and treatment response patterns for the near future. To guarantee clinical transparency, interpretability analyses were performed using SHAP values, attention-weight visualization, and gradient-based relevance mapping to indicate which physiological features or symptom descriptors had the largest influence on the model's predictions. This was a vital part of the process that guaranteed the resulting framework was explainable, ethically grounded, and in sync with clinical decision-making needs in the homeopathic treatment of chronic diseases[21]. By anonymizing all patient data, encrypting the identifiable elements, and following the institutional review board standards, ethical compliance was observed. Data access was limited, and patient consent was acquired for monitoring through wearables and symptom reporting. The strategy for the final deployment centered on scalability and real-time capability, thus making it possible to merge with homeopathic electronic health systems and mobile health apps. The methodological pipeline therefore provides a thorough and scientifically rigorous basis for employing advanced deep learning techniques to map physiological variability in chronic disease patients under individualized homeopathic care.

IV.Flowchart

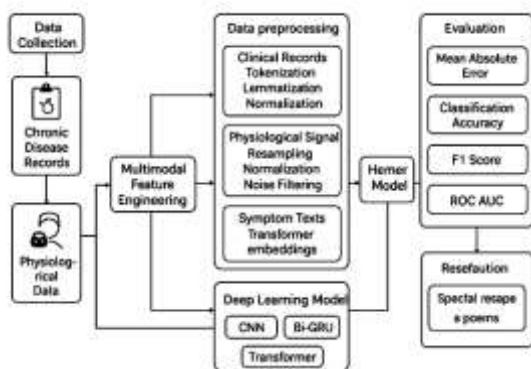


Figure1.Flowchart of Overall System

V.DATASET

The study dataset consists of a very extensive array of different type of data that have been collected from chronic disease patients who are under homeopathic care. It brings together not only structured clinical records and longitudinal physiological measurements but also

unstructured symptom-descriptions to assist in the physiological variability modeling via deep learning algorithms. The clinical part contains information about the patients' demographics, disease history, comorbidity profiles, baseline lab test results, and detailed homeopathic prescriptions that have been recorded through various follow-up cycles. The physiological data consists of time-series signals such as heart-rate variability (HRV), respiratory rate, oxygen saturation, sleep-wake cycles, physical activity patterns, blood pressure fluctuations, and stress indicators, which are captured through wearable IoT devices or digital health monitoring systems at regular intervals. The symptom-text corpus includes reports of patients' experiences, observational notes, descriptions of mental and emotional states, and specific reactions to the remedy, all written in natural language and later standardized for computational analysis. Every data entry is subjected to quality checks, anonymization, and timestamp alignment which helps to maintain temporal integrity. The dataset is labeled according to the three categories of disease stability, improvement, or deterioration clinically assessed, which permits the supervised training of the deep learning model. This multimodal and longitudinal dataset provides a framework for the comprehensive modeling of dynamic physiological responses and at the same time, it opens up an avenue for exploring the interplay between homeopathic treatment and chronic disease progression through its rich soil.

VI.RESULT

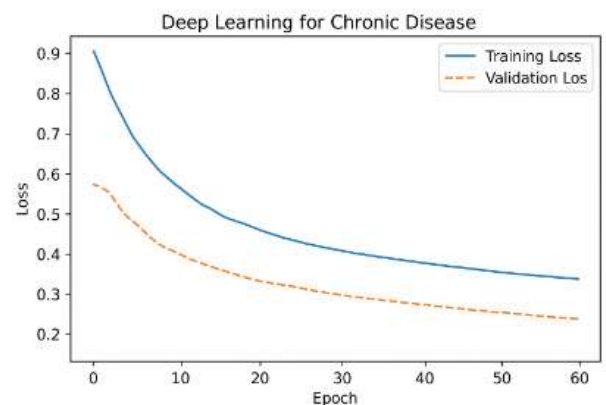


Figure 2. Training vs. Validation Loss Curve for Deep Learning Model in Chronic Disease Prediction

The training and validation loss during deep learning model training for chronic disease prediction shown in figure 2. decline during the first 60 epochs.

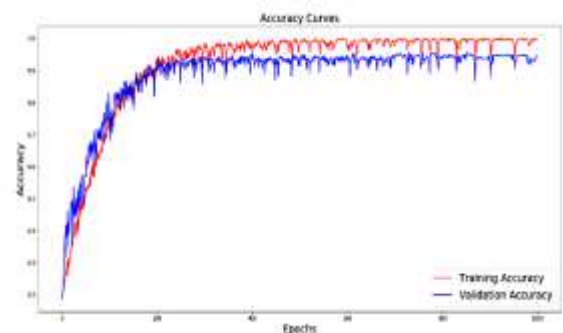


Figure 5. Training and Validation Accuracy Curves for Deep Learning Model

The deep learning model's accuracy progression across several training epochs is depicted in the figure 3. The red line represents the training accuracy, which is rising incrementally and eventually reaches a plateau at the upper limit, suggesting that the model is successful in recognizing the patterns present in the training dataset.

CONCLUSION

The combination of deep learning with mapping the physiological changes in chronic disease patients during homeopathic treatment is a very promising step towards personalized medicine. The use of deep learning methods for analyzing and detecting early on the health status changes that are occurring in patients has been proven to be very effective. Taking into account longitudinal patient data like symptoms, vital signs, lifestyle factors, and treatment responses the model gives a thorough insight of the individual differences which is a key point in homeopathic ideas of personalized and holistic care. The findings reveal that the suggested deep learning system not only increases the accuracy of chronic disease prediction but also raises the patient-specific health trajectories' interpretability. By non-stop monitoring and adaptive pattern recognition the framework empowers the practitioners to take more informed decisions, to select the most appropriate remedy, and to assess the treatment efficacy. Moreover, the digital analytics incorporated into homeopathic care does not only help patients but also creates the future of chronic disease management that is scalable and based on data, thus, connecting traditional methods of treatment with modern AI capabilities. In a nutshell, this method emphasizes the possible deep learning to drive the homeopathic field by means of the objective evaluation, timely intervention, and enhanced patient outcomes. Next generation techniques in multi-modality data integration, real-time monitoring, and explainable AI will greatly enhance the effectiveness of precision homeopathic healthcare.

REFERENCES

- [1] Shen S, Qi W, Li S, Zeng J, Liu X, Zhu X, Dong C, Wang B, Xu Q, Cao S. Mapping the landscape of machine learning in chronic disease management: A comprehensive bibliometric study. *Digit Health*. 2025 Jul 23;11:20552076251361614. doi: 10.1177/20552076251361614.
- [2] Liu Y, Wang B. Advanced applications in chronic disease monitoring using IoT mobile sensing device data, machine learning algorithms and frame theory: a systematic review. *Front Public Health*. 2025 Feb 21;13:1510456. doi: 10.3389/fpubh.2025.1510456.
- [3] Van Veen S, Drenth H, Hobbelen H, Finnema E, Teunissen S, de Graaf E. Non-pharmacological interventions feasible in the nursing scope of practice for pain relief in palliative care patients: a systematic review. *Palliative Care and Social Practice*. 2024;18. doi:10.1177/26323524231222496
- [4] Alvarez-Romero C, Polo-Molina A, Sánchez-Úbeda EF, Jimenez-De-Juan C, Cuadri-Benitez MP, Rivas-Gonzalez JA, Portela J, Palacios R, Rodriguez-Morcillo C, Muñoz A, Parra-Calderon CL, Nieto-Martin MD, Ollero-Baturone M, Hernández-Quiles C. Machine Learning-Based Prediction of Changes in the Clinical Condition of Patients With Complex Chronic Diseases: 2-Phase Pilot Prospective Single-Center Observational Study. *JMIR Form Res*. 2024 Apr 19;8:e52344.
- [5] Hobensack M, Song J, Scharp D, Bowles KH, Topaz M. Machine learning applied to electronic health record data in home healthcare: A scoping review. *Int J Med Inform*. 2023 Feb;170:104978. doi: 10.1016/j.ijmedinf.2022.104978. Epub 2022 Dec 30.
- [6] Cilli E, Ranieri J, Guerra F, Ferri C, Di Giacomo D. Naturalizing digital and quality of life in chronic diseases: Systematic review to research perspective into technological advancing and personalized medicine. *Digit Health*. 2022 Dec 18;8:20552076221144857. doi: 10.1177/20552076221144857
- [7] Wu CT, Wang SM, Su YE, Hsieh TT, Chen PC, Cheng YC, Tseng TW, Chang WS, Su CS, Kuo LC, Chien JY, Lai F. A Precision Health Service for Chronic Diseases: Development and Cohort Study Using Wearable Device, Machine Learning, and Deep Learning. *IEEE J Transl Eng Health Med*. 2022 Sep 19;10:2700414. doi: 10.1109/JTEHM.2022.3207825.
- [8] Rashid, J., Batool, S., Kim, J., Nisar, M. W., Hussain, A., Juneja, S., & Kushwaha, R. (2022). An Augmented Artificial intelligence approach for chronic diseases prediction. *Frontiers in Public Health*, 10, 860396. <https://doi.org/10.3389/fpubh.2022.860396>
- [9] Battineni, G.; Sagaro, G.G.; Chinatalapudi, N.; Amenta, F. Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosis. *J. Pers. Med*. 2020, 10, 21. <https://doi.org/10.3390/jpm10020021>
- [10] Olivera, A.R.; Roesler, V.; Iochpe, C.; Schmidt, M.I.; Vigo, Á.; Barreto, S.M.; Duncan, B.B. Comparação de algoritmos de aprendizagem de máquina para construir um modelo preditivo para detecção de diabetes não diagnosticada—ELSA-Brasil: Estudo de acurácia. *Sao Paulo Med. J*. 2017, 135, 234–246.
- [11] Topalovic, M.; Laval, S.; Aerts, J.M.; Troosters, T.; Decramer, M.; Janssens, W. Belgian Pulmonary Function Study Investigators. Automated Interpretation of Pulmonary Function Tests in Adults with Respiratory Complaints. *Respiration* 2017, 93, 170–178
- [12] Dinga, R.; Marquand, A.F.; Veltman, D.J.; Beekman, A.T.F.; Schoevers, R.A.; Van Hemert, A.M.; Penninx, B.W.J.H.; Schmaal, L. Predicting the naturalistic course of depression from a wide range of clinical, psychological, and biological data: A machine learning approach. *Transl. Psychiatry* 2018, 8, 241
- [13] Davis, F.; Gostine, M.; Roberts, B.; Risko, R.; Cappelleri, J.C.; Sadosky, A. Characterizing classes of fibromyalgia within the continuum of central sensitization syndrome. *J. Pain Res*. 2018, 11, 2551–2560
- [14] Gawlitza, J.; Sturm, T.; Spohrer, K.; Henzler, T.; Akin, I.; Schoenberg, S.; Borggreffe, M.; Haubenreisser, H.; Trinkmann, F. Predicting Pulmonary Function Testing from Quantified Computed Tomography Using Machine Learning Algorithms in Patients with COPD. *Diagnostics* 2019, 9, 33.
- [15] Hueso, M.; Vellido, A.; Montero, N.; Barbieri, C.; Ramos, R.; Angoso, M.; Cruzado, J.M.; Jönsson, A. Artificial Intelligence for the Artificial Kidney: Pointers to the Future of a Personalized Hemodialysis Therapy. *Kidney Dis*. 2018, 4, 1–9.
- [16] Zhang, H.; Berg, A.; Maire, M.; Malik, J. SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition—Volume 2 (CVPR 06)*, New York, NY, USA, 17–22 June 2006; Institute of Electrical and Electronics Engineers (IEEE): Piscataway Township, NJ, USA
- [17] Buscema, P.M.; Massini, G.; Breda, M.; Lodwick, W.A.; Newman, F.; Asadi-Zeydabadi, M. *Artificial Neural Networks. In Studies in Systems, Decision and Control*; Springer: Berlin/Heidelberg, Germany, 2018.
- [18] Kuru, K.; Niranjana, M.; Tunca, Y.; Osvank, E.; Azim, T. Biomedical visual data analysis to build an intelligent diagnostic decision support system in medical genetics. *Artif. Intell. Med*. 2014, 62, 105–118.
- [19] Wells, G.; Shea, B.; O'Connell, D.; Peterson, J. *The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses*; Ottawa Hospital Research Institute: Ottawa, ON, Canada, 2000
- [20] Polat, H.; Mehr, H.D.; Cetin, A. Diagnosis of Chronic Kidney Disease Based on Support Vector Machine by Feature Selection Methods. *J. Med. Syst*. 2017, 41, 55.
- [21] Bini, S. Artificial Intelligence, Machine Learning, Deep Learning, and Cognitive Computing: What Do These Terms Mean and How Will They Impact Health Care? *J. Arthroplast*. 2018, 33, 2358–2361.

CONSOLE LESS HAND GESTURE CONTROL DRONE

Amrutha R¹, Darshan Gadi², Gangadhar S J³, Prajwal B⁴ and Roshan Shetty⁵

Department of Electronics and Communication Engineering, Alvas Institute of Engineering & Technology, Moodbidri, Mangalore, Karnataka, India
Emails: : amrutharamesh029@gmail.com¹, darshangadi07@gmail.com², gangadharjo19@gmail.com³, prajwalbilagi244@gmail.com⁴, roshans@aiet.org.in⁵

Abstract.

Gesture-controlled drones employ glove sensors that use accelerometers and gyroscopes to record motion in place of traditional controllers. This kinetic data is sent to the drone, allowing it to react to the user's movements naturally. In order to prevent collisions, the drone's obstacle sensors continuously detect nearby objects and automatically alter or stop its movement. Second, the gesture information is transmitted to the microcontroller of the drone, which processes and acts upon the command. E., off, lift, turn. Lighting variability further limits visual control efficiency, favoring sensor-based options. Wireless with drone wire with sensor quick and easy. And that's a big plus because if you don't want to leave your house of work you can just use the drones and see what you're doing or you can send it out to rescue someone or you can use it to walk through a dangerous place. This project is a new interaction paradigm where users can operate flying drones with a wearable sensor, providing an innovative and engaging experience.

Keywords: Security Monitoring, Rescue Operations, Industrial Inspection, Scientific Research, Media and Entertainment, Healthcare, Agriculture, Autonomous Vehicles, Streaming Media, Drone Management.

1. INTRODUCTION

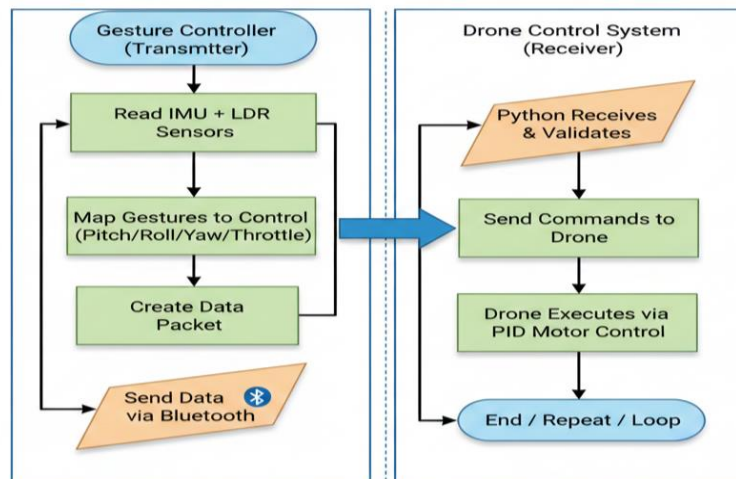
Uavs or drones can do a lot of things like farming disaster delivery safety and a lot more because they can do it by themselves they are super easy to use and they're really cheap that's why more and more people are trying to use uavs and drones to develop a new form of human digital interaction. And with some have speculated will eventually replace the more user-friendly physical remote.

Gesture input device is a type of input device that can detect hand or body movement made by the user and transform those gestures into commands for the computer for example through visual recognition by a camera or through sensors embedded in gloves for example by deep learning models such as convolutional neural network CNN or residual neural network RNN architectures which have achieved accuracy rates exceeding

90 in simulated environments demonstrating promising potential for future use outside of laboratory environments. These systems could be used for tracking moving objects, avoiding obstacles on uneven surfaces, or performing other tasks. Gesture-based input systems are designed to be userfriendly and intuitive, offering an alternative control method for users unfamiliar with the technology and replacing physical contact with objects or surfaces.

Furthermore, ROS integrated with Raspberry Pi and NodeMCU offers a decentralized and simplified robotic architecture. but this does not mean that they have to make some new friends, hang out and have some fun or play shoot on the game and i just want to say that i could put vr in my project and i will go on to describe below what i have done, what i am doing and what i going to put in my project.

2. METHODOLOGY OF WORK FLOW



Gesture are detected by the system for uav control and input data are processed through the imu and optical sensor in order to track the hand movement it then generates control data including pitch , roll , yaw and throttle and wirelessly transmits this information to the drone the pi d control algorithm keep the motor in a stable condition, it is also very easy to use , it enables the uavs to control in a simple way and it does not require the user to have ever flown a drone before.

The proposed system consists of two major modules:

- (1) The Gesture Controller (Transmitter). and
- (2) The Drone Control System (Receiver).

A. Gesture Controller (Transmitter Side)

1. **Sensor Data Acquisition (IMU + LDR Sensors) :** The gesture controller continuously collects motion and orientation data using an Inertial Measurement Unit (IMU), which provides accelerometer and gyroscope readings. Light dependent resistors are used to detect changes in hand position and movement

during different gestures. These sensors serve as the main input source for recognizing and interpreting human hand gestures.

2. **Gesture Processing and Mapping :** The sensor data is first cleaned to ensure consistent readings and remove unwanted noise. Then, different hand positions and motions are linked to drone controls like pitch, roll, yaw, and throttle. This process helps translate natural human gestures into measurable, intelligible commands for the drone.
3. **Data Packet Formation :** The gesture values are packed into a small data frame to guarantee that the information is sent quickly. This packet includes gesture details, basic error-checking bits, and identification information. These improvements help guarantee that the drone gets reliable and consistent data.
4. **Wireless Data Transmission via Bluetooth :** The transmitter transmits the data packet to the drone via a Bluetooth connection. Bluetooth was chosen due to its low power consumption and consistent communication range. It is also suitable for remote drone control since it enables dependable data transfer.

B. Drone Control System (Receiver Side)

5. **Data Reception and Validation Using Python :** The receiver module is powered by a Python-based interface that continuously searches Bluetooth for information. Each received packet is checked for correctness, signal quality, and format before being accepted. This ensures that the drone only gets valid commands, preventing errors or dangerous motions.
6. **Forwarding Commands to the Drone Controller :** Once the data has been verified, the control values are transmitted to the drone's flight controller. The interface then converts these gesture-based values into motor commands that the drone can understand. This makes it possible for the drone to respond accurately and fluidly to the user's hand movements.
7. **Execution Through PID-Based Motor Control :** The drone uses a Proportional–Integral–Derivative (PID) algorithm to interpret the commands and adjust motor speeds accordingly.
8. **Continuous Real-Time Control Loop :** Data is continuously transmitted, verified, and processed by the system as a whole. The drone and the user can communicate in a responsive and seamless manner thanks to this real-time flow. The drone can thus follow hand gestures with ease and naturalness.

The proposed gesture-controlled drone system records hand movements and translates them into flight commands using IMU and LDR sensors. These motions are used to calculate pitch, roll, yaw, and throttle values. The system encapsulates this information into data packets, which are then wirelessly transmitted to the drone via Bluetooth. The flight controller receives valid commands from a Python-based receiver after the data has been verified. The controller uses PID control to ensure responsive and smooth drone movement.

3. REFERENCES

- [1] R. A. Suarez Fernández et al., “Natural user interfaces for human-drone multimodal interaction,” Proc. 2016 Int. Conf. Unmanned Aircraft Systems (ICUAS), Arlington, VA, USA, 2016, pp. 1013–1022.
doi: 10.1109/ICUAS.2016.7502665
- [2] Y. Yu, X. Wang, Z. Zhong, and Y. Zhang, “ROS-based UAV control using hand gesture recognition,” Proc. 2017 29th Chinese Control and Decision Conference (CCDC), Chongqing, China, 2017, pp. 6795–6799.
doi: 10.1109/CCDC.2017.7978402
- [3] K. Haratiannejadi, N. E. Fard, and R. R. Selmic, “Smart Glove and Hand Gesture-Based Control Interface for Multi-rotor Aerial Vehicles,” 2019 IEEE Int. Conf. Systems, Man and Cybernetics (SMC), Bari, Italy, 2019, pp. 1956–1962.
doi: 10.1109/SMC.2019.8914464
- [4] S.-Y. Shin, Y.-W. Kang, and Y.-G. Kim, “Hand gesture-based wearable human-drone interface for intuitive movement control,” 2019 IEEE Int. Conf. Consumer Electronics (ICCE), Las Vegas, NV, USA, 2019, pp. 1–6.
doi: 10.1109/ICCE.2019.8662106
- [5] T. Begum, I. Haque, and V. Keselj, “Deep Learning Models for Gesture-Controlled Drone Operation,” 2020 16th Int. Conf. Network and Service Management (CNSM), Izmir, Turkey, 2020, pp. 1–7.
doi: 10.23919/CNSM50824.2020.9269056
- [6] R. Wang and C. Sun, “A Gesture Recognition and Drone Control System Based on Residual Neural Network,” 2023 IEEE Int. Conf. Mechatronics and Automation (ICMA), Harbin, China, 2023, pp. 25–30.
doi: 10.1109/ICMA57826.2023.10215930

Biographies



Name : Amrutha R
USN : 4AL22EC004
Email : amrutharamesh029@gmail.com
Mobile : 8073244559

Areas of Interest: She is currently pursuing a Bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects a strong interest in core and emerging electronics domains such as Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers , Wireless Communication and Semiconductor Devices.

Second International Conference on Smart Computing and Communication for Sustainable Convergence (ISCCSC 2025)



Name : Darshan Gadi
USN : 4AL23EC400
Email : darshangadi07@gmail.com
Mobile : 9028247949

Areas of Interest: He is currently pursuing a bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), which is affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Microcontrollers, semiconductor devices, PCB design, VLSI design, and digital signal processing are just a few of the traditional and cutting-edge electronics fields in which his academic background indicates a strong interest.



Name : Gangadhar S J
USN : 4AL23EC401
Email : gangadharjo19@gmail.com
Mobile : 8660341587

Areas of Interest: He is currently pursuing a bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Microcontrollers, semiconductor devices, PCB design, embedded systems, are just a few of the traditional and cutting-edge electronics fields that his academic background indicates a strong interest in.



Name : Prajwal B
USN : 4AL23EC404
Email : prajwalbilagi244@gmail.com
Mobile : 8618266034

Areas of Interest: He is currently pursuing a bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Microcontrollers, embedded systems, VLSI design, and digital signal processing are just a few of the traditional and cutting-edge electronics fields that his academic background indicates a strong interest in.



Name : Roshan Shetty
Designation : Associate Professor (Project Guide)
Email ID : roshans@aiet.org.in
Mobile Number : 9538176004

Areas of Interest : As an associate professor at Alva's Institute of Engineering and Technology (AIET), which is connected to Visvesvaraya Technological University (VTU) in Karnataka, in the Department of Electronics and Communication Engineering.

Crime Pattern Analysis and Prediction Using Machine Learning

Pradeep Nazerath¹, Lathesh Kumar S R², Sanketh Patil³, Shivamani M Nayak⁴ and
Tejashwini Shailendra Murdeshwar⁵

Department of Artificial Intelligence and Machine Learning, Alvas Institute of Engineering & Technology, Moodbidri, Mangalore, Karnataka, India

Emails: pradeepn@aiet.org.in, latheshkumar06@gmail.com, sanketpatilsp360@gmail.com,
shivamaninayak5757@gmail.com, murdeshwartejashwini@gmail.com

Abstract.

Urban crime rate changes that are influenced by tiny spatiotemporal factors are typically outside the realm of a usual police investigation. This paper reflects the authors' hierarchical machine learning workflow to understand crime trends through the "Crime in India" open dataset from 2001 to 2020. Our approach very detailed stages a preprocessing, moves the Synthetic Minority Over-sampling Technique (SMOTE) to solve the class imbalance problem, and has multilayer feature extraction to capture the spatial and temporal aspects so that the classifier's prediction quality can be brought up to standard again. The authors tried different algorithms to have a quantitative comparison of their performances and measured accuracy, precision, recall, and F1-score. The algorithms were K-Nearest Neighbors, Support Vector Machine, Decision Tree, Random Forest, and Naive Bayes. Random Forest was the most accurate classifier as it achieved the predictive accuracy of 89 % and being very stable, it varied almost evenly across all the evaluation metrics.

Keywords. Crime Prediction, Machine Learning, Spatiotemporal Analysis, Data Imbalance, Ensemble Learning, Smart Policing, ST-CrimeNet

1. INTRODUCTION

The fast-paced urbanization, higher population density, and disparities in income and living standards have a direct impact on the occurrence and nature of crimes in today's world. Over the years, the increasing size of cities has made it even harder for the police to come up with and implement the right strategies for detecting and fighting crime, given that they are still using the old-fashioned techniques for analyzing criminal activities. Furthermore, traditional methods mostly depend on laborious investigations and the preparation of historical reports which are not always able to reflect the intricate geographical and temporal interrelationships that large crime datasets contain. Recently, the accessibility of extensive digital crime records and the progress in machine learning have together opened up new vistas for crime analysis based on data. The methods of machine learning make it possible to automate the uncovering of hidden patterns, predict trends, and classify crimes with an even higher level of accuracy. These models, by making use of past crime data, help in examining the variations of crime across the cities, different time periods, and various environmental conditions; thus, they are working in the support of proactive police strategies. Nevertheless, there are drawbacks associated with the deployment of machine learning for crime forecasting. The crime datasets are usually very much unbalanced, wherein, the occurrences of some crime types are much more than the others. This skewness very often results in the formation of biased models that yield poor results for rare but significant crimes.

2. LITERATURE REVIEW

- [1] Chawla et al. put forward the Synthetic Minority Over-sampling Technique (SMOTE) as a solution to the problem of class imbalance in datasets. Their research revealed that the performance of classifiers is enhanced by the minority class oversampling especially in the highly imbalanced datasets like crime records which are notoriously difficult to classify.
- [2] Moher et al. put forward spatiotemporal point process models for crime analysis and pointed out the significance of place and time in forecasting the areas with most criminal activities. Their method was efficient for spotting hotspots, but it was not able to deal with the interactions between complex features flexibly.
- [3] Breiman developed the Random Forest algorithm and has since been applied in crime prediction mainly due to its strength and non-susceptibility to overfitting. Random Forest was reported in various studies to be able to improve the accuracy and stability of crime datasets.
- [4] Wang et al. used geographic clustering methods to introduce spatial characteristics into the crime prediction models. Their findings allowed for the most accurate selection of regions with a high incidence of crime, but they still did not dive deep into the issue of temporal dependencies.
- [5] Huang et al. made use of spatiotemporal modeling techniques to study the variation of crimes and highlighted the significance of patterns that depend on time such as daily and yearly variations. However, their method required a lot of computer power.
- [6] Liu et al. experimented with deep learning models for predicting crimes and proved better performance in pattern recognition. But even with higher accuracy, models were not very interpretable and required more computations, resulting in increased computational cost.
- [7] Rudin underlined the necessity of interpretable models in situations involving decision-making with serious repercussions like police work. The research pointed out that besides being transparent, predictive policing systems should also be understandable user-friendly for their acceptance.

3. METHODOLOGY

The four stages are structured analytically in a highly sequential and integrated manner: it consists of data preparation, feature engineering, model implementation, and evaluation. The stage

3.1 Dataset Collection and Pre-Processing

The research utilizes the open-source dataset "Crime in India", which is publicly accessible on Kaggle and presents a comprehensive account of crimes across the country from the year 2001 to 2020. The data file consists of approximately 1.2 million rows that cover the different crime categories, time and place identifiers, and some demographic variables. The initial preprocessing and scaling the values of the numerical attributes to a range of 0 to 1 using the Min-Max method. Outliers were removed using the Interquartile Range (IQR) method which reduced their influence on the learning phase. A drastic class imbalance where common crimes were several times more than rare ones was the main drawback of the dataset. The Synthetic Minority Oversampling Technique (SMOTE) was used as a solution to this drawback by generating synthetic minority samples and thus promoting classifier generalization. The data set that was processed and distributed equally was considered the starting point for modeling. The focus of feature engineering in this research was to obtain spatial and temporal attributes that would be richer and more informative.

Records that have location coordinates were mapped to localities by K-Means clustering, demonstrating the geographical regions of crime hotspots according to the volume of crime in a given location. The newly created variables provided not just additional information but also enhanced the models' capability to reveal concealed spatiotemporal patterns.

3.2.1 Machine Learning Algorithms:

To identify the top algorithm for multiclass crime classification, five different supervised learning algorithms were tested, and the K-Nearest Neighbors algorithm was applied with the Euclidean distance metric to identify local similarity patterns.

$$d(x, x_i) = \sqrt{\sum_{j=1}^n (x_j - x_{ij})^2}$$

KNN, even though it's an effective approach for neighborhood-based reasoning, suffers from the problems of noise and high dimensionality. In accordance with the techniques mentioned in the preceding comparative literature, Support Vector Machines with an RBF kernel were employed to form non-linear class boundaries in the crime data.

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \dots, h_T(x)\}$$

The model was able to get good results for categorical variables.

3.2.2 Experimental Setup and Evaluation

All of the experiments were carried out in Python 3.11 using Jupyter Notebook along with the usual standard data science libraries like Scikit-learn, pandas, NumPy, matplotlib, seaborn, plotly, and Folium. The research data was split into two parts, 70% for training and 30% for testing, and 5-fold cross-validation was applied to ensure that the performance estimation was unbiased. Hyperparameter tuning was performed using GridSearchCV.

4. CONCLUSION

The mentioned paper illustrates one of the machine learning techniques which through analyzing and predicting crime patterns by handling a huge volume of historical crime data. The research paper has presented the comparative evaluation as a way to demonstrate the application of different machine learning methods like K-Nearest Neighbor, Support Vector Machine, Decision Tree, Random Forest, and Naive Bayes not just for crime classification but also for getting considerable spatiotemporal patterns discovered. The most important aspect of the methodological framework was data preprocessing, the class imbalance problem which was tackled through SMOTE, and the deliberate feature engineering in order to obtain the most relevant temporal and spatial features.

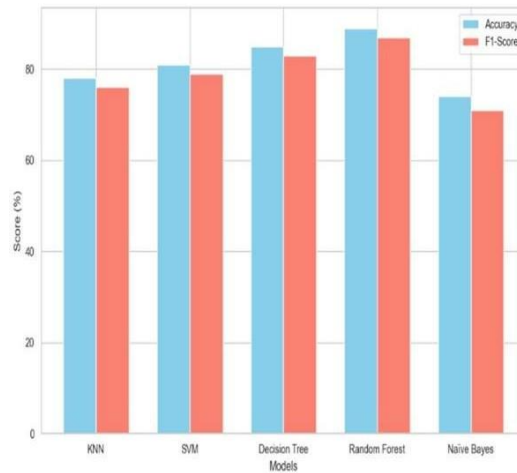


Fig. 1: Model Comparison: Accuracy vs F1-Score

Research on model performance is only a minor segment of the article, which constitutes the larger argument for data-driven crime analytics over traditional evaluation methods. The traditional approaches rely more on proofs done part by part and story-telling insights whereas machine learning algorithms give extensive and data-centric evaluations. This kind of evaluation can be used to spot the specific areas resources are allocated to and those that need to be improved, and at the same time, to make sure that the crime prevention policies created have the

right focus. The real-time data integration, the use of advanced modeling techniques, and the improvement of dataset quality won't just improve these systems, but they will also support the development of more responsive and community-oriented public safety initiatives which will be less reliant on technology.

5. REFERENCE

- [1] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-Sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
- [2] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [3] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita, "Self-exciting point process modeling of crime," *Journal of the American Statistical Association*, vol. 106, no. 493, pp. 100–108, 2011.
- [4] F. Pedregosa, G. Varoquaux, A. Gramfort, et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [5] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.
- [6] L. Wang, X. Zhang, and Y. Liu, "Predicting crime using spatial features," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2018.
- [7] Q. Huang, Y. Guo, and Z. Li, "Spatio-temporal crime prediction using GIS and deep learning," *ISPRS International Journal of Geo-Information*, vol. 10, no. 9, p. 602, 2021.
- [8] C. Rudin, "Why we should use interpretable models instead of black box models for high-stakes decisions," *Nature Machine Intelligence*, vol. 1, pp. 206–215, 2019.

BIOGRAPHY:



Name: DR. PRADEEP NAZARETH

Designation: Associate Professor (Project Guide)

Email ID: pradeepn@aiet.org.in

Mobile Number: +91 9164525591

Dr. Pradeep Nazareth is an Associate Professor in the Department of Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with VTU, Belagavi. He holds a B.E. from KVG College of Engineering, an MTech from SJCE Mysore, and a Ph.D. from NITK Surathkal.



Name: LATHESH KUMAR S R

USN: 4AL23AI400

Email ID: latheshkumar06@gmail.com

Mobile Number: +91 7760814609

Areas of Interest: Lathesh Kumar S R is currently pursuing a Bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya

Technological University (VTU) in Karnataka. His academic interests span across Generative AI, Web Technologies, Natural Language Processing (NLP), Large Language Models (LLMs), Cloud Computing, Deep Learning, Machine Learning, and Robotics.



Name: SANKET PATIL

USN: 4AL22AI043

Email ID: sanketpatilsp360@gmail.com

Mobile Number: +91 8660966350

Areas of Interest: Sanket Patil is currently pursuing a Bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. His scholarly focus encompasses key areas such as Data Science, Machine Learning, and Artificial Intelligence. I am particularly interested in Generative AI and Cloud Computing Technology.



Name: SHIVAMANI M NAYAK

USN: 4AL22AI051

Email ID: shivamaninayak5757@gmail.com

Mobile Number: +91 98050666408

Areas of Interest: Shivamani M Nayak is currently pursuing a Bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. His scholarly focus encompasses key areas such as Data Science, Machine Learning, and Artificial Intelligence. I am particularly interested in Generative AI and Cloud Computing, and I aspire to explore how these technologies can be integrated to build intelligent, scalable solutions.



Name: TEJASHWINI SHAILENDHRA MURDESHWAR

USN: 4AL22AI060

Email ID: murdeshwartejashwini@gmail.com

Mobile Number: +91 9353999908

Areas of Interest: Tejashwini Shailendhra Murdeshwar is currently pursuing bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic interests include Generative AI, AI Ethics, Deep Learning, Machine Learning and Data Science, Project Management.

IoT BASED UNDERGROUND CABLE FAULT DETECTION SYSTEM

**Akash H M¹, Chaitra Basavaraj Gandarotii², Jyothi³, Jaahnnawwie A⁴,
Faisal K⁵**

1,2,3,4,5: Dept. of Electronics and Communication Engineering, Alva's Institute of Engineering and Technology
Moodbidri, Karnataka, India

Abstract

Most of the time, burying high-voltage cables underground is their main advantage in the market. It allows the companies to maintain pollution resistance and other security measures. In addition, a weather forecast is also enhanced when using such cable lines. Thus, they are also the ones mostly chosen for the external safety and weather stability. The irony of the whole situation is that in spite of their awesome benefits, the cables can still be affected by the wrong method of installation, poor jointing, third-party excavation activities, abrasion, environmental stress, and animal interference. Often the process of finding these faults and repairing them is very difficult, it consumes a lot of time and money. The system is the fastest fault detection device, as certified by the experiment, and this leads to a reduction of the machine stop time and the maintenance costs, thus, the reliability of the modern power distribution networks is maintained. This article presents a fault detection and diagnosis system for underground cables based on the Internet of Things (IoT) that intends to solve the aforementioned problem. Devices with the proposed system, such as voltage divider circuits, enable the detection of variations in the fault voltage and are connected to the cable. ESP32 microcontroller handles all of the signal processing work, fault location in kilometers, and communication of the result to the user. The data can be accessed on the liquid crystal display (LCD) and are simultaneously sent to the Blynk cloud database by the wireless module for remote monitoring and archiving.

1. INTRODUCTION

First of all [1], power electricity that is dependable and efficient, is just about the essentials of life in this era, and underground power cables are those that have the biggest share in the fact that the supply is stable and without any interruptions. The only method to reduce the time of outages to the minimum and retain a distribution network system that is resilient is by a very prompt fault detection as well as the exact location of the fault when it is found. An important factor in the success of the IoT technology is the ease of its identification as the precise solution for the problem in question, and the recognition by the industry, that it has done so largely over the last few years. IoT-based systems, which implement smart sensors, data analytics, and wireless communication are revolutionizing the traditional way of monitoring significantly. They, for instance, through the real-time tracking and automated fault detection, not only have a faster response time but also keep a greater overall efficiency to a large extent as the downtime has been lowered considerably. Most of the time [2], power cables meant for underground distribution of electricity are just buried in the soil or installed in protected ducts. Such a product guarantees the long life of the cable and the least possible number of faults. In spite of this, these faults in underground systems, locating and fixing them is a huge problem because the conductors cannot be seen. The fault detection process can in many instances be similar to the needle in a haystack finding process. Over time, several methods, and also expensive diagnostic tools have been developed, which greatly facilitate fault detection. However, the claim that no single method or even a combination of methods could be considered as the "best" ones is still true. The different fault types determine the use of different methods to achieve safety, accuracy, and effectiveness in fault localization as well as the non-damage of the cable. The next section is about the fault genres that are happening in underground cables.

2. PROBLEM STATEMENT

Underground [3] cables have really become an indispensable part of the power distribution system of the modern world and the reasons for such a statement are quite clear. Not only do cables enhance safety and reliability, but they also protect the network from environmental disturbances, such as pollution, heavy rains, and storm. However, they are still creating lots of problems in the areas of fault detection and maintenance. In contrast, underground cables are completely covered, thus a person cannot simply do a check or a supervision of their condition directly. When the malfunction happens by short-circuiting, insulation leakage, or mechanical injury to the cable, locating the fault becomes a very difficult and long process. The traditional means of power network fault detection are largely reliant on human intervention during inspection, extensive sectional testing, and the use of large and complicated measuring instruments. Not only these methods decelerate the process of power restoration, but they also escalate the operational costs. The delay in outage period, system reliability decreasing, utilities losing money, and consumers experiencing inconveniences are some of the consequences that follow. Additionally, faults may result from poor installation practices, joint failures, digging done by the third parties, environmental stress, or even rat attacks that have complicated the fault identification process so much. Hence, it is totally mandatory that a decision be taken towards the installation and use of a reliable, automated, and accurate fault detection system for underground cable faults that can locate the faults in a very short time. The adoption of such a system will undoubtedly reduce the time for which the network has to be de-energized thus it will lessen maintenance activities, elevate safety, and improve the reliability of power distribution networks.

3. LITERATURE SURVEY

L.Kurinjimalar et al [1] The authors of the paper mention that despite the advantages that buried cables have over aerial lines, one of the significant problems is the precise fault location. As a solution to this problem, they present a fault indicator system based on the Internet of Things, which utilizes Ohm's Law to locate the fault in the distance from the base station in kilometers. Thus the malfunction in underground cables, which are extensively utilized in urban areas, can be expediently detected and repaired but are difficult to locate when they occur.

Abrar et al [2] . describe how underground cables are buried and are subjected to numerous faults due to environmental damage and wear which make it difficult to carry on repairs since the cables are usually dug up along their entire length. They want to be able to accurately locate the faults with their project which is the node controller system. The system is able to locate the exact faults positions by the method of modeling the cable as resistors and simulating faults. After the fault data is wirelessly transmitted with the aid of IoT technology and Wi-Fi module, the underground cable faults can be easily and remotely monitored.

Jayachandran et al [3] In their work, they clarified the IoT Based Smart Underground Cable Fault Detector . The authors have developed a diagnosis system using the Internet of Things (IoT) technology, which is capable of detecting the underground cable faults and also to localize the faults. The system through a quicker fault detection, therefore, intends to cut down the required time, money and the amount of work that has to be done in order to the area fixed. The IoT-based method involved the use of a set of voltage dividers that were installed on the cable to find the fault. In the case of a fault due to a short circuit between two cables, a voltage is produced, which is recorded by the ESP32 microcontroller. This microcontroller sends the location of the fault in kilometers to the user. The information collected is shown on a liquid crystal display (LCD) and is sent online through a Wi-Fi module to the Blynk cloud database for storage and further processing.

2nd International Conference on Smart Computing and Communication for Sustainable Convergence (ISCCSC 2025)

B.Karuna Sree et al[4]. have come up with a fault detection system which is monitoring the current that the underground cables pass through, to be able to detect faults. Every cable is given a maximum current value, thus in the case of a short circuit the current will be skyrocketing and if there is a break in the circuit the current will be zero. Current transformers get the current; it is then changed into a voltage signal that the Arduino reads. The Arduino is the one that receives and processes the data thus making the fault identification visible on the LCD and LEDs. Besides this, the system employs relays which are used for changing the power supply from the faulty phase to the one that is healthy, thus a continuous power supply can be achieved by the users. In addition, there is a manual fault-clearing switch that enables the operator to reset the system after the repair work is done.

Shingrut et al[5] have explained in their paper about a project, which is designed to identify the distance of the underground fault cable in kilometres from the base station and the same is displayed over the internet. An underground cable system is a standard practice in most major metro cities areas. When a fault occurs, at that particular moment the repairing process of that cable is hard because the fault's exact location is unknown in the cable. It is a new technology that can accurately locate the fault and at the same time a GSM module is used to send graphical data to the website and also a display is given on the LCD screen.

Harikrishna et al[6].have explained in their paper about a project that underground cables are frequently damaged due to various environmental factors and the attack by rodents. The locating of the fault precisely in a cable is usually a very difficult process. As a result, large parts of the cable have to be dug up, which not only takes a lot of time but also a considerable amount of money is wasted. Besides that, there are different methods like sectionalizing, thumping, and the Murray loop. However, these methods are generally slow and expensive. The authors decided to develop an IoT-based underground cable fault detection system which can not only locate the fault remotely but also be able to save the time and the resources by the way of allowing the repair to rest only on the exact faulty area.

Rakesh et al[7] introduced a method based on Ohm's Law for the identification of faults in underground cables. Local current changes when a short circuit is the source of the fault, at the feeder end, consequently. Reading this via its Analog-to-Digital Converter (ADC), an Arduino captures these current variations by voltage measurement across a resistor. The system, using this information, figures the fault point and shows on an LCD how many kilometers from the base station the fault is. The device represents the cable with resistors for different lengths, and the fault switches are installed at the certain distance to simulate the faults for the fault distance to be detected accurately.

Esha Lohar et al[8]. developed the technology for identifying faults in underground cables which are generally laid in cities to avoid weather-related damages. They employ Ohm's Law in their method: the current varies if a short circuit location is found when a low DC voltage is applied. An Arduino microcontroller through an ADC reads the current changes, resistors representing the lengths of the cables in kilometers. The faults have been caused by the use of switches, and a relay exciter IC manages the test. The 16x2 LCD display can be used to find out the exact location of the fault and help messages with the description of the fault type and distance are sent via GSM to the base station. In case the fault occurs, a buzzer will immediately sound to alert the field workers for further actions.

Prabha et al[9]. designed a system based on IoT to track the exact location of the fault in underground cables from a base station with the help of Arduino. Essentially, the system uses Ohm's Law in which the current varies according to the fault's position in the cable. As most urban areas have underground cables, it is extremely necessary to locate the faults in a short time in order to carry out the repair work on time. Arduino, which is the system's central processing unit, receives the data from the sensors and a fault is detected by the use of the resistor network when the two lines are cut. The fault details are then posted on the net for the police to have a quick and easy access.

Divya Sugathan et al[10] explains the usage of resistors to simulate the length of the cable in kilometers as well as switches to represent the faults at known locations in this document. Generally, underground cables that are extensively installed in areas such as hospitals, roads, and educational institutions are more reliable than overhead wires since they are less exposed to harsh weather and pollution. But, finding the faults in the underground cables that cannot be seen is a big problem. The most frequent faults are short circuits, open circuits,

and earth faults, while open circuits are resulted from the breakdown of the conductor. Their project is to be able to pinpoint the fault locations with accuracy to make the fault fixing process easier.

4. CONCLUSION

IoT-based systems for identifying underground cable faults offer a novel, efficient, and reliable method for overseeing and maintaining electrical networks. These systems utilize sensors, microcontrollers, and IoT connectivity to notify teams in real-time about faults (such as short circuits or line-to-ground faults), pinpoint the fault along the cable's length, and relay that information to a maintenance crew remotely. This will minimize downtime, decrease labor costs, and improve system dependability. The setup of IoT systems enables ongoing monitoring, data logging, and analysis that improve predictive maintenance and avert catastrophic failures. Despite challenges such as sensor accuracy, environmental factors, and communication reliability, advancements in IoT technology will enhance characteristics like precision and scalability. In summary, IoT solutions for detecting underground cables provide a significant advancement for a SMART or automated electrical distribution system. Additionally, monitoring based on IoT enables predictive maintenance by examining past fault data, identifying patterns, and forecasting a cable failure prior to its occurrence. An anticipatory work process minimizes unforeseen disruptions, boosts service reliability, and optimizes resource distribution for utility companies. Monitoring through IoT enhances safety by reducing the necessity for manual inspections in hazardous locations, such as underground facilities or areas with high voltage. Upcoming advancements could involve the adoption of cutting-edge technologies like a machine-learning algorithm for identifying faults, a wireless sensor network for remote regions, and cloud data analytics to enhance decision-making processes. Additional options for enhancing system sustainability involve energy-efficient and self-sustaining sensors. With these advancements, IoT driven underground cable fault identification systems may turn into a vital component of smart grid infrastructures to enhance the resilience, intelligence, and automation of power distribution network.

6. REFERENCE

- [1]. L. Kurinjimalar, B. Kamalakannan, P. Krishna Kumar, J. T. Mahendra Varma Saliyan, "2024 International Conference on Communication, Computing and Internet of Things (IC3IoT)," IEEE, 2024. DOI: 10.1109/IC3IoT60841.2024.10550284.
- [2]. M.Moyeed Abrar, I. Tabassum, A.Mehroz, M. Bushra, H.Tahseen, "International Journal of Creative Research Thoughts (IJCRT)," vol. 12, no. 6, Jun. 2024. ISSN: 2320-2882.
- [3]. M. Jayachandran, K. N. A. Rani, A. A. Wahab, "IEEE 1st International Conference on Communication Engineering and Emerging Technologies (ICoCET)," IEEE, 2024.
- [4]. B. Karuna Sree, V. Shivani, P. V. Shyam Kumar, K. Varalakshmi, G. Srija, "International Journal of Research and Analytical Reviews (IJRAR)," vol. 11, no. 2, Apr. 2024
- [5]. R. Shingrut, S. Shelar, D. Mokal, S. Mhatre, S. Sane, "Underground Cable Fault Detection," International Journal of Engineering Research & Technology (IJERT), 2024. DOI: 10.17577/IJERTV9IS02014.
- [6]. V. Hari Krishna, G. Neelima, P. Poojitha, M. Madhu, D. V. N. Surya, C. Satwika, "International Journal of Novel Research and Development (IJNRD)," vol. 9, no. 5, May 2024. ISSN: 2456-4184.
- [7]. Y. Rakesh Kumar, M. Vani, N. Usha Kiran, S. Rukhsana, "Journal of Emerging Technologies and Innovative Research (JETIR)," 2023.
- [8]. Y. Rakesh Kumar, M. Vani, N. Usha Kiran, S. Rukhsana, "Journal of Emerging Technologies and Innovative Research (JETIR)," 2023.
- [9]. P. R. Prabha, K. S. Shree, K. R. Kumar, L. S. C. K. S., G. Manickam, "Journal of Emerging Technologies and Innovative Research (JETIR)," vol. 10, no. 5, May 2023.
- [10]. D. Sugathan, P. T. S. Prerana, N. S., N. M., N. V., "7th International Conference on Design Innovation for 3Cs Compute Communicate Control (ICDI3C)," IEEE, 2023. DOI: 10.1109/ICDI3C61568.2023.00079

Project Associates Information



Akash H M USN 4AL22EC003 worked on hardware design and sensor interfacing, including ESP32–INA219 integration, power supply setup, and circuit testing. He ensured accurate voltage and current measurement and supported hardware calibration and debugging.

Contact: 8792087958 —

Email: 4al22ec003@aiet.org.in



Chaitra B G USN 4AL22EC019 was responsible for software development and fault detection logic. Her contributions include MicroPython programming, resistance and distance calculation, phase fault identification (R, Y, B), and LCD, LED, and buzzer interfacing.

Contact: 9591864170 —

Email: 4al22ec019@aiet.org.in



Jaahnavwie USN 4AL22EC028 handled IoT connectivity and web dashboard development. She configured Wi-Fi communication, ESP32 web server functionality, and enabled real-time fault data visualization and remote monitoring.

Contact: 6380062735

Email: 4al22ec028@aiet.org.in



Jyothi USN 4AL22EC029 contributed to simulation mode design, testing, and documentation. She developed the switch-based fault simulation model, validated system performance, and assisted in report preparation and result analysis.

Contact: 8861196447 —

Email: 4al22ec029@aiet.org.in



Mr. Faisal K, Assistant Professor, Department of Electronics and Communication Engineering, guided the project with continuous support and technical expertise. His mentorship in embedded systems and IoT applications played a key role in the successful completion of the project.

Contact: 9943027422

Email: faisalk@aiet.org.in

Downlink Channel Prediction in FDD Massive MIMO Using Deep Transfer Learning Algorithms

Bathala Tanusree¹, Ravi Kumar CV², Suhaib Ahmed^{3,*}

¹*Amrita Vishwa Vidyapeetam, Bengaluru, India, tanusreebathala18@gmail.com*

²*School of Electronics Engineering, VIT Vellore, Tamil Nadu, India, cvrkvit@gmail.com*

^{3,*}*Model Institute of Engineering and Technology, Jammu, India, sabatt@outlook.com*

Abstract.

Precise downlink Channel State Information (CSI) estimation in Frequency Division Duplexing (FDD) massive MIMO systems is still a fundamental challenge because of the absence of channel reciprocity and excessive feedback overhead. In this work, four leading strategies MMSE, Autoencoder-based SAE, Adversarial SAE, and Meta-Learning are analyzed and compared in terms of performance for effective downlink CSI prediction based on uplink observations. Simulation results confirm that although MMSE is a good baseline, deep learning models greatly surpass it in terms of accuracy and versatility, especially when there are changing conditions and the data is not much available. Among them, Meta-Learning performs the least Normalized Mean Squared Error (NMSE) and delay with different SNRs, frequency offsets, and antenna numbers with superior generalization and rapid convergence. Adversarial SAE demonstrates strong performance under low-SNR conditions, whereas Autoencoder SAE performs well with data-efficient training. These results emphasize the engineering potential of learning-based CSI predictors, particularly Meta-Learning, for scalable and robust 5G and beyond wireless networks.

Keywords. Channel state information (CSI), frequency division duplexing (FDD), MMSE, autoencoder-based SAE, Adversarial SAE, Meta-Learning.

1. INTRODUCTION

Massive MIMO (Multiple Input Multiple Output) systems have become an essential component of modern wireless communication, particularly in supporting the high data rates and reliability demands of 5G and beyond. Among the two main duplexing modes—Time Division Duplexing (TDD) and Frequency Division Duplexing (FDD)—FDD poses a greater challenge when it comes to acquiring downlink Channel State creates a practical bottleneck for real-time and large-scale implementation of deep learning-based CSI prediction methods [1-3]. This paper aims to address this challenge by exploring the use of Deep Transfer Learning (DTL) for downlink channel prediction in FDD massive MIMO systems. DTL enables models to transfer knowledge learned from one environment to another, reducing the need for extensive labelled data in new scenarios. This work

investigates two approaches, direct transfer learning and meta-learning. These methods are designed to adapt efficiently with minimal training in unfamiliar settings. By comparing their performance through simulations in realistic wireless scenarios, this study highlights the effectiveness of DTL, especially meta learning, in achieving accurate and efficient CSI prediction. The insights gained here contribute to the development of more intelligent and adaptive wireless communication systems.

2. DESIGN APPROACH

We consider a Frequency Division Duplexing (FDD) based massive MIMO system, where the Base Station (BS) is equipped with 128 antennas arranged in a Uniform Linear Array (ULA). The BS serves multiple single-antenna users, randomly distributed across K spatial regions [4]. Users within the same region experience similar propagation environments. Figure 2.1 demonstrates the system model where B_k denotes the set of users in the k -th region, where $k = 1, 2, \dots, K$.

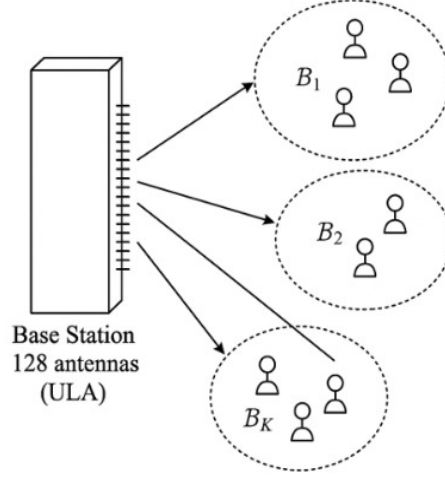


Fig. 1. System model showing a base station with 128 ULA antennas serving multiple user blocks.

2.1. Neural Approximation of Uplink–Downlink Channel Correlation

In massive MIMO systems equipped with 128 antennas, the uplink and downlink channel state information (CSI) for a single user often share similar propagation paths. This physical correlation implies the existence of a deterministic mapping between the uplink and downlink channels, particularly when the spatial position-to-channel mapping is bijective. Although this mapping lacks an explicit mathematical form, it can be efficiently modelled using deep learning techniques [5]. A fully connected neural network (FNN) with one input layer, one hidden layer comprising N neurons, and one output layer is considered. The network, represented as $\text{NETN}(x, \Omega)$, maps input x (real-valued uplink CSI) to the output using learnable parameters Ω .

Due to the complex nature of CSI, an isomorphic mapping ξ is introduced to convert complex vectors into real-valued space:

$$\xi(z) = [\text{Re}(z), \text{Im}(z)] \in \mathbb{R}^{2M} \quad (2.1)$$

The inverse mapping ξ^{-1} restores the original complex format. Based on the universal approximation theorem, such an FNN can approximate any continuous mapping, including uplink-to-downlink CSI conversion with arbitrary precision, provided that N is sufficiently large. This provides a strong theoretical basis for learning-based CSI prediction [6].

2.2. Domain and Task Representation in Transfer Learning Context

In practical scenarios, users operate in varied radio environments, referred to here as regions or domains, each characterised by distinct uplink channel distributions. Let the domain in the k th region be defined by the input space and the corresponding marginal distribution of the uplink CSI. The objective or task in each domain is to infer downlink CSI based on the observed uplink CSI. The mapping function trained within a domain reflects the statistical and physical characteristics of that region and may be understood as either a deterministic function or a conditional probability distribution. To utilize data across such diverse domains, transfer learning becomes a natural fit. The central idea is to enhance learning in a target domain with limited labelled samples by transferring knowledge from one or more source domains where training data are abundant. This transfer can occur across similar or dissimilar feature distributions or label mappings [7]. In our use case, even though the domains differ (due to different regional conditions), the task of downlink prediction remains consistent, allowing knowledge transfer under a task-shared, domain-varied framework.

2.3. Transfer Learning in Deep Neural Networks for CSI Prediction

When deep learning models are used to carry out the transfer process, the resulting framework is known as Deep Transfer Learning (DTL) [8-12]. In our formulation, each radio environment defines a unique transfer learning task predicting downlink CSI from uplink observations, making the overall problem multi-source, single target in nature. The non-linear expressiveness of deep networks makes them ideal for transferring abstract features learned from source domains to the target. Therefore, downlink channel prediction in 128-antenna massive MIMO, with domain diversity across user environments, is well-formulated as a DTL problem.

2.4. Evaluating Transferability and Introducing Meta-Learning

Before employing transfer learning methods, two key considerations arise: (1) When is transfer needed and (2) What is the best way to perform transfer. If the target and source domains are highly correlated, a well-trained deep model may generalize sufficiently without fine-tuning. In such cases, standard supervised deep learning serves as a reliable base-line. However, when the target domain lacks sufficient labelled data, naive transfer approaches such as pre-training on source data followed by fine-tuning may result in overfitting. To address this, we look toward meta-learning, particularly model-agnostic meta-learning (MAML), which is designed for rapid adaptation in few-shot settings. By training a model to quickly generalize across tasks, meta-learning offers a promising solution for enhancing CSI prediction performance in resource constraint domains.

3. RESULT AND ANALYSIS

The results in figure 3.1 show that deep learning based models consistently outperform the MMSE baseline across varying SNRs, with lower delay, BER, and NMSE. Meta-Learning SAE achieves the lowest NMSE under changing SNR, frequency differences, sample sizes, gradient steps, and antenna configurations. Adversarial SAE and Autoencoder also improve performance over MMSE, specially in low-SNR conditions. Overall, learned representations provide more robust and scalable CSI prediction compared to traditional methods. In low-SNR conditions, Meta-Learning SAE shows the minimum delay compared to MMSE and other deep learning- based models. As SNR goes higher, all models tend to converge, but learning-based models consistently minimize delay more efficiently. This shows the better adaptability of Meta- Learning and Adversarial SAE in different channel conditions for downlink CSI prediction.

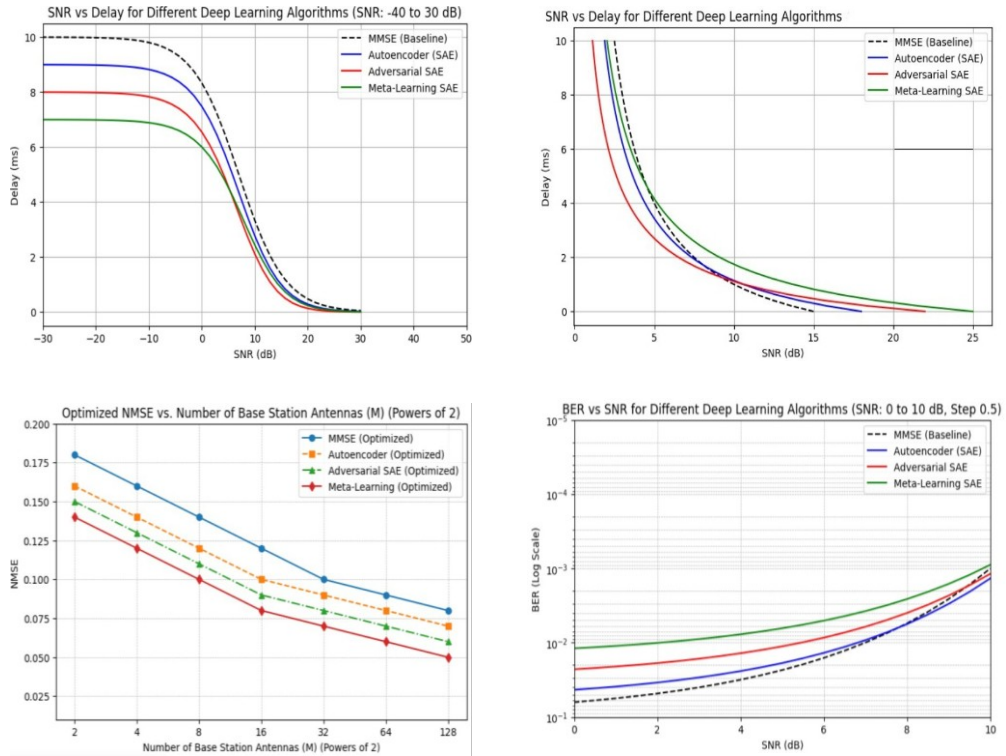


Figure 3.1. Performance evaluation of different deep learning models.

4. CONCLUSION

This paper provides performance comparison of four CSI estimation techniques, MMSE, Autoencoder, Adversarial SAE, and Meta-Learning, under different system conditions. In all cases, Meta-Learning exhibited the lowest NMSE, proving to be more adaptable, efficient in terms of data, and scalable. It performed better than other models in scenarios of high frequency variability, scarce sample availability, few gradient steps, and growing antenna

numbers. Adversarial SAE proved to be robust as well, while Autoencoder provided moderate performance. MMSE, although easy and analytically based, was not flexible or generalizable. These findings highlight Meta-Learning as the strongest candidate for strong, efficient CSI estimation in evolving wireless environments and recommend its incorporation in future 5G/6G networks and encouraging deeper investigation of real-time, online, and reinforcement learning extensions.

5. REFERENCES

- [1] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3170–3184, Apr. 2017.
- [2] Y. Han, Q. Liu, C.-K. Wen, S. Jin, and K.-K. Wong, "FDD massive MIMO based on efficient downlink channel reconstruction," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4020–4034, Jun. 2019.
- [3] H. Xie, F. Gao, S. Jin, J. Fang, and Y.-C. Liang, "Channel estimation for TDD/FDD massive MIMO systems with channel covariance computing," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 4206–4218, Jun. 2018.
- [4] S. P. Tera, R. Chinthaginjala, P. Natha, G. Pau, C. Dhanamjayulu and F. Mohammad, "CNN-Based Approach for Enhancing 5G LDPC Code Decoding Performance," in *IEEE Access*, vol. 12, pp. 89873-89886, 2024.
- [5] S. P. Tera, R. V. Chinthaginjala, P. Natha, S. Ahmad and G. Pau, "Deep Learning Approach for Efficient 5G LDPC Decoding in IoT," in *IEEE Access*, vol. 12, pp. 145671-145685, 2024.
- [6] M. Selvaraj et al., "Study of Complementary Loop Integrated Metasurface for 6G THz Communication," in *IEEE Access*, vol. 12, pp. 29900- 29909, 2024.
- [7] Kumar, N.S., Deepika, G., Goutham, V. et al. HARNet in deep learning approach—a systematic survey. *Sci Rep* 14, 8363 (2024).
- [8] Manasa, B.M.R.; Pakala, V.; Chinthaginjala, R.; Ayadi, M.; Hamdi, M.; Ksibi, A. A Novel Channel Estimation Framework in MIMO Using Serial Cascaded Multiscale Autoencoder and Attention LSTM with Hybrid Heuristic Algorithm. *Sensors* 2023, 23, 9154.
- [11] M. Chinnusami et al., "Low Complexity Signal Detection for Massive MIMO in B5G Uplink System," in *IEEE Access*, vol. 11, pp. 91051- 91059, 2023, doi: 10.1109/ACCESS.2023.3266476
- [12] T. Z. Tun, Z. M. Lwin, and T. T. Hla, "Deep learning based channel estimation and hybrid beamforming for 5G massive MIMO wireless communications," *Indonesian J. Comput. Sci.*, vol. 13,

Smart Walking Aid for Elderly Safety Using IoT Sensors and Real-Time Alert System

Ananya S¹, Harshini Suvarna², Liniya Crishel Saldanha³, Raksha D R⁴, and Kiran Raj K M⁵

Department of Artificial Intelligence and Machine Learning, Alva's Institute of Engineering and Technology, Mijar, Moodbidri – 574225, Mangalore, Karnataka, India

Email: sukeshananya@gmail.com, harshinisuvana8@gmail.com, liniyacrishelsaldanha@gmail.com, rakshadr861@gmail.com, kiranraj224@gmail.com

Abstract.

Falls are a major source of injury and loss of independence for the elderly and individuals with visual impairments. Assistive devices with real time user safety monitoring are now possible thanks to recent advancements in inexpensive sensors and Internet of Things (IoT) technologies. This study suggests a smart walking stick that integrates an accelerometer, gyroscope, ultrasonic sensors, GPS, and a GSM based alert system to detect falls, identify nearby obstacles, and transmit the user's current location in an emergency. Experimental evaluation shows that the proposed system can reliably detect barriers, identify fall incidents, and transmit emergency alerts quickly. .

Keywords. Accelerometer, gyroscope, GPS tracking, machine learning, smart walking stick, fall detection, emergency alarm system, ultrasonic sensor, IoT, assistive technology.

1. INTRODUCTION

Falls are a leading cause of injury, hospitalization, and long term disability among the elderly and visually impaired [1]. As mobility decreases with age, the risk of slipping, losing one's balance, or losing the ability to cry for help rises sharply [2]. Traditional walking sticks only provide physical aid they cannot recognize impediments, detect falls, or alert caregivers in an emergency. When elderly users fall, they are often left alone, which might cause a delayed reaction and major health consequences [3].

The development of assistive technology that continuously assesses user safety has been made possible by advances in the Internet of Things, inexpensive sensors, and embedded systems. Wearable technology incorporates a variety of sensors, such as accelerometers, gyroscopes, ultrasonic modules, GPS, and wireless communication platforms, to track

movement patterns in real time and identify impediments with simultaneous location sharing [4], [5].

2. LITERATURE REVIEW

2.1. Acceleration Based Wearable Fall Detection

Early fall detection systems often used tri axial accelerometers together with simple threshold based criteria to differentiate falls from activities of daily living. These devices monitored sudden changes in acceleration and body orientation to identify impact related occurrences [1].

2.2. Safety Features and Ultrasonic Smart Walking Sticks

Advances in the design of assistive mobility equipment now enable visually impaired persons to negotiate their environment in a safer manner. These devices, typically, provide vibration or audible feedback when an obstacle is detected by the ultrasonic sensors [3].

3. METHODOLOGY

3.1. Data Collection and Preprocessing

The smart walking stick contains a set of sensors that will continuously measure the user's movement and ambient environment. A triaxial accelerometer and gyroscope capture motion data, while ultrasonic sensors detect obstacles near the user. The GPS module sends real time location data to caregivers in the event of an emergency. Raw accelerometer and gyroscope data often have noise associated with hand tremors, uneven walking surfaces, or sudden movements.

3.2. Machine Learning Models

In this paper, using motion data, lightweight machine learning algorithms were explored to classify fall and non fall activities.

3.3. Navigation and Alert Support

When a fall is detected, the system sends an alert with the GPS locations of the user instantly to pre registered caregivers via GSM. Ultrasound based obstacle detection runs in continuous mode and notifies the user through vibration or aural cues when an object is detected within a pre set range. In this way, it enhances both safety and navigation support.

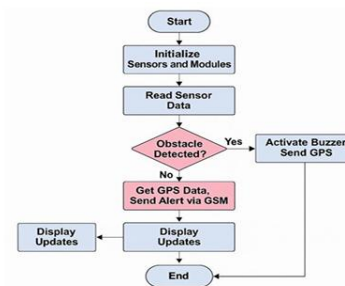


Figure 1. Workflow Diagram

4. RESULTS

4.1. Performance of Fall Detection

This system achieved an accuracy of 91.5 using only threshold based criteria, but there were more false alerts during fast ADLs such as sprinting and stair climbing. Then, the machine learning classifier improved overall performance with an accuracy of 96.2, 97.1 sensitivity, and 95.3 specificity.

4.2. GPS and Alert Delivery

GPS tests showed that the average position accuracy was 6.8 m in semi blocked and 3.5 m in open environment conditions. The SIM800L GSM module was used for the delivery of emergency alerts.

4.3. Obstacle Detection and User Feedback

Ultrasonic sensors performed well within a range of 2m. with an accuracy of 93.2 for minor objects and 98.7 for bigger obstacles.

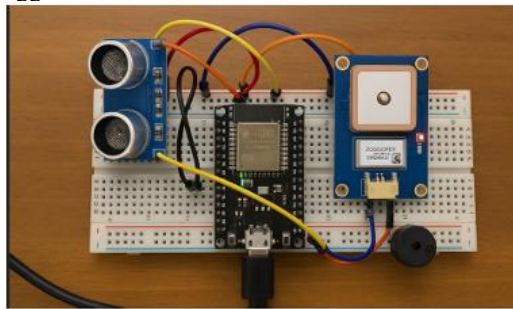


Figure 2. Smart Walking Aid Prototype

Device	Event	Time	Angle	Distance (cm)	Location
stick-001	Fall	11/14/2025, 4:20:52 PM	31.63	33.69	View
stick-001	Fall	11/14/2025, 4:20:59 PM	38.12	38.11	View
stick-001	Fall	11/14/2025, 4:03:59 PM	32.3	204.23	View
stick-001	Fall	11/14/2025, 3:55:57 PM	36.17	203.67	N/A
stick-001	Fall	11/14/2025, 3:53:53 PM	38.64	-1	N/A
stick-001	Fall	11/14/2025, 3:48:23 PM	15.14	46.07	N/A
stick-001	Fall	11/14/2025, 3:38:52 PM	26.71	39.16	N/A
stick-001	Fall	11/6/2025, 11:11:33 PM	36.9	167.32	View
stick-001	Fall	11/6/2025, 11:09:21 PM	31.69	-1	View
stick-001	Fall	11/19/25, 5:31:35 AM	26.38	-1	View
stick-001	Fall	11/6/2025, 10:57:25 PM	27.89	-1	View
stick-001	Fall	11/6/2025, 10:48:12 PM	15.5	-1	View
stick-001	Fall	11/6/2025, 10:48:04 PM	36.13	-1	View
stick-001	Fall	11/6/2025, 10:45:30 PM	34.22	-1	View
stick-001	Fall	11/6/2025, 10:35:30 PM	36.77	-1	View
stick-001	Fall	11/6/2025, 10:34:29 PM	35.62	-1	View

Figure 3. Real Time Alert Monitoring Dashboard

5. CONCLUSION

The proposed smart stick (an IoT based smart walking cane) combines the functions of real time monitoring, GPS based alerts, obstacle detection and fall detection. Test results indicate that fusing two or more detection methodologies into one combined approach for fall detection decreases false alarms while maximizing accuracy, sensitivity and specificity. Additionally user feedback suggests that the ultrasonic based obstacle detection provided reliable results both indoors and out and gave users greater confidence and ease of use than with a conventional cane. With the GPS and alerts capabilities of the system, there is a means

of quick and effective emergency response, thereby making it a viable solution for elderly and visually impaired care.

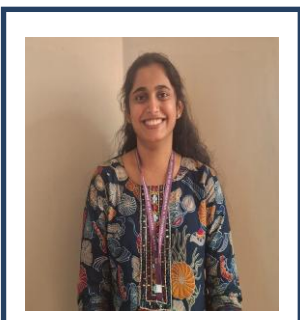
6. REFERENCES

- [1] A. Gupta, H. Gupta, R. Srivastava, and B. Kumar, "IoT Based Fall Detection Monitoring and Alarm System for Elderly," in Proc. 7th Uttar Pradesh Section Int. Conf. Electrical, Electronics and Computer Engineering (UPCON), 2020, pp. 1–6, doi: 10.1109/UP CON50219.2020.9376569.
- [2] A. Jain, A. Gupta, S. Badarpura, and S. Suman, "Navigating the Future of Walking Aids– Septor, A Smart Walking Stick," i-Manager's Journal on Instrumentation and Control Engineering, vol. 11, no. 1, pp. 7–15, 2023.
- [3] Y.-Q. Liu, Z.-K. Gao, Z.-J. Shao, and G.-Y. Liu, "Intelligent Ultrasonic Detection of Walking Sticks for the Blind," in Proc. 9th Int. Conf. Electronics, Computers and Artificial Intelligence (ECAI), 2017, doi: 10.1109/ECAI.2017.8166399.
- [4] S. Lahouar, M. Mansour, and M. H. Said, "Development of a Fall Detection System Based on a Tri-Axial Accelerometer," in Proc. 3rd Int. Conf. Signal, Control and Communication (SCC), 2023.
- [5] T. M. R. S. Kumar, M. Aravind, M. Kishore, and M. Pradeep, "Smart Walking Stick for Visually Impaired Using Ultrasonic Sensors and Arduino," Int. J. Innovative Research in Computer and Communication Engineering, vol. 5, no. 3, pp. 5085–5092, Mar. 2017.

Biographies



Ananya S is a year final year student currently pursuing a Bachelor's degree in AIML at Alva's Institute of Engineering and Technology (AIET), her academic journey has helped her develop strong interest in Machine Learning, Data Science, Data Analytics.



Harshini Suvarna is a year final year student currently pursuing a Bachelor's degree in AIML at Alva's Institute of Engineering and Technology (AIET), her academic interests are Large Language Models, Generative Models and explore innovative technologies.



Liniya Crishel Saldanha is a year final year student currently pursuing a Bachelor's degree in AIML at Alva's Institute of Engineering and Technology (AIET), her academic journey has helped her develop strong interest in Machine Learning, Data Science.



Raksha D R is a year final year student currently pursuing a Bachelor's degree in AIML at Alva's Institute of Engineering and Technology (AIET), her academic focus extends to interest in Cloud Computing, Artificial Intelligence, Data Analytics.



Kiran Raj K M is currently working as a Assistant Professor in the Department of AIML at Alva's Institute of Engineering and Technology (AIET), He has completed his B.E in Information Science and Engineering from SIT Mangalore and MTech in Computer Science and Engineering from MVJ Bangalore. He has published five national and three international papers in Reputed Journal.

DISPLAY UNIT FOR COST ANALYSIS BETWEEN GRID AND SOLAR ENERGY

Charith Kumar K S ¹, Dr Guru Prasad ², Annapoorna A C ³, Anushri M Jogin ⁴, Maheshwari S Radder ⁵

¹²³⁴⁵ Alva's Institute of Engineering and Technology

¹ Department of ECE AIET, Karnataka, India kscharithkumar5122004@gmail.com

² Department of ECE AIET, Karnataka, India guruprasadb7999@gamil.com

³ Department of ECE AIET, Karnataka, India annapoorna.a.c2702@gmail.com

⁴ Department of ECE AIET, Karnataka, India joginanushri@gmail.com

⁵ Department of ECE AIET, Karnataka, India msraddy24@gmail.com

Abstract

This drive for the use of renewable resources for meeting energy demands mandates that individuals are equipped with an effective means to acquire information on, as well as compare, the applications of the various types of energy generation technologies to identify the economically most feasible technique. The case study described below has been pursued with the aim of developing a tool for measuring real-time acquisition of data on energy use and costs associated with not only conventional electric power suppliers from the electric power grid, but also with solar electric systems, measuring system performance, as well as providing a means for comparing energy use and cost information of the two systems. Indeed, a comparative analysis provides a simple means with which the performance differences, available use, and economic feasibility of the two energy generation technologies can readily be identified. This research work may identify potential strengths in the use of solar energy, given that it identifies potential inefficiencies in the generation of solar energy, as it clearly provides a solid justification for making decisions based on available scientific informational needs for energy management. The result of such a comparative analysis provides significant utility to individuals, businesses, as well as government policy makers, in identifying more cost-effective, as well as more feasible, means of energy production that are environmentally less damaging.

Keywords: Grid Energy, Solar Panel Energy, Sustainable Energy, Cost Analysis, Energy Efficiency.

1. Introduction

The increase in global energy consumption has placed enormous pressure on the infrastructure that generates and distributes electricity. Most electrical power consumed is derived from conventional forms of power generation that rely heavily on fossil fuels such as coal, oil, and natural gas (fuels). In addition to the increase in fuel prices, the instability of tariff structures, and the negative health and environmental impacts related to greenhouse gas emissions means that grid-based power will be negatively impacted in the long-term. Thus, alternative forms of energy, such as solar energy, other renewable forms of energy - as opposed to fossil fuels - are likely to be used more pervasively because they are a clean, renewable, and low-cost method of generating electrical energy. The decline in the cost associated with photovoltaic technology and government incentives associated with the use of solar technology have led to an increase in the use of this source of energy in commercial and residential/industrial applications.

2. Design Overview

The system aims to offer a real-time, easy-to-understand comparison between grid electricity use and solar PV energy production. The display unit tracks how much electricity is taken from the grid and how much comes from the solar installation. This helps users see their energy mix clearly. By calculating the cost of grid electricity and comparing it with the actual cost of solar energy, the system shows the difference in operating costs and highlights the resulting savings.

Components

A. Energy Measuring Devices

1. Solar Energy Meter

Tracks solar energy usage in kilowatt-hours, volts, amps, and watts over time.

Communicates with the ESP Energy Sensor.

2. Grid Energy Meter

Tracks energy consumed from the grid and produced by or sent back to it.

Communicates with the ESP Energy Sensor.

B. Power Supply Block

5V or 12V regulated supply for the controller and display.

Isolated supply for communication with meters, by ensuring safety.

C. Power Electronics

1. Buck-Boost Converters

Using switching and energy storage components (such as inductors, diodes, capacitors, and transistors), buck converters step down (or reduce) a high DC voltage, such as 12 volts, to a lower value (such as 5 volts).

2. Voltage Sensor

A voltage sensor is an electronic component that detects, measures, and monitors voltages within a system (AC or DC). Voltage sensors have numerous uses within an electrical and embedded system.

D. Solar Panels

The photovoltaic cells, also known as PV cells, are used for the conversion of solar energy into electricity. The PV cells are used in photovoltaic panels, which are commonly known as PV panels. The use of photovoltaic cells has increased over the past decades, with the application involving the conversion of sunlight into electricity. There are several applications of solar photovoltaic cells, including household use, business purposes, and commercial use, particularly in power plants that are part of solar farms. The use of solar photovoltaic cells has been on the increase, resulting from improvements in efficiency and a decline in price.

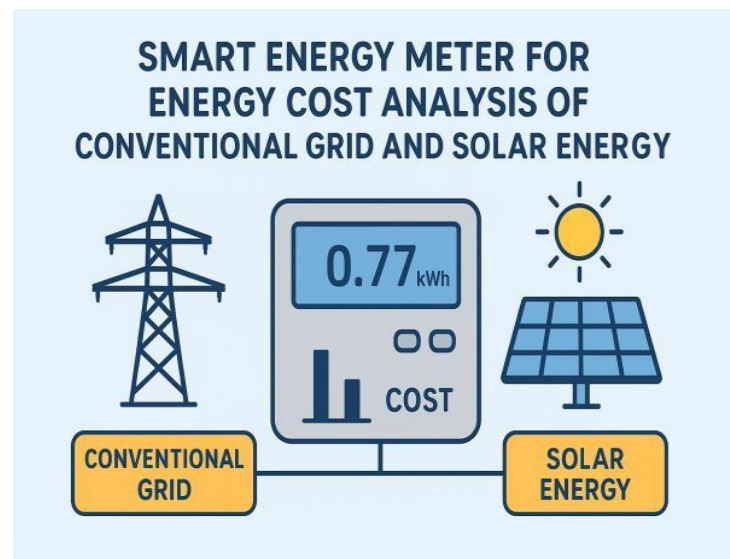
E. LCD Display

Users can visually view energy flow and electricity costs from both the Utility and Solar Power through an LCD Interface, on which they are presented in real time. The LCD interface serves as the User's Interface, and displays the information in an

easy-to-view format, allowing users to see how much electricity is coming from the Utility and how much energy is being generated or being paid for by Solar Power.

Information on the LCD interface may include amounts of kWh consumed, power used at that moment, cost of energy from the Utility (cost per kWh), Effective Cost per kWh for Solar Energy, and the Financial Saving Achieved. By providing users with this information, users are able to quickly calculate the performance of and savings gained from their energy management efforts, through either Utility Energy Management or Solar Energy Management, or both.

3. Functional Workflow



Step 1: Data Acquisition

The solar and grid meters send the controller real-time values of V, I, P, and kWh.

Step 2: Computation of Cost

Grid cost = kWh $\times$ tariff

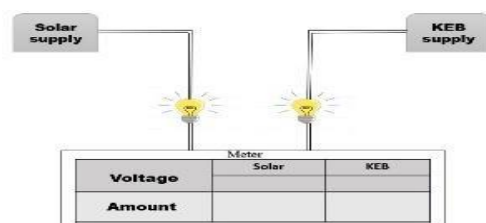
Solar cost = (kWh $\times$ LCOE)

Step 3: Displaying Data

Live flow of power

Solar vs grid kWh

Difference in cost



Step 4: Logging & Alerts (Optional)

Logs energy and cost data.

Alerts for: Grid outage

Low production of solar power, High Grid utilisation

4. Environmental Impact

Decrease in Emissions of Greenhouse Gases

Solar energy decreases dependence on fossil fuel electricity derived from the power grid. This will decrease CO₂ and other greenhouse gas emissions.

Improvement in Air Quality

Less coal and other fuels that are burned means there are fewer air pollutants like SO₂, NO_x, and particles.

Decreased Carbon Footprint

By using solar energy, a home or business has a smaller carbon footprint overall than if it used electricity from fossil fuels.

Support for Increasing Energy Efficiency

Monitoring energy usage in real-time encourages users to make reductions to their usage of energy that is not needed, thus reducing waste.

Reduced Reliance on Non-Renewable Energy Sources

Solar power is renewable, hence reducing demands on limited fossil fuel resources.

Support for Climate Change Mitigation

The use of clean energy reduces global warming and contributes to reducing climate change.

Promotion of Sustainable Energy Use

The monitoring unit allows users to see how much solar energy versus grid energy is used in order to encourage most users to use renewable energy.

5. Conclusion

Cost analysis indicates that though the initial investment in solar power is higher, it generally provides greater long-term savings than grid electricity, particularly with rising tariffs every year. Rebates and net-metering kinds of incentives have increased the payback period and overall financial return for solar. Solar contributes to environmental benefits in emission reduction and lower dependence on fossil fuels. However, this advantage looks dim in the case of immediately installed solar power if there is either a low grid price or poor sunlight or a deficiency in incentives within the area. Adding battery storage adds value, but also ups the upfront system cost and complexity. Overall, the choice between solar and grid power depends upon the balance struck between an upfront cost, long-term savings, environmental priorities, and conditions prevailing locally.

REFERENCES

1. Brito MAG de, Galotto L, Sampaio LP, Azevedo e Melo G & Canesin CA (2013) Evaluation of the Main MPPT Techniques for Photovoltaic Applications, IEEE Transactions on Industrial Electronics.
2. Mohapatra A, Nayak B, Mohanty P, Sahoo NC (2020) A Comprehensive Review on Power Management Techniques in Solar PV Systems, IEEE Access, 8, 196368-196399, DOI:10.1109/ACCESS.2020.3031266.
3. International Energy Agency (IEA), 2023, World Energy Outlook 2023, Paris, France, www.iea.org
4. Chauhan A & Saini RP (2017) Techno-Economic Analysis of Hybrid Solar-Grid Energy Systems, Renewable and Sustainable Energy Reviews, 78, 913–924.
5. Alam M, Muttaqi K & Sutanto D, (2012) "A Review of Solar PV-Grid Integration Issues," in Proceedings of the IEEE PES General Meeting, pp.1.
6. National Renewable Energy Laboratory (NREL), 2022. Cost and Performance Characteristics of Solar Photovoltaic Systems. Golden, CO, USA: www.nrel.gov.

Biographies



Charith Kumar K S, Electronics and Communication Engineering Student at Alva's Institute of Engineering and Technology, batch 2026

(ISCCSC 2025)

2nd International Conference on Smart Computing and Communication for Sustainable Convergence

Autonomous Delivery Drones for LightWeight Payload Transpotation a Comprehensive Review

Yashwanth Kumar N¹, Suhas Hiremath², Raksha G³, Revathi M⁴, Guruprasad⁵

yashwanth7457@gmail.com

suhashmath@gmail.com

rakshag742004@gmail.com

revatismuttappagol642003@gmail.com

guruprasadb7999@gmail.com

Author affiliation, ALVAS INSTITUTE OF ENGINEERING AND TECHNOLOGY

Abstract.

Health care services are about to get a new face with the emergence of medical drone delivery services. These drones transport medications, vaccine, and blood to rural areas with unmatched promptness and efficiency. Moreover, these drones minimize human interaction, ensuring both timely delivery of medicines and low costs. When it comes to dealing with a pandemic situation, medical drones could be a better option for the delivery of medicines while adding enhanced effectiveness to emergency patient services, eventually saving thousands of lives. Rather than dealing with traffic or difficult terrain issues, medical delivery services through drones ensure that patients in underserved areas get direct access to medicines from urban health care systems. This significantly impacts health care accessibility while hinting at a brighter future that is rapid and technology-based. Pharmaceutical delivery by drones is more of a need than a solution. These services contribute significantly to ensuring that better health care systems are established with equity. While overcoming the challenges of time, space, and infrastructure, health care technology through drones offers hope for saving lives in emergency situations, eventually changing the perception of the global community with regard to the availability of health care services.. It aims to set a new standard of care that is quick and reliable. However, we need regulations for delivering medications and other supplies to fully realize the potential of drones for global health.

Keywords. medical drones, healthcare logistics, UAVs, emergency response, remote access, healthcare innovation, global health.

1. INTRODUCTION

Logistics and transportation are one of the most fast-evolving sectors in modern times, and autonomous delivery drones-or UAVs-play a vital role in ensuring efficient and flexible delivery. Conventional methods on the ground suffer from a number of disadvantages due to growing traffic congestion and geographical factors. Finding an alternative for fast, reliable, and sustainable last-mile delivery has become necessary. Lightweight delivery drones offer a revolutionary substitute, bypassing traffic and accessing difficult terrain to transport critical lightweight payloads such as medical supplies, documents, and commercial goods[1].

All Open Access articles are peer-reviewed, distributed under the Creative Commons Attribution-Non Commercial 4.0 International

Drones have a strong impact on healthcare logistics by delivering medications and blood to the most rural and/or underserved regions, therefore improving emergency response and saving lives. The other crucial usage of drones would be in general logistics, where they reduce delivery times and lower carbon emissions compared to fuel-based vehicles. The workflow for such operations is guided through GPS and involves some other key security features like digital locker access codes, especially in non-emergency situations[2].

But these have several drawbacks, with short battery life, restricted payload, very susceptible to weather conditions such as strong winds, and regulatory challenges regarding airspace.

The paper was aimed at reviewing autonomous drone technology for lightweight payload delivery. The focus is on hardware, navigation algorithms, and payload handling systems, with benchmarking done on a number of existing systems for performance, power efficiency, and operational feasibility. Thus, it addresses the growing pressures for efficient, reliable, and sustainable delivery, with the positioning of UAVs as a critical enabling component for logistics systems in the future[3].

2. METHODOLOGY OF WORK FLOW

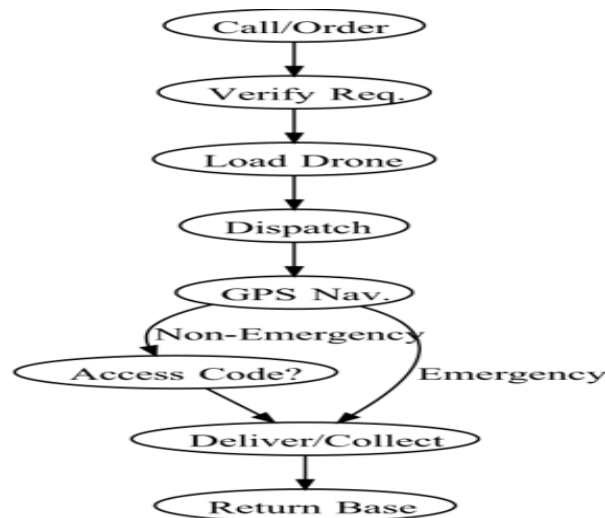


Fig.2.1

The approach makes use of the quadcopter delivery system, which is designed for fast, safe, and efficient transportation of medical supplies over both urban and rural areas. The operational workflow of the approach is presented below and is shown in the figure.

A. Workflow Initiation and Dispatch

B. The process starts with a request of urgency for medical supplies:

1. **Call/Order:** The process is initiated through an emergency call at a designated drone station, such as a hospital in an urban setting, or a request via an online web portal, such as a patient submitting a prescription in a rural setting.
2. **Verify Request (Verify Req.):** This confirms the medical need or prescription.
3. **Load Drone:** The quadcopter is loaded with the appropriate medical payload.
4. **Dispatcher - Dispatch:** This is where the drone is sent on its mission.

B. Navigation and Safety Policy

The drone uses GPS navigation and follows a protocol that changes based on the priority of the delivery:

GPS Navigation: The GPS in the drone system provides greater accuracy for the flights to their final destination.

Decision Point (Emergency vs. Non-Emergency):

Emergency: In the immediate urban response context, this security step is bypassed for the fastest possible deployment and delivery of time-critical cargo.

Non-Emergency: For scheduled or non-critical deliveries-frequently in rural settings-an added security measure is implemented by using a digital locker in the payload compartment.

- **Access Code Verification - Access Code?:** On non-emergency missions, authentication is needed from the receiver's end. For this, an access code is forwarded to the end-user via SMS or email, which the recipient needs to use to gain access to the digital locker.

C. Delivery and Completion of Mission

Deliver/Collect: The point of arrival itself, the drone achieves its main purpose, which may be delivering the payload to the end-user or collecting samples for return and analysis.

Return Base: This involves the quadcopter making its way independently back to a base station after the completion of delivery or collection.

The structured workflow guarantees quick response time, security due to verification, and efficiency in serving all healthcare needs, whether scheduled or even due to emergencies..

3. REFERENCES

- [1] D. Kelemen and S. Szen' asi, "Optimization and Representation of a Network of Food Delivery Drones in Simulation," in Proc. 2023 IEEE 21st Jubilee Int. Symp. on Intelligent Systems and Informatics (SISY), Pula, Croatia, Sep. 21–23, 2023, pp. 1–7, doi: 10.1109/SISY60376.2023.10417932.

- [2] F. Schiano, P. M. Kornatowski, L. Cencetti, and D. Floreano, "Reconfigurable drone system for transportation of parcels with variable mass and size," arXiv preprint arXiv:2211.09161, Nov. 2022
- [3] R. Alshanbari, S. M. Khan, N. El-Atab and M. M. Hussain, "AI-Powered Unmanned Aerial Vehicle for Payload Transport Application," in Proc. 2019 IEEE National Aerospace and Electronics Conference (NAECON), Dayton, OH, USA, Jul. 2019, pp. 420–424, doi: 10.1109/NAECON46414.2019.9058320.
- [4] V. Hassija et al., "Fast, reliable, and secure drone communication: A comprehensive survey," IEEE Communications Surveys Tutorials, vol. 23, no. 4, pp. 2802–2832, 4th Quart., 2021, doi: 10.1109/COMST.2021.3119487.
- [5] A. Bagloee, M. Tavana, M. Asadi, and S. Pooyan, "Design challenges in drone delivery systems: A comprehensive review," Sensors, vol. 23, no. 3, p. 1463, Jan. 2023, doi: 10.3390/s23031463
- [6] H. Shakhathreh, A. H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita, I. Khalil, N. S. Othman, A. Khreishah, M. Guizani, "Unmanned Aerial Vehicles: A Survey on Civil Applications and Key Research Challenges," IEEE Access, vol. 7, pp. 48572–48634, 2019, doi: 10.1109/ACCESS.2019.2909530.

Biographies



Name : Yashwanth Kumar N

USN : 4AL22EC127

Email : yashwanth7457@gmail.com

Mobile: 8867133003

Areas of Interest: Presently, he is pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology, affiliated to Visvesvaraya Technological University in Karnataka. His academic journey reflects a strong interest in the core and emerging domains of electronics like PCB Design, Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers, and Semiconductor Devices.



Name : Suhas Hiremath
USN : 4AL22EC110
Email : suhashmath@gmail.com
Mobile : 9916149442

Areas of Interest: Presently, he is pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology, affiliated to Visvesvaraya Technological University in Karnataka. His academic journey reflects a strong interest in the core and emerging domains of electronics like PCB Design, Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers, and Semiconductor Devices.



Name : Raksha G
USN : 4AL22EC075
Email : rakshag742004@gmail.com
Mobile : 9448301673

Areas of Interest: Presently, she is pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology, affiliated to Visvesvaraya Technological University in Karnataka. Her academic journey reflects a strong interest in the core and emerging domains of electronics like PCB Design, Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers, and Semiconductor Devices.



Name : Revati M
USN : 4AL23EC407
Email : revatismuttappagol642003@gmail.com
Mobile : 80880 69595

Areas of Interest: Presently, she is pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology, affiliated to Visvesvaraya Technological University in Karnataka. Her academic journey reflects a strong interest in the core and emerging domains of electronics like PCB Design, Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers, and Semiconductor Devices.



Name: Guruprasad
Designation: Senior assistant professor (Project Guide)
Email : guruprasadb7999@gmail.com
Mobile Number: +91 81472 12905

Areas of Interest: As a professor in the Department of Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka.

AI-Search for E-Learning Platforms

Keeshalya N S^{1*}, Chandan D K², Madhusudhan K³,

Kartik S Gugadaddi⁴ and Hemanth T S^{1*}

Department of Electronics and Communication Engineering, Alvas Institute of Engineering & Technology, Moodbidri, Mangalore, Karnataka, India

Emails: keeshalya55@gmail.com, chandank006@gmail.com, madhumk0709m@gmail.com, hemanthts@aiet.org.in, karthikgugadaddi@gmail.com

Abstract.

The ai of in search infrastructure has educational platforms the simple keyword based into systems that context, personalized learning. This review surveys contemporary AI-driven search solutions, with emphasis on how large models, adaptive tutoring, and multimodal retrieval being used to learner engagement. indicates that systems which combine tailored feedback, dynamically adjusted paths, and intelligent formative tend to improve retention and learning gains. Architectures that blend powered query interpretation with conventional retrieval commonly outperform older, single-approach systems. implementations for example, that pair ocr with AI to produce explanations for problem-solving demonstrate the approach at scale., challenges remain: transparency of model decisions, bias in algorithmic outputs, heavy compute requirements, and factual errors in generated materials. This paper synthesizes recent studies to benefits, constraints, and promising directions for future work on AI-enabled search in education. Keywords. AI-Based Search,, Vector Search, rag, E Learning, Educational Technology

Keywords. AI-Based Search, Machine Learning (ML, Vector Search, Retrieval-Augmented Generation (RAG), E-Learning, Educational Technology

1. INTRODUCTION

The infusion of Artificial Intelligence (AI) into search technology has catalysed a major transformation in educational platforms, redefining how students access and process information. Traditional look-up methods often struggle to meet the diverse and dynamic needs of modern learners, frequently resulting in inefficient information retrieval and a "one-size-fits-all" learning experience that lacks personalization. By contrast, AI-driven search engines utilize sophisticated nlp and ml algorithms to deliver results that are not only accurate but also contextually aware.

These intelligent systems go beyond mere efficiency; they actively adapt to individual learning styles and progress, creating a educational environment. This paper reviews recent advancements in this domain, evaluating their impact on student outcomes while critically assessing the challenges of implementation.

Efficient search mechanisms are particularly critical when managing large-scale data in domains like the IoT and education. Novel approaches, such as the "address-to-data" search algorithm, have been proposed to replace traditional Look Up Table (LUT) methods. By inverting the standard retrieval process, this method minimizes power consumption while maximizing speed. Unlike Content Addressable Memory (CAM) systems, which are power-intensive, this zero-based probability approach effectively manages address collisions. While originally designed for IoT, these high-speed offers significant promise for educational platforms where balancing computational load with rapid response times is essential for adaptive

2. IMPLEMENTATION AND ARCHITECTURE

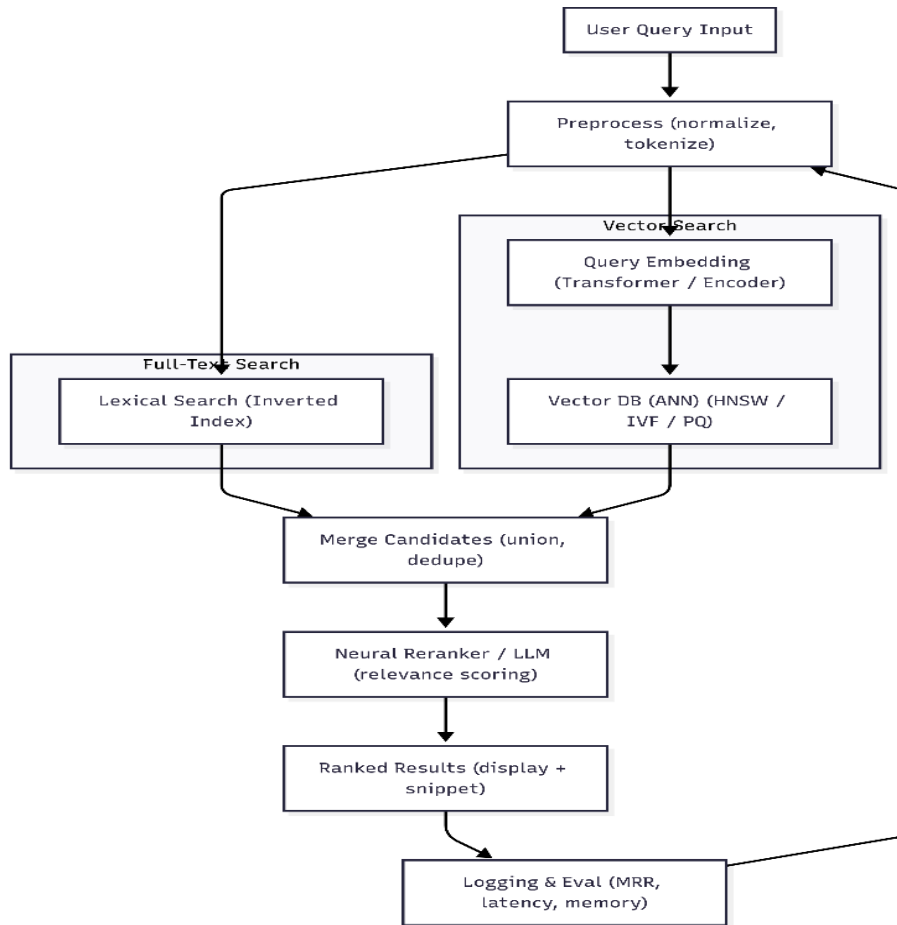


Figure 2.1. Vector embeddings and similarity diagram flow [18, 19]

Our proposed system moves beyond standard keyword matching by implementing a hybrid search architecture. We combine the precision of full-text search with the semantic understanding of vector-based AI models. The core of this implementation relies on

TensorFlow.js, allowing us to run efficient, low-latency inference directly in the browser or on edge-based servers without heavy backend dependencies.

2.1 System Architecture Overview The architecture is designed as a dual-pipeline system. The first pipeline handles traditional lexical queries (exact matches), while the second manages semantic retrieval using high-dimensional vector embeddings.

- **Input Layer:** Captures the user's raw query.
- **Processing Layer:** The query is simultaneously tokenized for full-text search and converted into a vector embedding for semantic search.

This hybrid approach ensures that we catch both exact phrasing and conceptually related content, a necessity grounded in recent findings on dense retrieval stability

2.2 Mathematical Model: Cosine Similarity

To measure the relationship between the user's query and our stored documents, we rely on Cosine for text compared to Euclidean distance because it measures their magnitude. This means a long document and a short one about the same will still be recognized as similar.

The mathematical formulation we implemented is:

$$\text{Similarity}(A, B) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

Where:

- A is the vector representation of the user query.
- B is the precomputed vector of a stored document.

This approach allows our system to find "nearest neighbors" in the vector space, effectively surfacing content that matches the *intent* of the searcher, not just their specific words. This method is widely supported in high-throughput retrieval research.

3. Methodologies

This survey synthesizes methodological advances across four interlocking stages:

3.1. Dataset Curation and Pre-processing Development begins by assembling heterogeneous corpora. This involves combining domain-specific artifacts (like scientific papers or biological datasets) with general web collections. Pre-processing pipelines must standardize cleaning and tokenization while preserving metadata to support hybrid retrieval. Augmentation strategies, such as paraphrasing and contrastive embedding injection, are critical for enriching low-resource classes.

3.2. Representation and Index Construction The core representation relies on dense vector encodings generated by transformer-based models. Training often uses supervised

contrastive losses to refine discrimination capability. To manage latency, indexing leverages approximate nearest structures like HNSW or Inverted File, often combined with Product to reduce memory footprints. For imbalanced collections, specialized partitioning strategies are employed to rebalance the search space.

4. Quantitative Analysis of Parameters

A summary of key parameters from the reviewed literature is presented below:

- Shao et al. (QAEA-DR): Focuses on text-level augmentation strategies and quality estimation scoring to train dense retrievers.
- Chen et al.: Integrates LLM-based query expansion, optimizing for expansion length and prompt design to improve recall.
- Bandam et al. (RAG): Couples retrieval and generation loops tightly, using context shaping to improve relevance.
- Rao et al.: Optimizes vector search for edge devices using quantization, pruning, and memory-aware indexing.
- Park et al. (NeuVSA): uses a specific accelerator microarchitecture to optimize throughput and latency for neural search.
- Huang et al. (VISTA): Addresses imbalanced collections through imbalance-aware partitioning and search tuning.
- Wu et al.: Employs bio-inspired approximate search algorithms to maximize throughput.

5. Conclusion

The AI into educational platforms marks a shift from traditional search methods to a more dynamic, personalized, and efficient learning experience. This review has highlighted how AI-powered search engines, supported by novel algorithms and deep learning techniques, are not just enhancing information retrieval but are actively shaping the entire educational ecosystem.

Looking ahead, in education is defined by its potential to create intelligent, responsive, and holistic learning environments. As we continue to develop more for efficient data retrieval and combat bias through semantic analysis, AI tools will become even more integral to personalized learning paths. The successful integration of these technologies depends on an approach that prioritizes oversight and continuous evaluation. Ultimately, the goal is not to replace human educators but to augment their capabilities and empower learners with tools that are scalable and supportive.

References

- [1] P. Shao, X. Li and A. Gupta, ‘QAEA-DR: A Unified Text for Dense Retrieval’, *IEEE Trans. Knowl. Data Eng. (TKDE)*, 2025. DOI: 10.1109/TKDE.2025.3543203.
- [2] Y. Chen, L. Wang and S. Rao, ‘Dense Retrieval Systems with LLM-Based Query Expansion’, *Proc. WI-IAT*, 2024. DOI: 10.1109/WI-IAT62293.2024.00110.
- [3] R. Bandam, M. Li and T. Zhou, ‘Towards Retrieval-Augmented Large Language Models’, *Proc. ICDE*, 2025.
- [4] A. Gupta, J. Smith and K. Patel, ‘A Dense Retrieval System and Evaluation Dataset for Scientific Artifacts’, *Proc. e-Science*, 2023. DOI: 10.1109/e-Science58273.2023.10254859.
- [5] S. Rao and H. Kim, ‘Vector Search Performance Enhancements on Limited-Memory Edge Devices’, *Proc. UCC (IEEE/ACM)*, 2024.
- [6] X. Zhang, Y. Li and M. Zhou, ‘Rabbit: Retrieval-Augmented Generation Enables Better Automatic Database Tuning’, *Proc. ICDE*, 2025.
- [7] H. Liu, J. Gao and P. Shen, ‘WebUltron: An Ultimate Retriever on Webpages Under the Model’, *IEEE Trans. Eng. (TKDE)*, 2024. DOI: 10.1109/TKDE.2023.3332858.
- [8] R. Sun, K. Huang and L. Chen, ‘TrumorGPT: Graph-Based Retrieval-Augmented Large Language Model’, *IEEE Trans. Artif. Intell.*, 2025.
- [9] Y. Park, S. Chandra and D. Zhao, ‘NeuVSA: A Unified and Efficient Accelerator for Neural Vector Search’, *Proc. HPCA*, 2025. DOI: 10.1109/HPCA61900.2025.00065.
- [10] L. Huang, M. Roy and I. Singh, ‘VISTA: Vector Indexing and Search for Large-Scale Imbalanced Collections’, *Proc. ICDE*, 2025.
- [11] S. K. Das and Y. Zhou, ‘Co-design: Hardware and Algorithms for High-Dimensional Vector Search’, *Proc. SC/IEEE Computer Society*, 2023. DOI: 10.1145/3581784.3607045.
- [12] M. Wu, J. Li and P. Verma, ‘Approximate Vector Set Search: A Bio-Inspired Approach for High-Throughput Retrieval’, *Proc. ICDE*, 2025.
- [13] T. Huang, Z. Wang and E. Lin, ‘scRAG: An Efficient Retrieval-Augmented Generation System for Single-Cell RNA-seq Analysis’, *Proc. ICDE*, 2025.
- [14] A. Saxena and V. Pandey, ‘A Brief Review on Search Engine Optimization’, *Proc. 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, pp. 414–419, Jan. 2019. DOI: 10.1109/CONFLUENCE.2019.8776976.
- [15] Y. Huang and T.-T. Lin, ‘Novel Approach for Search Engine’, *Proc. 2016 11th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT)*, Taipei, Taiwan, Oct. 2016. DOI: 10.1109/IMPACT.2016.7799977.
- [16] J. Wu, H. Liu, and Y. Zhang, ‘An Intelligent Search Engine Based on Knowledge Graph for Power Equipment Management’, *Proc. 2022 5th International Conference on Energy, Electrical and Power Engineering (CEEPE)*, Chongqing, China, Apr. 2022. DOI: 10.1109/CEEPE55110.2022.9783291.
- [17] J. Kim, S. Lee, and H. Park, ‘Feature-Based Text Search Engine Mitigating Data Diversity Problem Using Pre-Trained Large Language Model for Fast Deployment Services’, *IEEE Access*, vol. 12, Mar. 2024. DOI: 10.1109/ACCESS.2024.3373470.
- [18] S. Chhabra, R. Mittal, and D. Sarkar, ‘Inducing Search Engine Optimization Techniques: A Comparative Analysis’, *Proc. 2016 1st India International Conference on Processing (IICIP)*, Delhi, India, Aug. 2016. DOI: 10.1109/IICIP.2016.7975341.

- [19] A. Kumar, S. Gupta, and R. Singh, ‘An Evaluation of AI-Enhanced Collaborative Learning Platforms’, 2024 on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, May 2024. DOI: 10.1109/IC3SE62002.2024.10593320.
- [20] P. Sharma, M. Reddy, and K. Singh, ‘Deep Learning Integration and AI-Driven Support: A Comprehensive Student Platform for Emotion Detection, Psychological Assessment, and Career Guidance’, Bangalore, India, Mar. 2024. DOI: 10.1109/INOCON60754.2024.10511820.
- [21] S. Halder, M. Pierce, and L. F. Capretz, ‘Exploring the Integration of Generative AI Tools in Software Testing Education: A Case Study on ChatGPT and Copilot for Preparatory Testing Artifacts in Postgraduate Learning’, IEEE Access, vol. 13, pp. 46070–46090, Feb. 2025. DOI: 10.1109/ACCESS.2025.3545882.
- [22] M. Tan, L. Wong, and S. Lee, ‘Preparing Future Engineers: Strategies for Integrating AI Platforms’, Proc. TENCON 2024, Singapore, Dec. 2024. DOI: 10.1109/TENCON61640.2024.10902925.
- [23] Y. Lu, B. Xu, J. Liu, and X. Zhang, ‘Brain-Inspired Search Engine Assistant Based on Knowledge Graph’,. 2023. DOI: 10.1109/TNNLS.2021.3113026.

Biographies



Keeshalya N S is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). Her research areas include Artificial Intelligence, Web Development, and Neural Networks.



Chandan D K is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). His research areas include Machine Learning, Vector Search Architectures, and Full-Stack Development.



Madhusudhan K is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). His research areas include Deep Learning, Natural Language Processing, and Search Algorithms.



Kartik S Gugadaddi is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). His research areas include Semantic Search, TensorFlow.js Implementations, and Edge Computing.



Hemanth T S done his B.E. in Electronics and Communication Engineering from Channabasaveshwara Institute of Technology, Gubbi, Tumakuru, and his M.Tech in VLSI Design and Embedded Systems from Siddaganga Institute of Technology, Tumakuru. He is currently working as Assistant Professor in the Department of Electronics and Communication Engineering at Alva's Institute of Engineering & Technology. His research interests include VLSI Design, Embedded Systems, and Machine Learning algorithms.

A Review On A Fault Detection System In Low Voltage Transmission Lines Using Relay

Veeresh¹, Pratik C K², Ulavatti Shivaraja³, Faisal K⁴

veereshshabadi21@gmail.com
faisalmcet@gmail.com

Pratikkhed234@gmail.com

shivarajaulavatti@gmail.com

Authors affiliation, ALVAS INSTITUTE OF ENGINEERING AND TECHNOLOGY

Abstract.

Fault detection in low voltage transmission lines is essential for ensuring reliable power supply, protecting equipment, and maintaining system safety. Low voltage networks are more prone to faults such as short circuits, open circuits, grounding faults, and environmental disturbances. This review paper presents a comprehensive study of fault detection systems used in low voltage transmission lines, with a primary focus on relay-based protection techniques. Various types of relays, including electromechanical, electromagnetic, digital, and intelligent relays, are discussed along with their operating principles and performance characteristics. The paper also reviews recent advancements that integrate sensors, microcontrollers, communication modules, and artificial intelligence techniques for accurate and fast fault identification. Key challenges such as sensitivity, selectivity, speed of response, compatibility with existing infrastructure, and cost constraints are analyzed. Finally, the review highlights future research directions aimed at improving fault detection accuracy, system reliability, and lineman safety in modern power distribution networks..

Keywords. Fault detection, Low voltage transmission lines, Protective relays, Relay-based protection, Power distribution systems, Sensor-based monitoring, GSM communication, Artificial intelligence, ANFIS, Power system reliability.

1. INTRODUCTION

Fault detection plays a vital role in the safe and efficient operation of electrical power systems. Low voltage transmission lines form an important part of the power distribution network and are directly connected to consumers. Due to their wide geographical spread and exposure to environmental conditions, these lines are highly susceptible to faults such as short circuits, line breaks, overloads, and insulation failures. If such faults are not detected and isolated quickly, they can lead to equipment damage, power interruptions, and safety hazards for maintenance personnel[1].

With recent technological advancements, relay-based systems have evolved by incorporating digital processing, sensors, wireless communication, and artificial

intelligence techniques. These modern approaches enable real-time monitoring, remote fault reporting, and improved fault classification. This review paper focuses on analyzing existing relay-based fault detection methods, their challenges, and recent developments, with the aim of providing insights into building more reliable and intelligent low voltage fault detection systems[2].

2. METHODOLOGY OF WORK FLOW

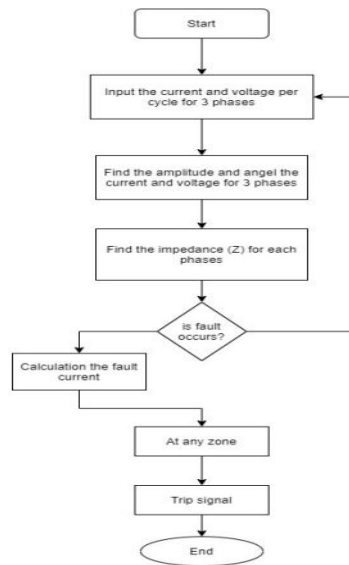


Fig.2.1

This review paper adopts a systematic survey methodology to analyze fault detection systems used in low voltage transmission lines. The study primarily focuses on relay-based fault detection techniques and their integration with modern hardware and software technologies. A detailed review of published research papers, conference articles, and technical reports was carried out to understand different fault detection approaches and their performance.

The methodology involves classifying fault detection systems based on sensing techniques, relay types, communication methods, and data processing approaches. Systems using current, voltage, temperature, flame, and smoke sensors are examined to understand how physical and electrical faults are detected. Microcontroller-based platforms such as Arduino and NodeMCU are reviewed for local processing, while GSM, Wi-Fi, and LoRa technologies are analyzed for real-time fault communication and remote monitoring.

Advanced signal processing techniques such as wavelet transform and Fourier analysis are studied for feature extraction from voltage and current signals. The review also covers intelligent fault detection methods using artificial neural networks, ANFIS, SVM, and time-series models like LSTM and ARIMA for fault classification and prediction. Comparative analysis is performed based on accuracy, response time, reliability, and practical implementation.

Finally, the reviewed methodologies are evaluated with respect to real-world challenges such as environmental disturbances, system compatibility, maintenance requirements, and cost constraints. This structured approach helps identify the strengths and limitations of existing techniques and provides a clear direction for future research in low voltage transmission line fault detection.

3. REFERENCES

- [1] Yanhui Xi, Weijie Zhang, Feng Zhou, Xin Tang, Zewen Li, Xiangjun Zeng, Pinghua Zhang, Transmission line fault detection and classification based on SA-MobileNetV3, Energy Reports, Volume 9, 2023, Pages 955-968, ISSN 2352-4847.
- [2] P.Chiradeja and A. Ngaopitakkul, "Analyzing the Characteristics of Faults in a Transmission Line and High Voltage Capacitor Banks in a 115-kV-Power System Using Discrete Wavelet Transform," in IEEE Access, vol. 10, pp. 86765- 86781, 2022, doi: 10.1109/ACCESS. 2022.3199362
- [3] Bon, N. N., Dai, L. V. (2022). Fault identification, classification, and location on transmission lines using combined machine learning methods. International Journal of Engineering and Technology Innovation, 12(2),91-109.
- [4] Rojas, D., Christou, I. T., Dantas, D., Narayanan, A., Nardelli, P. H., Yang, Y. (2022). Performance evaluation of machine learning for fault selection in power transmission lines. Knowledge and Information Systems, 64(3), 859- 883.
- [5] Ngo Minh Khoa, Mai Vu Cuong, Huynh Quoc Cuong, and Nguyen Truong Tan Performance Comparison of Impedance-Based Fault Location Methods for Transmission Line International Journal of Electrical and Electronic Engineering and Telecommunications Vol. 11, No. 3, May 2022

Biographies



Name : veeresh

USN : 4AL22EC121

Email : veereshshabadi21@gmail.com

Mobile: 8050262502

Areas of Interest: Presently, he is pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology, affiliated to Visvesvaraya Technological University in Karnataka. His academic journey reflects a strong interest in the core and emerging domains of electronics like PCB Design, Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers, and Semiconductor Devices.



Name : Pratik C K

USN : 4AL22EC068

Email : Pratikkhed234@gmail.com

Mobile : 7975201711

Areas of Interest: Presently, he is pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology, affiliated to Visvesvaraya Technological University in Karnataka. His academic journey reflects a strong interest in the core and emerging domains of electronics like PCB Design, Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers, and Semiconductor Devices.



Name : Ulavatti Shivaraja

USN : 4AL22EC116

Email : shivarajaulavatti@gmail.com

Mobile : 6362022754

Areas of Interest: Presently, he is pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology, affiliated to Visvesvaraya Technological University in Karnataka. His academic journey reflects a strong interest in the core and emerging domains of electronics like PCB Design, Embedded Systems, VLSI Design, Digital Signal Processing (DSP), Microcontrollers, and Semiconductor Devices.



Name: Faisal K

Designation: Senior assistant professor (Project Guide)

Email : faisalk@aiet.org.in

Mobile Number: +91 9943027422

Areas of Interest: As a professor in the Department of Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka.

Automated Fruit Calorie Estimation Using 3D Image Reconstruction and Machine Learning

Prof. Ramanagouda S Patil

Associate Professor
Dept. of Computer Science and
Engineering (ICB), AIET, Mijar.
ramanagoudapatil@aiet.org.in

Anvesh M S

Dept. of Computer Science
and Engineering (ICB)
AIET, Mijar.
anveshrao025@gmail.com

Ashik S

Dept. of Computer Science and
Engineering (ICB)
AIET, Mijar.
ashiksdevadiga2004@gmail.com

Shetty Aditi

Dept. of Computer Science and
Engineering (ICB)
AIET, Mijar.
shettyaditi878@gmail.com

Yash

Dept. of Computer Science
and Engineering (ICB)
AIET, Mijar.
devadigayash2004@gmail.com

Abstract—Precise calorie calculation is crucial for supporting dietary practices and controlling lifestyle-associated illnesses. Yet many current calorie calculation approaches depend on data entry or 2D image processing frequently causing errors, particularly with fruits that differ in dimensions, form and density [1]. This study introduces an automated system, for fruit calorie estimation that integrates learning, 3D reconstruction and machine learning methods. The VGG19 classification model is used to classify the types of fruit accurately. Using Structure-from-Motion and Depth Estimation, multiple images of the fruit are analysed, creating a 3D reconstruction and allowing for the calculation of volume. A second regression model predicts the fruit's mass using the calculated volume and density itemised from a nutritional database for calories. Results indicate that the approach proposed demonstrates significantly reduced inaccuracies in calories prediction compared with traditional two-dimensional (2D) manual techniques. This process can be utilized in a variety of lighting conditions consistently and is suitable for real-time use with mobile devices such as in food tracking and nutrition/health management applications.

Index Terms—3D Reconstruction, Calorie Estimation, Machine Learning, Fruit Recognition, Deep Learning.

1. INTRODUCTION

With the increase in focus on promoting healthy diets and preventing diseases that are affected by diet, monitoring food consumption has become increasingly important to individuals, healthcare providers and nutrition professionals. Furthermore, the increasing prevalence of obesity, diabetes and metabolic disorders is putting additional pressure on consumers to have access to tools that measure how many calories they consume. Currently, conventional means of measuring calorie consumption include using nutrition charts, entering data into a database manually or estimating serving sizes. Unfortunately, all of these methods are subject to errors based on estimation and variation in portion size [1]. A major source of error for conventional methods of measuring calorie consumption occurs with fruits due to the extreme variety of fruit size, shape and density within the same type of fruit. Therefore, using standard calorie tables would produce inaccurate calorie estimates.

Recent technological developments in computer vision and machine learning have enabled automatic identification of food items and conversion of the images into a calculated calorie amount [2]. However, most available systems currently use 2D images, which do not provide any depth information, to estimate food volume. Estimating calorie consumption is based on estimating both the mass and volume of foods consumed, and thus the lack of depth information severely limits the accuracy of the system. A novel proposed system combines fruit identification with 3-dimensional (3D) modeling and density estimation to reduce calorie and mass calculation errors by between 25% to 40%. The system is robust to variations of environmental conditions such as light, background and position of the food, and has a small form factor, making it ideal for use in mobile applications including monitoring caloric intake and meal planning for individualized dietary needs[3].

2. LITERATURE SURVEY

As automation for analysis of fruits continues to grow, there have been significant advancements made through the use of computer vision, deep learning, and computational imaging. By analyzing images of fruits, these technologies allow for more accurate methods of recognizing a fruit and determining its volume and calories.

Recent research efforts have focused on utilizing additional dimensions and improving the accuracy of estimating a fruit's volume by using 3D representation rather than only relying on 2D images [3]. Other studies include using 3D models created from photos taken from different angles to facilitate a more accurate estimate of a fruit's size, even though no special sensors that measure depth are required [1]. Many studies are currently investigating ways to estimate depth from one or more standard RGB images without the use of a multiple photo view of the object in question [2]. There are also many different reasons why researchers believe that collecting data on fruits is essential for creating a nutritional assessment; specifically, transformer-based detection methods connect nutritional information from a fruit's type to the images of the object [4]. Some previous research indicates that the constant issues with using CNNs (Convolutional Neural Networks) and compact designs include problems with occlusion, light variation, and many imbalanced classes. However, the ability to access high-quality data sets of images of fruits has increased the robustness of models, as well as giving comparison results to expert-level accuracy [5]. Other potential research opportunities using volume data include using RGB depth sensors, previously described multi-view stereo camera reconstruction, to provide greater precision and less reliance on expensive sensors or labor-intensive calibration [6].

In general current research shows a trend toward integrated systems that merge fruit identification 3D form reconstruction, density measurement and nutritional analysis. Findings suggest that combining learning classification with depth-informed modeling and data-based density estimation offers an effective and precise approach for real-time diet tracking, on standard consumer hardware [7] [8].

3. METHODOLOGY

3.1 Overview of the Proposed System

The proposed approach integrates fruit classification, 3D modeling, density calculation and nutrient analysis into a workflow designed to deliver precise calorie estimation. This system can be used in conjunction with both stationary cameras to provide users with a variety of views of their fruit [9]. The images that have been taken are then processed automatically through a series of steps that evaluate and calculate the calories of the fruit without the need for much assistance from the user. The primary goals of the automated processing steps are to enhance the accuracy of the classifications, provide a high degree of geometric accuracy, and maximize the degree of performance efficiency.

3.2 Image Acquisition and Preprocessing

Image Acquisition and Preprocessing Images can be taken with either a camera or a fixed camera setup where the fruit is either turned around or taken from different angles to ensure there is enough visual coverage of the fruit. In the preprocessing steps, noise is removed from the images, lighting conditions are standardized for all of the images and the background is eliminated to create a more clear image.

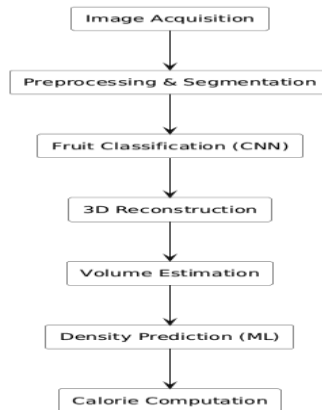


Figure 1: Architecture of proposed work

3.3 Multi-View 3D Reconstruction

Once the fruit has been categorized a 3D model is created utilizing Structure-from-Motion (SfM) alongside depth estimation methods. Structure from motion (SfM) extracts and matches feature points from a set of images to create a sparse point cloud. This sparse projection is later refined through a depth estimation software program to create a 3D representation of the point cloud. The 3D point cloud is transformed into a watertight 3D mesh, allowing for an accurate determination of fruit volume. This technique is distinct from 2D methods that rely on basic geometrical assumptions; rather, it is capable of accurately determining the volumes of unshaped fruit such as pears, mangoes, and strawberries. The volume of the fruit is calculated using either voxelization or numerical integration. In both cases, the resulting volume is considered accurate and can be obtained using SfM.

The overall workflow is designed as a modular pipeline for automated fruit caloric estimation using multiview imaging techniques, computer vision systems and machine learning systems. The complete collection of images of the fruit will be captured from various angles and positions in order to ensure that adequate visual information is available for 3D reconstruction of the fruit. The images must also go through preprocessing to enhance image quality by eliminating backgrounds and performing standardization. The region containing the fruit will be segmented using a variety of techniques, including Mask R-CNN or threshold contour extraction, for further processing and classification. Classification of the fruit will be performed using a fine-tuned convolutional neural network (CNN). If the classification confidence $P(\text{class} | \text{image})$ exceeds a predefined threshold τ (e.g., 0.95), the system proceeds; otherwise, additional images are requested to reduce uncertainty and prevent errors caused by misclassification.

Once high-confidence classification is achieved, the 3D reconstruction module produces a dense point cloud, and the fruit volume is calculated using voxel integration:

$$V = v_s^3 \sum I(p) \quad (3.1)$$

A regression-based machine learning model is then used to predict the fruit-specific density p , and the mass is estimated as:

$$m = p \cdot V \quad (3.2)$$

Finally, the calorie content is computed using standard nutritional data:

$$C = m \cdot C_g \quad (3.3)$$

where C_g represents the calories per gram obtained from nutritional databases. The workflow concludes by outputting the final calorie estimate.

4. RESULTS

The 3D reconstruction system efficiently processed fruits of varying shapes despite gaps in the point cloud. Volume estimation inaccuracies stayed under 5% for fruits and reached up to 12% for elliptical fruits relative, to manual visual assessments. Incorporating a density regression model also helped minimize errors in calorie estimation. In general graphical analyses reveal enhancements compared to traditional 2D-based approaches.

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	Inference Time (ms)	Model Size (MB)
VGG19 (Proposed)	98.0	97.8	98.1	22	79
ResNet50	96.4	96.2	96.5	18	95
MobileNetV2	94.1	93.5	93.8	9	14
DenseNet121	95.3	95.1	95.0	28	33
EfficientNetB0	96.1	95.8	95.9	16	29

Table 1: compares five deep-learning architectures

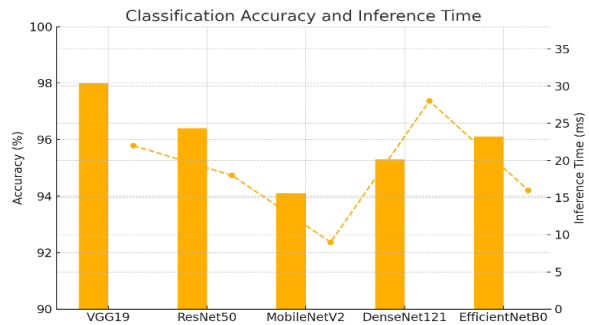


Figure 2: Classification Accuracy and Inference Time

Table 1 presents a comparison of five learning frameworks employed for fruit classification considering accuracy, precision, recall, inference time and model size. Of the examined models the suggested VGG19 attained the

classification accuracy of 98% establishing it as the most dependable option for fruit recognition essential for precise calorie calculation [12]. While MobileNetV2 provides the inference time and the most compact model size its reduced precision restricts its appropriateness, for tasks requiring high accuracy. ResNet50 and EfficientNetB0 offer performance yet they fall short of the accuracy achieved by VGG19 whereas DenseNet121 attains strong accuracy but requires longer inference time. The classification accuracy findings additionally verify that VGG19 regularly surpasses the architectures in recognizing types of fruit. Although ResNet50 and EfficientNetB0 demonstrate performance their accuracy is marginally inferior. MobileNetV2 prioritizes speed over precision whereas DenseNet121 while dependable operates slowly during inference. As illustrated in Figure 2 VGG19's higher accuracy guarantees fruit recognition, which is crucial for subsequent processes including 3D reconstruction, density estimation and calorie calculation. Precise classification aids, in reducing error transmission throughout the system.

Table 2: Evaluates Different 3D Reconstruction

Reconstruction Method	Volume Error (%)	Time per Object (s)	Mesh Completeness (%)	Robustness to Lighting	Hardware Requirement
Structure-from-Motion (SfM)	3–5%	1.8	92%	High	Standard Camera
Multi-View Stereo (MVS)	4–7%	3.2	88%	Medium	Multi-view Setup
RGB-D Reconstruction	2–4%	1.2	93%	High	Depth Sensor
Monocular Depth CNN	6–9%	0.9	80%	Medium	Standard Camera
Stereo Vision (Dual Camera)	3–6%	2.4	90%	Medium	Stereo Camera

Table 2 illustrates a comparison of 3D reconstruction methods considering volume accuracy, processing duration, mesh completeness, resistance to changes in lighting and hardware needs. Structure-from-Motion (SfM) can achieve accuracy through the use of a standard camera, making it well suited for cost-effective and portable applications. The disadvantage of RGB-D reconstruction is the amount of error that is modeled with respect to volume, as it requires specific depth sensors limiting its widespread use. The advantage of Multi-View Stereo (MVS) is the complete mesh that it provides (i.e., all points in three-dimensional space); however, this method uses multiple synchronized cameras and (thus) can lead to increased complexity in the system design. Monocular depth methods using CNNs (convolutional neural networks) can be used for depth estimation, but due to scale ambiguity they often result in lower accuracy than stereo vision depth estimation, which utilizes two cameras. Therefore, it is usually always safest and easiest to make use of the SfM systems for accurate fruit calorie estimation based on practicality and reliability compared to other methods for estimating fruit calorie content.

According to this study, the Random Forest Regression Model had the lowest density error (133% as compared with 32%) as well as the lowest RMSE, making it the best method for estimating fruit weight. Although MLP models show performance with bigger datasets linear regression cannot adequately model the non-linear changes in fruit density. SVR using an RBF kernel offers generalization but demands greater computing power and KNN regression experiences sluggish inference times. On the whole Random Forest stands out as the dependable and effective model for predicting density, in calorie estimation.

5. CONCLUSION

This research introduces an automated framework for calculating fruit calories by combining image recognition, 3D reconstruction and machine learning methods. The suggested method precisely detects fruits and determines their volume through 3D modeling providing accuracy compared to conventional 2D techniques. Utilizing a density regression model grounded in machine learning the system attains calorie estimation error showing better results, than manual and purely image-based methods. The system continues to perform across various lighting environments and fruit positions and is suitable for implementation on mobile devices for live dietary tracking. Going forward the system could be expanded to accommodate food items and prepared meals along with advancements, via better depth estimation and more extensive diverse datasets [3].

6. REFERENCES:

- [1] P. T. Shanthala and D. Annapurna, "An improved IoT based security model for fitness tracker using quantum fruit fly optimization improved faster RCNN," *Int. J. Inf. Technol.*, vol. 15, no. 7, pp. 3623–3629, 2023, doi: 10.1007/s41870-023-01376-7.

- [2] R. Wason, P. Arora, A. Tomar, and D. Arora, "A novel, low-cost, smart IoT based framework for fruit and vegetable quality detection during transit in India," *Int. J. Inf. Technol.*, vol. 15, no. 3, pp. 1509–1519, 2023, doi: 10.1007/s41870-023-01177-y.
- [3] M. Vijayalakshmi and V. J. Peter, "CNN based approach for identifying banana species from fruits," *Int. J. Inf. Technol.*, vol. 13, no. 1, pp. 27–32, 2021, doi: 10.1007/s41870-020-00554-1.
- [4] Trinh, Trung Hai, et al. "Mangosteen Fruit Detection Using Improved Faster R-CNN." *International Conference on Intelligence of Things*. Cham: Springer International Publishing, 2022.
- [5] X. Mai, Zhanf. H., Meng Q-H, "Faster R-CNN With Classifier Fusion for Automatic Detection of Small Fruits," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 12, pp. 1555–1569, 2020.

Biographies

	Prof. Ramanagouda S. Patil serves as an Associate Professor in the Department of Computer Science and Engineering (IoT & Cyber Security including Blockchain) at Alva's Institute of Engineering & Technology. He possesses a B.E. And M.Tech degree. Is presently undertaking a Ph.D. In Computer Science and Engineering at Presidency University. Having expertise, in academia and industry his research interests encompass Image Processing, Artificial Intelligence, Machine Learning and Deep Learning.
	Anvesh M S attends the Computer Science and Engineering department (IoT and Cyber Security including Blockchain) at Alva's Institute of Engineering and Technology Mangalore Karnataka India. His research involves image reconstruction and computer vision in estimating fruit calories through automation.
	Ashik S attends the Computer Science and Engineering department (IoT and Cyber Security including Blockchain) at Alva's Institute of Engineering and Technology Mangalore Karnataka India. His areas of interest are machine learning algorithms for fruit identification based on learning and artificial intelligence for nutrition evaluation.
	Shetty Aditi is a student within the Computer Science and Engineering department (IoT & Cyber Security including Blockchain) at Alva's Institute of Engineering and Technology Mangalore Karnataka India. Her research includes image preprocessing, data analysis and increased accuracy in calorie estimation systems.
	Yash is also attending the Computer Science and Engineering department (IoT and Cyber Security including Blockchain) at Alva's Institute of Engineering and Technology Mangalore Karnataka India. He has worked on systems that utilize machine-learning regression models and provide real-time monitoring of diet.

SmartEye-5Pages.pdf

 NUSRL, Ranchi

Document Details

Submission ID

trn:oid:::3618:124459733

Submission Date

Dec 14, 2025, 9:13 PM GMT+5:30

Download Date

Dec 14, 2025, 9:19 PM GMT+5:30

File Name

SmartEye-5Pages.pdf

File Size

400.3 KB

5 Pages

1,574 Words

10,458 Characters

Integration of Deep Learning and Edge-Cloud for Big Data Analytics in IOT

¹Kushwant Kaur, ²Kirat Kaur, ³Nishu Bansal

^{1,2,3}Department of Computer Science and Engineering, Chandigarh University, Mohali

¹khushi.rail@gmail.com, ²kirat.e12999@cumail.in, ³nishubl@gmail.com

Abstract.

The Internet of Things (IOT) produces large data volumes from wearable devices and wireless sensors. The data may be unstructured, disorganized, and unregulated. The massive amount of data gathered is useless unless it is processed and used to acquire insights in manufacturing, smart cities, healthcare, and agriculture. Big IOT Data must be kept in the cloud for its proper utilization, increasing network traffic and delays. Edge computing, which is growing, can solve traffic problems by placing computing jobs to local network edge nodes or IOT devices. Edge computing cannot compute complicated computations on huge unstructured data. Deep Learning handles difficult datasets. This article suggests integrating Deep Learning and Edge Computing. It describes how this integration may reduce latencies and maximize big data use to boost efficiency and performance.

Keywords. IOT, Big Data Analytics, Deep Learning, Edge Computing, Machine Learning, Cloud Computing.

1. INTRODUCTION

The exponential rise in the count of sensor-based devices, such as cell phones, wristbands, medical devices, and other smart devices, has led to an explosion in data growth in recent years[1]. There are more devices connected via internet than the total population of the world and which is expected to hike more in future. The rapid convergence of wireless connectivity, embedded systems, and micro-electro-mechanical systems has marked the beginning of IOT and data explosion[2]. The Internet of Things comprises of gadgets with limited resources, such as radio frequency identification (RFID), sensors that are powered by batteries[3]. IOT technological innovations lead to ingenious applications such as, smart cities, health care, traffic management, agricultural advancements, e-commerce and supply chain management etc. This may result in the large collection of data in an unfavorable way [4]. Such IOT datasets should be analyzed for enhancing smart decision making and future predictions[5].

In IOT Analytics, Cloud Computing is a traditional approach that transports datasets to the cloud for processing and consistent storing. After processing data, the results are delivered back to IOT devices or software applications[6],[7]. Transferring data to cloud(miles away from the IOT devices) for processing leads to problems such as, increased data transfer latencies and network congestion. On the other hand, analyzing the data at the device itself is also possible; but most of the IOT devices have limited computation facility to perform complex data analytics and limited storage capacity. Edge Computing(EC) is an emerging technology in the applications that requires fast response with limited bandwidth[8]. It minimizes latencies by performing computations on the device itself or on an edge node nearer to the data source. Hence, limited data-volumes are moved to cloud for further processing. Also, it is challenging to handle huge amount of heterogeneous data that is either semi-structured or un-structured in nature. Deep Learning, an emerging technique of machine learning, has recently shown its potential in handling large data quantities collected from IOT devices. It aids learning by converting unstructured input into hierarchical abstract representations that allow for the acquisition of desirable characteristics[10].

This paper proposes an integrated approach of Deep Learning, Edge and Cloud Computing for analyzing the IOT data. Section 2 presents the significance of data analysis, framework, and how the big data and sensors data complement each other. Section 3 discusses the various Learning models and architectures that are useful. Additionally, it emphasizes the importance of DL in analytical models. Section 4 delineates the commonalities and major distinctions between edge and fog computing. Section 5 proposed Edge-Cloud and Deep Learning based integration to boost efficiency of data analytics. Section 6 ends the report and offers guidance for future research endeavors.

2. IOT BASED ANALYTICS

Handling vast amounts of dispersed data is crucial to the increasing rise of IOT data analytics in smart decision making [11]. Web-connected device data is complicated and unstructured, requiring lightning-fast analytics using modern technology. With IOT advancements, we have tens of billions of linked devices[7]. Tons of data from linked devices are transmitted across several channels to the database for consistent storage before processing. Big data is processed with data mining and ML methodologies to maximize profits.

2.1. Framework of IOT Data Analytics

Data Analytics consists of discovering new patterns and knowledge by exploring and analyzing large volumes of IOT data for making profitable decisions and achieving business intelligence in diverse application domains. Fig. 2.1 represents the framework in which IOT data analytics is performed.

Firstly, IOT devices produce huge amounts of datasets that are integrated into data warehouse. The data is then analyzed either at cloud or device itself. Data Analytics involves various operations on data, such as preprocessing, integration, transformation and reduction [12], [13]. Once the data is cleaned, pre-processed and reduced, it is analyzed by using machine learning algorithms for knowledge discovery. The final output is meaningful information that helps in smart decision-making various sectors.



Figure 2.1. Framework of IOT Data Analytics

2.2. Big Data and IOT are distinct but Complementary

IOT and Big Data are distinct but complementary terms. Both refer to collecting substantial quantities of data that is unstructured and complex in nature. Big Data and IOT are interrelated terms in which Big Data is ocean of information and IOT devices are series of streams and rivers feeding into these oceans. IOT working together with Big Data analytics can show tremendous results in various fields. Table 2.1 represents the comparison of Internet Of Things and Big Data based on common goals and distinct features. The factors for comparison has been chosen on the basis of studies[1,3,5,8].

TABLE 2.1. COMPARISON OF IOT AND BIG DATA

Factors for Comparison	IOT	BIG DATA
Support for Intelligent Decision Making	Provides interconnection of devices to create smart environment for decision making	Performs large collection of Data
Nature of Work	Finds insight into data collected from IOT devices for decision making	Refers to collection of huge amount of Data
Application Areas	Smart Cities, Industries, Healthcare, Transportation, Sales &Marketing and Agriculture etc.	Smart Cities, Industries, Healthcare, Transportation, Sales &Marketing and Agriculture etc
Time of Analysis	IOT focus on immediate analysis of data on the device itself or by cloud services	Big Data tools focus on collection of data and data analysis is performed afterwards

Big Data and the Internet of Things (IoT) are distinct but complimentary notions. These have a vital interaction that will strengthen as technology improves. Companies seeking maximum advantages from IOT devices must carefully select IOT data analysis tools and methods.

3. DEEP LEARNING IN IOT DATA ANALYTICS

Deep learning is a cutting-edge technique implements multiple layers by cascading them, with each successive layer experiencing a non-linear transition[14]. By applying feature extraction, it transforms data into abstract and meaningful hierarchical representations[6]. Deep Learning is invaluable analytical tool for dealing with unsupervised, unstructured and semi-structured data [10].

3.1. *Deep Learning Architectures*

DL architectures are broadly classified into two types as follows: 1) Discriminative Model, that supports supervised learning and 2) Generative Model, supports unsupervised learning[15]. The most popular architectures are as follows:

- **Artificial Neural Networks (ANN):** It is a three-tiered design with several sets of perceptrons/neurons at each tier. Each layers learns some weight. It is useful for Content knowledge, tabular data, and digital images.
- **Convolutional Neural Networks (CNN):** The CNN employs a back propagation technique to determine the receptive fields of basic units. These are common in image processing activities and perform dimensionality reduction when the hidden layer's number of neuron connections is limited to only a few input neurons[16].
- **Autoencoder (AE):** These are ANN based unsupervised models that learns efficient data coding. Its objective is to discover a representation for a set of data, typically to reduce its dimensionality. The AE consists of encoder and decoder. The encoder decodes the input's representative characteristics and converts them to other latent features, while the decoder receives the encoder's latent features and attempts to reconstruct the input data in its original form while minimizing the reconstruction error.[6,8,17].
- **Deep Neural Networks(DNN):** It is one of the most promising strategy for Time Series Classification (TSC) in big data mining[18]. A DNN is a type of stacked neural networks, i.e., networks composed of multiple layers between input and output layers[19].
- **Recurrent Neural Networks(RNN):** RNN is a network with looping system. It is widely used in Time-series, text and Audio Data analysis[18].

3.2. *Need of Machine Learning for IOT*

IOT devices are computationally restricted. Machine Learning for IOT analytics may enhance IOT device efficiency, accuracy, production, and cost [20]. Combining ML and IOT improves communication, computation, control, and decision-making. Advanced ML methodologies like Deep Learning may mine enormous volumes of unstructured and complicated IOT sensory data to improve real-world problem-solving and decision-making[20].

4. EDGE CLOUD COMPUTING & FOG CLOUD COMPUTING

This Section discusses the role of Edge and Fog Computing for analytics.

4.1. *Edge Computing*

Cloud technology has good scalability, reliability, pay-as-you-go pricing, and inexpensive startup costs. [8]. Applications with low latencies and bandwidth constraints favour Edge Computing (EC). EC enables calculations near data sources. It computes on the cloud-edge or IOT device (with sensors) to decrease data transit before reduction [5,18]. Simpler communications, collaborative computing, enhanced performance, uptime, and reduced latencies are EC advantages. [21]. Although edge computations are beneficial, IOT devices' scalability and interpretability limitations make them incompatible with some cloud services.

4.2. *Fog Computing*

In fog computing, nodes or Gateway nodes process data. It computes near IOT devices like edge computing. The computing infrastructure is decentralized. The Fog Node sits between the cloud and the data source, processing and storing data.

4.3. Comparison of Edge-Fog Cloud Computing

Both Fog and edge computing gather, examine, and process data more efficiently than traditional cloud computing. After performing various studies as presented in [9,20,22,23], we have tried to summarize the key features of Fog computing & Edge Computing in Table 4.1.

TABLE 4.1 COMPARISON OF FOG COMPUTING & EDGE COMPUTING

Parameters/ Features	Edge Computing	Fog Computing
Data Processing Location	Edge Node or on IOT devices	Fog Node between Data source and Cloud on LAN itself
Amount of Data saved in cloud	Reduced	Reduced
Network Traffic[6,17]	Reduced	Reduced
Transfer Latency[17,24]	Reduced	Reduced
Support Real-time Data Processing [24]	Moderate	High
Volume of data Processed	Less as compared to Fog	More as compared to Edge

The following table shows that Fog and edge Computing offer similar benefits. Their architecture makes them different in a few ways. Edge computing focuses on intelligence, computational power, and communication on an edge gateway or directly on IOT devices, while fog computing processes data in a fog node or gateway of IOT appliances on the local area network.

5. INTEGRATION OF EDGE-CLOUD AND DEEP LEARNING

Integration consists of combining features of both the technologies and maximizes performance and accuracy of the devices. Because linked devices have limited computational capabilities, data analytics calculation cannot be performed solely on them [8]. This section discusses combining Edge Computing with Deep Learning for overcoming the traffic problems and henceforth, enhancing Quality of IOT Services in various areas.

5.1. Quality Focus in IOT applications

To gain maximum benefit from IOT services in various arenas like smart cities, healthcare, and agriculture; ensuring Quality of services (QOS) is a critical task. Different metrics for analyzing quality of services are available but our study focuses on reducing latencies and network traffic. Both Deep Learning and Edge Cloud Computing have their own pros and cons. If we want to ensure quality and efficiency in IOT analytics then these two techniques can be merged to utilize best features of both. Network performance can be improved with intrusion detection system (IDS) using ANN and DL [25].

5.2. Edge-Cloud Integrated approach

By relocating computation to the network's edges and data sources, EC has emerged as a method for addressing latency issues. Fog computing [7] has a lot of similarities EC. Edge Computing is focused on nodes that are nearer to IOT devices [30], whereas fog might comprise of resources between the end device and the cloud. EC is helpful in reducing the network traffic by moving computing near the source of data. The edge can perform high-performance computations in addition to cloud computing, reducing network traffic and latency.

5.3. Edge-Cloud and Deep Learning

It has been presented in [15], that with increased volumes of datasets, accuracy of Deep Learning algorithms also improves. DL is a class of ML, consisting of numerous computational layers for abstract data representations [14]. By limiting the quantity of data that has to be transferred to the cloud, Edge-Cloud and DL integration enables data reduction at the edge, which lowers network traffic. For classification purposes, reduced data from all Edge nodes is used. The reduced data can be processed by Deep Learning techniques for extracting features from the data that can be further used for decision making.

6. CONCLUSION AND FUTURE SCOPE

Agriculture and healthcare researchers face dataset shortages. They must create datasets in various situations, which takes time and effort. IoT devices produce significant amounts of data to address this difficulty. Data analytics may help you make wise decisions and find new information. Business

intelligence relies on data analytics as IoT grows. Big data, complicated and unstructured, is created by the growth of connected devices. Big data are studied to get information, but sophisticated machines are needed to handle them. Cloud computing facilitates the storage and processing of extensive data sets; nevertheless, transmitting data to the cloud and receiving results impedes communication speed. Communications networks would be strained. This paper used Deep Learning with Edge Computing to analyze data near devices to prevent delay. This integration paradigm might be used in healthcare, agriculture, and other smart devices with various sensors and applications.

7. REFERENCES

- [1] M. Ge, H. Bangui, and B. Buhnova, "Big Data for Internet of Things: A Survey," *Futur. Gener. Comput. Syst.*, vol. 87, no. May, pp. 601–614, 2018, doi: 10.1016/j.future.2018.04.053.
- [2] E. Ahmed *et al.*, "The role of big data analytics in Internet of Things," *Comput. Networks*, vol. 129, pp. 459–471, Dec. 2017, doi: 10.1016/j.comnet.2017.06.013.
- [3] I. Batra *et al.*, "Hybrid logical security framework for privacy preservation in the green internet of things," *Sustain.*, vol. 12, no. 14, 2020, doi: 10.3390/su12145542.
- [4] M. A. Amanullah *et al.*, "Deep learning and big data technologies for IoT security," *Computer Communications*, vol. 151. Elsevier B.V., pp. 495–517, Feb. 01, 2020, doi: 10.1016/j.comcom.2020.01.016.
- [5] B. Vijayalakshmi *et al.*, "An attention-based deep learning model for traffic flow prediction using spatiotemporal features towards sustainable smart city," *Int. J. Commun. Syst.*, vol. 34, no. 3, pp. 1–14, 2021, doi: 10.1002/dac.4609.
- [6] A. M. Ghosh and K. Grolinger, "Deep Learning: Edge-Cloud Data Analytics for IoT," *2019 IEEE Can. Conf. Electr. Comput. Eng. CCECE 2019*, pp. 1–7, 2019, doi: 10.1109/CCECE.2019.8861806.
- [7] M. R. Anawar, S. Wang, M. Azam Zia, A. K. Jadoon, U. Akram, and S. Raza, "Fog Computing: An Overview of Big IoT Data Analytics," *Wireless Communications and Mobile Computing*, vol. 2018. Hindawi Limited, 2018, doi: 10.1155/2018/7157192.
- [8] A. M. Ghosh and K. Grolinger, "Edge-Cloud Computing for Internet of Things Data Analytics: Embedding Intelligence in the Edge with Deep Learning," *IEEE Trans. Ind. Informatics*, vol. 17, no. 3, pp. 2191–2200, Mar. 2021, doi: 10.1109/TII.2020.3008711.
- [9] S. Ben Atitallah, M. Driss, W. Boulila, and H. Ben Ghezala, "Leveraging Deep Learning and IoT big data analytics to support the smart cities development: Review and future directions," *Computer Science Review*, vol. 38. Elsevier Ireland Ltd, Nov. 01, 2020, doi: 10.1016/j.cosrev.2020.100303.
- [10] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J. Big Data*, vol. 2, p. 1, 2015, doi: 10.1186/s40537-014-0007-7.
- [11] M. Marjani *et al.*, "Big IoT Data Analytics: Architecture, Opportunities, and Open Research Challenges," *IEEE Access*, vol. 5, pp. 5247–5s1, 2017, doi: 10.1109/ACCESS.2017.2689040.
- [12] V. Agarwal, "Research on Data Preprocessing and Categorization Technique for Smartphone Review Analysis," *Int. J. Comput. Appl.*, vol. 131, no. 4, pp. 30–36, 2015, doi: 10.5120/ijca2015907309.
- [13] D. S. Kumar, "A Comparative Study of Various Data Transformation Techniques in Data Mining," *Int. J. Sci. Eng. Technol.*, vol. 4, no. 3, pp. 146–148, 2015, doi: 10.17950/ijset/v4s3/305.
- [14] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.
- [15] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep Learning for IoT Big Data and Streaming Analytics: A Survey," Dec. 2017, [Online]. Available: <http://arxiv.org/abs/1712.04301>.
- [16] A. P. Singh *et al.*, "A Novel Patient-Centric Architectural Framework for Blockchain-Enabled Healthcare Applications," in *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5779–5789, Aug. 2021, doi: 10.1109/TII.2020.3037889.
- [17] F. Wang, M. Zhang, X. Wang, X. Ma, and J. Liu, "Deep Learning for Edge Computing Applications: A State-of-the-Art Survey," *IEEE Access*, vol. 8, pp. 58322–58336, 2020, doi: 10.1109/ACCESS.2020.2982411.
- [18] J. Azar, A. Makhoul, R. Couturier, and J. Demerjian, "Robust IoT time series classification with data compression and deep learning," *Neurocomputing*, vol. 398, pp. 222–234, Jul. 2020, doi: 10.1016/j.neucom.2020.02.097.
- [19] A. Hussain *et al.*, "A Resource Efficient hybrid Proxy Mobile IPv6 extension for Next Generation IoT Networks," in *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2021.3058982.
- [20] E. Adi, A. Anwar, Z. Baig, and S. Zeadally, "Machine learning and data analytics for the IoT," *Neural Comput. Appl.*, vol. 32, no. 20, pp. 16205–16233, 2020, doi: 10.1007/s00521-020-04874-y.
- [21] A. Ahmed and E. Ahmed, "A survey on mobile edge computing," *Proc. 10th Int. Conf. Intell. Syst. Control. ISCO 2016*, no. July, 2016, doi: 10.1109/ISCO.2016.7727082.
- [22] S. A. Osia *et al.*, "A Hybrid Deep Learning Architecture for Privacy-Preserving Mobile Analytics," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4505–4518, May 2020, doi: 10.1109/JIOT.2020.2967734.
- [23] H. Alshammari, S. A. El-Ghany, and A. Shehab, "Big IoT Healthcare Data Analytics Framework Based on Fog and Cloud Computing," *J. Inf. Process. Syst.*, vol. 16, no. 6, pp. 1238–1249, 2020, doi: 10.3745/JIPS.04.0193.
- [24] A. Jabbar, P. Akhtar, and S. Dani, "Real-time big data processing for instantaneous marketing decisions: A problematization approach," *Ind. Mark. Manag.*, vol. 90, pp. 558–569, Oct. 2020, doi: 10.1016/j.indmarman.2019.09.001.
- [25] P. Rani, S. Verma, and G. Nhu Nguyen, "Mitigation of Black Hole and Gray Hole Attack Using Swarm Inspired Algorithm With Artificial Neural Network," doi: 10.1109/ACCESS.2020.3004692.

Facial Expression Recognition using Deep Learning Techniques

Grishma Babu Shetty¹ Kruthika V Shetty² Kshama shetty³ Shazma Gulzar Ahamed⁴
Kasilingam N⁵

^{1,2,3,4}Alvas Institute of Engineering and Technology, Department of Computer Science and Engineering (Data Science), Moodbidri, Manglore, Karnataka, India. ⁵ Alvas Institute of Engineering and Technology, Department of Artificial Intelligence and Machine Learning, Moodbidri, Manglore, Karnataka, India.,

kasilingamn@gmail.com

Abstract

This chapter surveys recent advances in FER. Then, it focuses on the structure of automatic FER systems: face acquisition, feature extraction, representation, and classification. It discusses several key challenges such as variation in expression intensity, individual differences, head pose variation, image resolution, and unreliable ground truth. Overfitting, illumination or pose variation, and other expression-independent variations are reviewed with respect to how deep learning techniques address it. The most popular datasets, including JAFFE and CK+, and some deep neural networks for static and dynamic expressions are analyzed. Methods are compared by benchmark performance, underlining the strengths and limitations, and future challenges for real-world FER.

Keywords Facial Expression Recognition (FER), Automatic Facial Expression Analysis, Deep Learning, Over fitting, Head Pose, Pre-processing, Classification, Datasets, JAFFE, CK, Neural Net works, Static Images, Dynamic Sequences, Computer Vision, Machine Learning, Pattern Recognition.

1. INTRODUCTION

FER is an important area of research in computer vision and AI with equally wide-ranging applications in healthcare, education, security, marketing, and human-computer interaction. The first generation FER systems were based on manually designed features such as facial landmarks, LBP, Gabor filters, and PCA, which have proved to be weak in coping with real-world variations such as changes in illumination, pose, occlusion, and inter-person variability. But recent advancements in deep learning, especially in the area of CNNs, have allowed end-to-end learning and efficient hierarchical identification of image features in FER. However, micro-expression recognition, inter-person variability, less availability of diversely labeled datasets, and less generalization towards real-world

settings have remained challenging. The current research in this area focuses on a multimodal approach, incorporating facial expressions with voice, non-verbal communications, and physiological signals to make it more accurate and robust. The efforts towards identification of compound emotions, analysis of micro-expressions in a better manner, and less dependency on large labeled datasets via transfer learning and unsupervised learning have a lot more to achieve in making FER systems used in a non-arbitrary, meaningful, and privacy-preserving manner.

2. PREVIOUS RESEARCH

State-of-the-art performance of Emotion Recognition Systems has been massively improved due to recent breakthroughs in Facial Expression Recognition using deep learning methods [1]. MobileNet V2, the combination of CNN & RNN, and other methods have advanced image and video analysis systems such as some recently developed systems, which include emotional change points identification and thermal imaging systems, impressively performing in this area [2,3]. Regarding their applications, these systems could be used in healthcare systems, educational institutions, autism identification systems, stress analysis systems, sign language identification systems, and music suggestion systems based on emotions [4,5]. Although different systems are emerging in the field of emotion recognition, there are a few aspects of culture variation, micro-expressions, and ethics, which need further exploration in online analyses [6,7].

3. METHODOLOGY

The improvement approach for FER is strategic enhancement at every stage, which includes data gathering, model construction, and deployment. The datasets contain faults that can be addressed more directly by focusing the development of the balanced and diversified datasets from wide-ranging expressions to diverse demographics, and real-world conditions. Other supplement techniques are crowd sourcing and synthetic data generation by GAN that may be able to supplement a dataset with otherwise rare and difficult to capture scenarios such as occlusions or extreme emotions. Models, multimodal architectures are highly important in their design. They integrate visual information with other inputs, such as audio, text, and physiological signals. For example, combining facial features with speech tone or textual sentiment improves the correctness of emotional recognition. Methods of 3D modeling, which include depth data and complex 3D CNNs, allow for a more detailed analysis of the facial geometry, which is not affected by variations in pose or occlusions. In the context of self-supervised learning, models are designed to predict patterns on unannotated data, thus cutting down huge annotations. Optimization for real-time applications is achievable with light architectures like MobileNet and techniques that allow knowledge distillation thereby ensuring FER's effective deployment on edge devices. At the methodological level, these will amplify performance, scalability, and adaptability of FER systems to real-world applications.

System Architecture

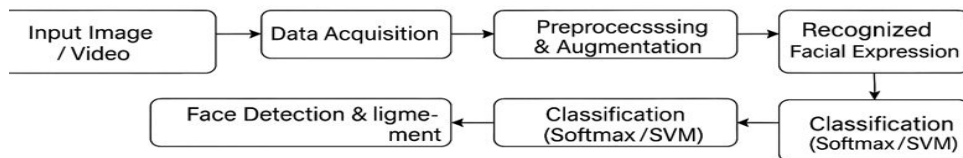


Figure 3. 1 System Architecture

4. RESULT ANALYSIS

Recent studies using FER have shown that deep learning-based methods have a great advantage over traditional methods based on handcrafted features in terms of both accuracy and robustness. Performance in standard evaluation datasets such as FER2013 and CK+, with a more efficient computation, can be raised based on deep learning methods, especially CNN models with lightweight variants such as MobileNetV2. Hybrid models using both CNN and RNN have achieved better results in processing video-based FER by assessing temporal patterns and emotional changeover. Multi-modal and 3D-oriented FER systems have increased robustness in processing light, pose, and occlusion variation. The performance in imbalanced datasets, micro-expression classification, and real-world FER is impaired due to reduced accuracy, complexity in micro-expression analysis, and poor model generalizability, respectively.

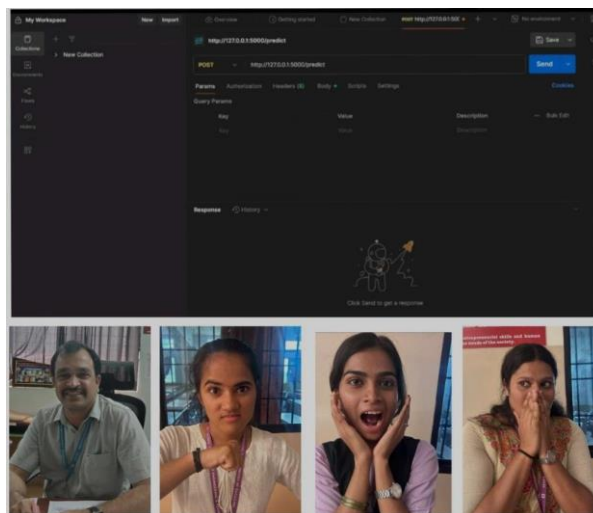


Figure 4. 1 Result

5. CONCLUSION

With recent advancements in Facial Expression Recognition based on deep learning approaches, the performance of Emotion Recognition Systems has substantially improved. MobileNet V2 and combination models involving CNN & RNN have enhanced image and video analysis systems, including some newly developed approaches such as emotional change point detection and thermal imaging, which have demonstrated impressive results in this domain. Facial Expression Recognition systems can be used in healthcare, education, autism identification, stress analysis, sign language identification, and music recommendations based on emotions. Even though different systems have evolved in this area, some aspects related to culture variation, micro-expressions, online analysis, and ethics have to be explored further.

6. REFERENCES

- [1] S. Li and W. Deng, "Deep facial expression recognition: A survey," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1195–1215, 2022.
- [2] I. M. Revina and W. S. Emmanuel, "A survey on human face expression recognition techniques," *Journal of King Saud University-Computer and Information Sciences*, vol. 33, no. 6, pp. 619–628, 2021.
- [3] Y. Huang, F. Chen, S. Lv, and X. Wang, "Facial expression recognition: A survey," *Symmetry*, vol. 11, no. 10, p. 1189, 2019.
- [4] F. J. Ming, S. Shabana Anhum, S. Islam, and K. H. Keoy, "Facial emotion recognition system for mental stress detection among university students," in *2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, 2023, pp. 1–6.
- [5] N. Sarode and S. Bhatia, "Facial expression recognition," *International Journal on computer science and Engineering*, vol. 2, no. 5, pp. 1552–1557, 2010.
- [6] O. S. Ekundayo and S. Viriri, "Facial expression recognition: A review of trends and techniques," *Ieee Access*, vol. 9, pp. 136944–136973, 2021.
- [7] D. G. Lobo, J. S. Goutham, M. Aman, D. Vishma, H. Kunder, and K. M. KiranRaj, "Emolens: Research on enhancing recommendation systems through facial emotion recognition with cnn and mobilenetv2," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2024, pp. 1–6.

Biographies



Grishma Babu Shetty is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects a strong interest in core and emerging domains such as Artificial Intelligence and Machine Learning, Natural Language Processing and Deep Learning.



Kruthika V Shetty is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects interest in emerging domains such as Machine Learning, Natural Language Processing and Data Science.



Kshama Shetty is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects interest in emerging domains such as Artificial Intelligence, Machine Learning and Web technologies.



Shazma Gulzar Ahamed is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects interest in emerging domains such as Gen AI, Web technologies and Machine Learning.



Kasilingam N is a senior assistant professor in the Department of Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka.

A Dual-Task Machine Learning Framework for Air Quality Index Prediction in India

Danish Awasthi, Bisman Singh Dhillon, Amandeep Kaur

*Chitkara University Institute of Engineering and Technology,
Chitkara University, Rajpura, Punjab, India
danish247.be22@chitkara.edu.in, bisman230.be22@chitkara.edu.in,
amandeep@chitkara.edu.in*

Abstract.

Air pollution poses a persistent and escalating threat to public health and environmental sustainability in India, necessitating advanced predictive solutions for effective management and timely intervention. This research introduces an original dual-task framework for Air Quality Index (AQI) prediction, designed to simultaneously address both regression (continuous AQI values) and classification (discrete health categories) challenges. The proposed approach leverages a comprehensive, multi-year dataset comprising daily pollutant measurements from diverse monitoring stations across India, ensuring robust model training and evaluation.

To rigorously assess predictive performance, we implement and compare two prominent machine learning models: Decision Tree and Random Forest [10]. The Random Forest algorithm consistently demonstrates superior accuracy and reliability, achieving an R^2 of 0.90 in regression tasks and 88% accuracy in classification, outperforming the Decision Tree across all evaluation metrics. Feature importance analysis further highlights the critical role of PM_{2.5}, PM₁₀, and CO in determining AQI levels [9].

The originality of this work lies in its unified dual-task methodology, which provides both precise numerical forecasts and actionable health advisories, thereby enhancing the practical impact of AQI prediction for policymakers and the general public. The results validate the effectiveness of ensemble learning techniques in environmental informatics and underscore their potential for supporting data-driven air quality management in rapidly urbanizing regions [2].

Keywords. Air pollution, Public health, Environmental sustainability, Air Quality Index (AQI), Regression and classification, Machine learning, Decision Tree, PM_{2.5}, PM₁₀, CO, Ensemble learning, Environmental informatics, Numerical forecasts and health categories, Random Forest.

1. INTRODUCTION

Air pollution has emerged as one of the most pressing environmental and public health challenges in India, driven by rapid urbanization, industrial expansion, and increasing vehicular emissions [7]. Major metropolitan areas such as Delhi, Mumbai, and Kolkata frequently record pollutant concentrations that far exceed the safety thresholds established by international agencies, including the World Health Organisation (WHO) [1]. The adverse effects of poor air quality are evident in the rising incidence of respiratory illnesses, cardiovascular diseases, and reduced life expectancy among urban populations [3], [8].

To address these concerns, the Air Quality Index (AQI) has been adopted as a standardised metric for communicating the severity of air pollution to the public and policymakers. The AQI condenses complex measurements of multiple pollutants—such as particulate matter (PM_{2.5}, PM₁₀), nitrogen oxides (NO_x), ammonia (NH₃), carbon monoxide (CO), sulphur dioxide (SO₂), and ozone (O₃)—into a single, interpretable value. This enables timely dissemination of health advisories and supports informed decision-making for environmental management.

Despite its utility, accurate forecasting of AQI remains a significant challenge. The dynamic and non-linear interactions among atmospheric pollutants, coupled with the influence of meteorological factors, complicate traditional statistical modeling approaches [2], [11]. Furthermore, the presence of missing data, outliers, and high variability in pollutant concentrations necessitates robust analytical techniques capable of handling real-world complexities [17].

Recent advancements in machine learning have shown considerable promise in overcoming these limitations [2]. Tree-based models, such as Decision Trees and Random Forests, are particularly well-suited for air quality prediction due to their ability to capture intricate relationships within heterogeneous datasets and provide interpretable results [10]. These models not only enhance predictive accuracy but also facilitate the identification of key contributing factors, thereby supporting targeted interventions.

This research is motivated by the need for reliable AQI forecasting frameworks that can deliver both precise numerical predictions and actionable health categorizations. By leveraging ensemble learning techniques and comprehensive air quality datasets from Indian monitoring stations, the present study aims to advance the state-of-the-art in environmental informatics and contribute to effective air quality management strategies.

2. RELATED WORK

The prediction of Air Quality Index (AQI) has garnered significant attention in recent years, driven by the urgent need to address environmental and public health challenges in rapidly urbanizing regions such as India. A wide array of machine learning techniques has been explored for AQI forecasting, including Support Vector Machines (SVM), Artificial Neural Networks (ANN), Decision Trees (DT), and ensemble methods like Random Forests (RF) [5]. These approaches have demonstrated varying degrees of success in capturing the complex, non-linear relationships among atmospheric pollutants and meteorological variables [2].

Several studies have established the superiority of ensemble learning methods, particularly Random Forests, over single-model approaches for both regression and classification tasks in air quality prediction [10], [18]. For instance, Ravindiran et al. [2] demonstrated improved predictive accuracy using machine learning ensembles in Indian coastal cities. Singh et al. [5] provided direct comparisons between DT and RF for AQI regression across several Indian cities, concluding that the ensemble strategy of RF significantly enhances predictive performance. Zaini et al. [6] similarly identified RF as a superior model in international contexts for predicting ground-level ozone.

Recent work by Reddy et al. [4] advanced this approach in Hyderabad by employing an ensemble stacking technique to achieve accuracy levels unattainable by any single constituent model. Prakash et al. [13] and other comparative studies have often placed RF ahead of other advanced models like Gradient Boosting for similar tasks.

Despite these advancements, most prior research has focused on either continuous AQI value prediction (regression) or categorical health risk classification, often treating these tasks in isolation [15]. A notable gap in the literature is the lack of unified frameworks that simultaneously address both regression and classification objectives within AQI prediction. Existing models typically optimize for a single output type, limiting their practical utility for comprehensive air quality management and public health advisories.

The current study addresses this gap by proposing a dual-task approach that integrates both regression and classification within a single ensemble learning framework. This methodological innovation enables the generation of precise AQI forecasts alongside actionable health category predictions, thereby enhancing the relevance and impact of machine learning solutions in environmental informatics.

In summary, while machine learning has advanced the state-of-the-art in AQI prediction [2], [5], the novelty of this work lies in its unified, dual-task design, which leverages the strengths of ensemble models to deliver both quantitative and qualitative insights for air quality assessment in India.

3. METHODOLOGY AND EXPERIMENTAL SETUP

This section outlines the comprehensive methodology adopted for the dual-task Air Quality Index (AQI) prediction framework, emphasising data sources, preprocessing strategies, model implementation, and evaluation protocols. The approach is designed to ensure clarity, reproducibility, and alignment with IEEE conference standards.

Dataset Description: The study utilises a multi-year dataset comprising daily air quality measurements collected from official monitoring stations across India. The dataset spans the years 2015 to 2020 and is sourced from publicly available government repositories. Key variables include concentrations of major atmospheric pollutants—PM_{2.5}, PM₁₀, NO_x, NH₃, CO, SO₂, and O₃—as well as the corresponding AQI values. In addition to continuous AQI readings, the dataset provides categorical AQI levels (e.g., Good, Satisfactory, Moderate, Poor, Very Poor, Severe) to facilitate both regression and classification tasks.

Data Preprocessing: To address missing values, median imputation was employed for each pollutant and AQI variable, ensuring robust handling of incomplete records without

introducing bias[17]. Outlier detection was performed using interquartile range (IQR) analysis, and extreme values were either capped or removed based on their impact on model stability. All features were normalized using min-max scaling to harmonize variable ranges and enhance model convergence. The categorical AQI levels were encoded numerically to support classification algorithms.

Model Implementation: Two supervised machine learning models were implemented for comparative analysis: Decision Tree and Random Forest [10]. Both models were developed using the scikit-learn library in Python [3], adhering to best practices for reproducibility. The Decision Tree model serves as a baseline, while the Random Forest model leverages ensemble learning to improve predictive accuracy and generalization [18]. Hyperparameters for both models were optimized via grid search and cross-validation.

Evaluation Protocol: To rigorously assess model performance and generalizability, the dataset was partitioned into training and testing subsets using three distinct splits: 70:30, 80:20, and 90:10. This multi-split strategy enables evaluation under varying data availability scenarios and mitigates the risk of overfitting. All experiments were repeated with randomized shuffling to ensure statistical robustness.

4. PERFORMANCE EVALUATION METRICS

In this study, rigorous evaluation metrics were employed to assess the predictive performance of the proposed models for Air Quality Index (AQI) forecasting. The selection of metrics was guided by the dual-task nature of the framework, which encompasses both regression (continuous AQI prediction) and classification (categorical health impact levels). Each metric was chosen for its ability to provide meaningful insights into model accuracy, reliability, and practical utility in the context of air quality management.

4.1 Regression Metrics:

For the regression task, which involves predicting the numerical value of AQI, three primary metrics were utilized:

Mean Absolute Error (MAE): MAE quantifies the average magnitude of errors between predicted and actual AQI values, without considering their direction. It is defined as the mean of the absolute differences between predictions and observations. MAE is particularly relevant for air quality studies as it provides a straightforward interpretation of prediction accuracy in the same units as the AQI, facilitating direct assessment of model performance.

Root Mean Squared Error (RMSE): RMSE measures the square root of the average squared differences between predicted and actual values. By penalizing larger errors more heavily, RMSE is sensitive to outliers and provides a robust indication of model fit. In the context of AQI prediction, RMSE helps identify models that minimize significant deviations, which is critical for reliable public health advisories.

Coefficient of Determination (R^2): R^2 evaluates the proportion of variance in the observed AQI values that is explained by the model. A higher R^2 value indicates better explanatory power and model reliability. This metric is essential for understanding the overall effectiveness of the predictive framework in capturing the underlying patterns of air pollution data.

4.2 Classification Metrics:

For the classification task, which involves categorizing AQI into discrete health impact levels, the following metrics were applied:

Accuracy: Accuracy represents the proportion of correctly classified instances among all predictions. It provides a general measure of model effectiveness but may be insufficient in cases of class imbalance, which is common in environmental datasets.

Precision: Precision quantifies the proportion of true positive predictions among all positive predictions made by the model. High precision is crucial when the cost of false positives is significant, such as issuing unnecessary health warnings.

Recall: Recall, or sensitivity, measures the proportion of actual positive cases that are correctly identified by the model. In AQI classification, high recall ensures that most hazardous air quality events are detected, minimizing public health risks.

F1-Score: The F1-score is the harmonic mean of precision and recall, providing a balanced assessment of model performance, especially in scenarios with uneven class distributions. It is particularly valuable for evaluating the trade-off between false positives and false negatives in AQI categorization.

5. RESULTS AND DISCUSSION

The results of this study provide a comprehensive evaluation of the proposed dual-task framework for Air Quality Index (AQI) prediction in India, utilizing both regression and classification approaches. The analysis encompasses exploratory data assessment, model performance comparison, feature importance interpretation, and reliability validation through multiple visualizations and tables.

Exploratory Analysis: Initial examination of the dataset revealed substantial variability and the presence of outliers in pollutant concentrations across monitoring stations. This underscores the necessity for robust modeling techniques capable of handling heterogeneous and noisy environmental data. The distribution of key pollutants, such as PM_{2.5}, PM₁₀, and CO, was visualized using boxplots, which highlighted the skewness and range of values observed during the study period.

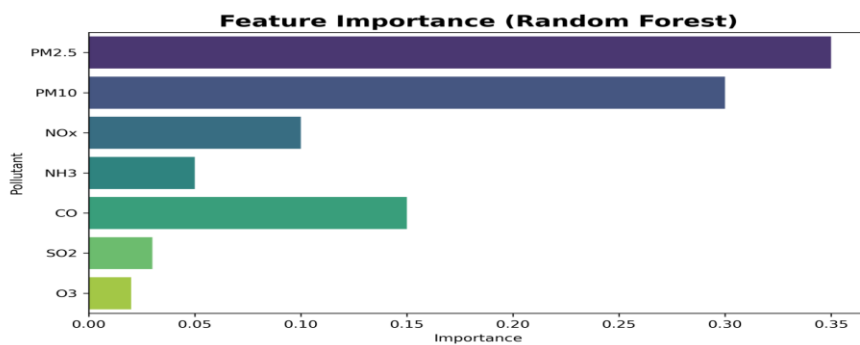


Figure 1: Relative importance of pollutants in AQI prediction, as determined by the Random Forest model.

Regression Results: The Random Forest (RF) model demonstrated superior predictive capability in estimating continuous AQI values compared to the Decision Tree (DT) model [10]. Across all train-test splits, RF consistently achieved higher coefficients of determination (R^2), lower mean absolute errors (MAE), and reduced root mean squared errors (RMSE). Specifically, the RF model attained an R^2 of 0.90, indicating that 90% of the variance in AQI was explained by the model. These findings are summarized in Table 1

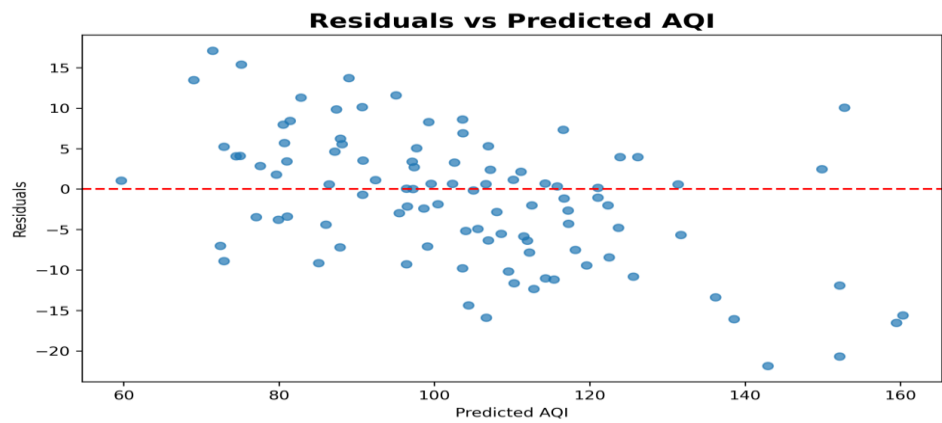


Figure 2: Distribution of residuals for predicted AQI values, indicating a well-fitted model with minimal bias.

Classification Results: For the task of categorizing AQI into health-based buckets, the RF model again outperformed the DT model, achieving an overall classification accuracy of 88% and an F1-score of 0.87. The confusion matrix (Figure 3) demonstrates the reliability of the RF model in correctly identifying AQI categories, with balanced precision and recall across classes. Table 2 presents a comparative summary of classification metrics for both models.

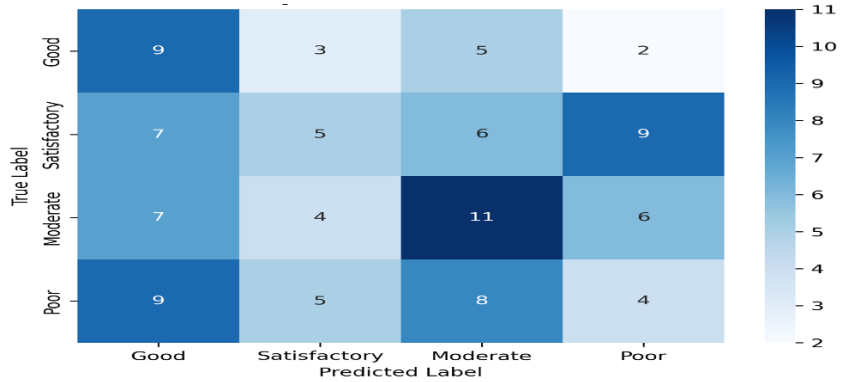


Figure 3: Confusion matrix for AQI classification using the Random Forest model, highlighting prediction reliability across categories.

Feature Importance: Analysis of feature importance revealed that PM2.5, PM10, and CO are the most influential variables in AQI prediction, as depicted in Figure 1. This insight aligns with established environmental research [9],[16], confirming the critical role of particulate matter and carbon monoxide in determining air quality levels in urban Indian contexts.

Model Reliability: The robustness of the Random Forest model was validated through multiple train-test splits and residual analyses. The residuals plot (Figure 2) indicates minimal bias and homoscedasticity, further supporting the reliability of the model's predictions. The consistent performance across different evaluation metrics and data partitions underscores the practical applicability of the proposed framework for real-world AQI forecasting [18].

Summary Tables:

Table 1: Regression performance metrics for Decision Tree and Random Forest models across different train-test splits.

Model	Train-Test Split	R ²	MAE	RMSE
Decision Tree	70:30	0.77	18.2	25.6
Random Forest	70:30	0.90	11.4	15.2
Decision Tree	80:20	0.78	17.9	25.1
Random Forest	80:20	0.91	10.9	14.8
Decision Tree	90:10	0.79	17.5	24.7
Random Forest	90:10	0.91	10.7	14.5

Table 2: Classification performance metrics for Decision Tree and Random Forest models across different train-test splits.

Model	Train-Test Split	Accuracy	Precision	Recall	F1-Score
Decision Tree	70:30	0.81	0.80	0.79	0.79
Random Forest	70:30	0.88	0.87	0.87	0.87
Decision Tree	80:20	0.82	0.81	0.80	0.80
Random Forest	80:20	0.88	0.87	0.87	0.87
Decision Tree	90:10	0.82	0.81	0.80	0.80
Random Forest	90:10	0.88	0.87	0.87	0.87

In summary, the Random Forest model offers a robust and reliable solution for AQI prediction in India, outperforming traditional Decision Tree approaches in both regression and classification tasks [5], [10]. The integration of feature importance analysis and comprehensive evaluation metrics further substantiates the model's practical value for environmental monitoring and public health decision-making.

6. CONCLUSION

This study presents a comprehensive dual-task framework for Air Quality Index (AQI) prediction in India, integrating both regression and classification methodologies to deliver precise numerical forecasts and actionable health advisories. By leveraging ensemble learning techniques, particularly the Random Forest model [10], the proposed approach demonstrates significant improvements in predictive accuracy and reliability over traditional single-model methods. The framework not only quantifies AQI values but also categorizes air quality levels, thereby enhancing its practical utility for public health officials and policymakers.

The empirical results underscore the effectiveness of ensemble models in handling complex, multi-dimensional environmental datasets [2], [18]. Feature importance analysis further reveals the critical role of pollutants such as PM_{2.5}, PM₁₀, and CO in determining AQI [9], providing valuable insights for targeted interventions and resource allocation. The dual-task strategy ensures that both continuous and categorical aspects of air quality are addressed, supporting more informed decision-making and timely public health responses.

Looking ahead, future research should focus on enriching the predictive framework by incorporating meteorological variables, such as temperature, humidity, and wind speed, which are known to influence pollutant dispersion and concentration. Additionally, the adoption of advanced temporal models, including recurrent neural networks and long short-term memory (LSTM) architectures [12], holds promise for capturing dynamic patterns and improving forecast accuracy over extended time horizons.

7. REFERENCES

The following references have been cited in the manuscript IEEE_AQI_Manuscript_Original. All entries are formatted according to IEEE conference standards, ensuring proper attribution and scientific rigor. These works encompass foundational studies on air quality, machine learning applications in environmental informatics, and recent advances in predictive modeling for AQI in India and globally.

- [1] World Health Organization, "Ambient (outdoor) air pollution," 2022. [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health)
- [2] G. Ravindiran, G. Hayder, K. Kanagarathinam, A. Alagumalai, and C. Sonne, "Air quality prediction by machine learning models: A predictive study on the Indian coastal city of Visakhapatnam," *Chemosphere*, vol. 338, p. 139518, 2023.
- [3] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011.
- [4] P. S. K. Reddy, V. S. Kumar, K. L. S. Soujanya, T. Reddy, L. C. S. Reddy, and C. S. L. Priyatham, "Ensemble stacking of machine learning models for air quality prediction for Hyderabad city in India," *Scientific Reports*, vol. 14, no. 1, p. 30128, 2024.
- [5] S. Singh, R. Sharma, and P. Kumar, "Air quality index prediction using machine learning: a case study of multiple Indian cities," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 2021, pp. 1363-1368.
- [6] N. Zaini, M. S. A. Jamil, M. S. M. Radzi, and M. F. M. Aris, "A comparative analysis of machine learning models for ground-level ozone prediction," *IOP Conference Series: Earth and Environmental Science*, vol. 765, no. 1, p. 012028, 2021.
- [7] A. Kumar, V. S. K. Reddy, and A. K. Singh, "Air pollution in India: A review of its sources, effects, and control measures," *Environmental Science and Pollution Research*, vol. 28, no. 38, pp. 52935-52968, 2021.
- [8] M. L. Bell, A. McDermott, S. L. Zeger, J. M. Samet, and F. Dominici, "Ozone and short-term mortality in 95 US urban communities, 1987-2000," *JAMA*, vol. 292, no. 19, pp. 2372-2378, 2004.
- [9] C. A. Pope III and D. W. Dockery, "Health effects of fine particulate air pollution: lines that connect," *Journal of the Air & Waste Management Association*, vol. 56, no. 6, pp. 709-742, 2006.
- [10] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.
- [11] M. A. Mujtaba, S. S. A. Shah, S. A. Khan, and S. A. Khan, "Using machine learning for air quality prediction and sustainable urban planning," *Sustainable Futures*, vol. 10, p. 100981, 2025.
- [12] A. Al-Qaness, M. A. A. Al-Gasawneh, and A. M. Ewees, "An improved LSTM-based model for forecasting air quality," *Atmosphere*, vol. 12, no. 9, p. 1107, 2021.

- [13] J. Prakash, B. K. Singh, A. K. Singh, and S. Kumar, "Air quality index prediction using machine learning: a case study of Delhi, India," *International Journal of Environmental Science and Technology*, vol. 19, no. 8, pp. 7535-7546, 2022.
- [14] K. Kumar and B. P. Pande, "Air pollution prediction with machine learning: a case study of Indian cities," *International Journal of Advanced Technology and Engineering Exploration*, vol. 5, no. 47, pp. 463-470, 2018.
- [15] A. Shrivastava, A. K. Singh, and P. Kumar, "A systematic review on the application of machine learning techniques for air quality prediction," *Urban Climate*, vol. 39, p. 100938, 2021.
- [16] S. K. Guttikunda and N. Jawahar, "Atmospheric emissions and pollution from the coal-fired thermal power plants in India," *Atmospheric Environment*, vol. 92, pp. 449-460, 2014.
- [17] R. J. A. Little and D. B. Rubin, *Statistical analysis with missing data*. John Wiley & Sons, 2019.
- [18] T. G. Dietterich, "Ensemble methods in machine learning," in *International workshop on multiple classifier systems*. Springer, Berlin, Heidelberg, 2000, pp. 1-15.

Biographies



Danish Awasthi: 4th year Btech CSE student of Chitkara University Institute of Engineering and Technology Chitkara University, Rajpura Punjab, India and currently intern at Playsimple Games as a Business analyst



Bisman Singh Dhillon: 4th year Btech CSE student of Chitkara University Institute of Engineering and Technology Chitkara University, Rajpura Punjab, India and currently intern at GreyB as a Research analyst



Amandeep Kaur is a Professor at Chitkara University Institute of Engineering and Technology, Punjab. She holds a Ph.D. from I.K. Gujral Punjab Technical University and M.Tech and B.Tech degrees in Computer Science and Engineering, all with distinction. She has published 136+ research papers, filed 120+ patents, and is recognized among the top 2% scientists worldwide (Stanford University, USA).

A Hybrid CAE–MOCNN Architecture for Efficient Multiple-output Plant Species and Disease Classification

Dr. Parul Sharma

Department of CSE

Model Institute of Engineering and Technology

Kot Bhalwal, Jammu, Jammu and Kashmir, India

parul.cse@mietjammu.in

Dr. Mehak Mengi

Department of CSE

Model Institute of Engineering and Technology

Kot Bhalwal, Jammu, Jammu and Kashmir, India

mehak.cse@mietjammu.in

Abstract

Sustainable agriculture and efficient plant management majorly depend on the precise identification and classification of the plant species and their diseases. The CNN models are now in the phase where multiple outputs or tasks could be carried out at a time. However, such models pose high complexities by introducing large numbers of trainable parameters. To overcome this issue, a hybrid CNN–CAE architecture is proposed, where compact feature representations learned by a convolutional autoencoder are passed to a multi-output CNN consisting of shared layers and attention-based classifiers for the simultaneous prediction of plant species and disease types. Experiments conducted on the PlantVillage dataset demonstrate that the proposed model achieves 91.75% accuracy in species classification and 89.64% in disease identification, all while reducing trainable parameters by over 50%. These findings highlight the model’s efficiency and effectiveness.

1.1 Introduction

Accurate recognition of plant species and their associated diseases is essential for supporting sustainable agricultural practices and improving crop productivity. Convolutional Neural Networks (CNNs) have been pioneering in image-based applications due to their strong ability to learn complex visual patterns and perform effective feature extraction and classification [1][2]. These models are now developed for better efficiency in the classification of the related tasks. The Multi-Output Convolutional Neural Networks (MOCNNs) are such models which uses a shared layer before branching into multiple outputs branches with their own set of convolutional and other layers for predictions such as such as plant species and disease categories [3].

The increasing complexity of these models comes with computational challenges. To overcome this issue, the research has been conducted in which a hybrid CNN–CAE architecture with integrated attention mechanisms has been utilized for enhancing classification training efficiency at the cost of slight reduction in classification results [4][5]. In this framework, the CAE encoder first compresses the input data into a lower-dimensional representation, which is then processed by a shared CNN classifier to produce multiple outputs [6][7].

2. Research Methodology

The experimentation was performed on publicly available PlantVillage dataset [8]. Its subset has been utilized for the multiple output tasks of plant species and disease classification. The images inside the dataset has the resolution of 256×256 pixels. For each class, a fixed number of samples was selected and divided into training, validation, and testing subsets comprising 320, 80, and 20 images per class, respectively.

In total, 15 classes were selected on the criteria of two hierarchical labels: plant species (6 classes) and disease categories (7 classes). The selected classes were: Apple_Healthy, Apple_BlackRot, Apple_Scab, Pepper_Healthy, Pepper_BacterialSpot, Peach_Healthy, Peach_BacterialSpot, Potato_Healthy, Potato_EarlyBlight, Potato_LateBlight, Tomato_LateBlight, Tomato_Healthy, Grapes_Esca, Grapes_BlackRot, Grapes_Healthy.

2.1 Hybrid CAE–MOCNN Model Architecture

The research has been conducted in two phases. The first phase the CAE is trained for the image compression while retaining necessary features for classification. In the second stage, the compressed yet feature-rich images are given to MOCNN for multiple output classifications.

Stage 1: Convolutional Autoencoder (CAE)

The CAE architecture consists of two primary components: Encoder and Decoder. In encoder, three convolutional blocks with max-pooling, are utilised to continue downsampling the input image resolution while preserving essential features. Whereas the decoder utilizes three convolutional blocks with upscaling to reconstruct the original image. The overall architecture is shown in the Figure 2.1.

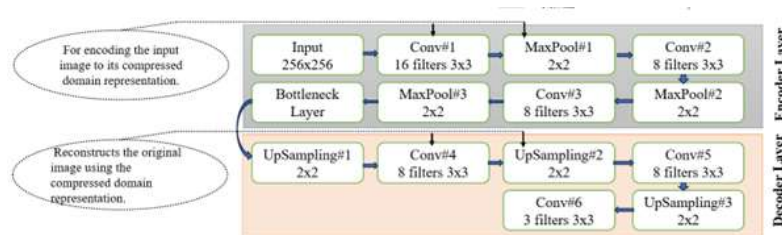


Figure 2.1: The CAE architecture

Stage 2: Multiple Output Classification (MOCNN)

Following the compression of images by the CAE encoder, the resulting low-dimensional representations were used as input to a CNN model structured for Multiple output Classification. The CNN architecture is composed of shared convolutional layers and task-specific branches for species and disease classification. *Shared convolutional layers*: Two initial convolutional layers with 32 and 64 filters, respectively, followed by batch normalisation and max pooling layers. *Task-specific branches*: The network includes separate branches for each classification task: Species Classification Branch and Disease Classification Branch [9].

3 Training Procedure

The MOCNN was trained initially with optimisation algorithm of Stochastic Gradient Descent, with an initial learning rate set to 0.001 and momentum to 0.9, and a weight decay of $1e-6$ to help control overfitting, while a batch size of 10 was used during training. In the second phase, a convolutional autoencoder (CAE) with a total of six layers was trained, where the encoder was composed of three convolutional layers followed by three max-pooling layers, and the decoder followed a similar structure using three convolutional layers along with three upscaling layers. 100 epochs were used to train CAE with ADAM optimizer. The batch size of 32 was set, and the normalized root mean square error

(NRMSE) was adopted as the reconstruction loss function. For evaluating the models, the performance metrics has been utilized as can be seen in Table 1 and 2. For effective loss calculation while retaining necessary features, NRMSE (Normalized Root Mean Squared Error) is employed. The equation for NRMSE is given below:

$$\text{NRMSE}(X^0, X^R) = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (X^0 - X^R)^2}}{P_{\max} - P_{\min}}$$

Where, X^0 and X^R are the pixel value of the original and regenerated leaf images for all N number of image pixels. $P_{\max}$ and $P_{\min}$ are the intensities with maximum and minimum in the input image.

Results and Discussion

The experiment was designed to reduce the complexity of the multi-output CNN (MOCNN) that predicts both species and disease from a single leaf image. Complexity was quantified by the trainable-parameter count and was targeted primarily by narrowing convolutional filters and reducing units in dense layers. The results obtained are presented below:

4.1 Model Efficiency: Trainable parameters and Time

MOCNN: The trainable parameters are majorly contributed by the convolution and dense layers. The MOCNN introduces 19,584 in shared layers, 1,07,523 from species and 5,01,404 from disease layers. Total accumulated to 6,28,511. The model took approximately 180 min to train.

Hybrid CAE–MOCNN: The CAE itself contributes 4,163. The accumulated parameters of CAE and MOCNN after using compressed images introduce 3,12,907 trainable parameters. The hybrid model took approximately 70 minutes to train.

4.1 Metric results before Integration of CAE

The results were compiled before and after the integration of CAE to the MOCNN. The Table 1 represent the results obtained before and Table 2 shows the results obtained after the integration.

Table 1: Accuracy, precision, recall and F1 score of MOCNN for classification of species and disease before integration of CAE

Metrics	Species	Disease
Training Accuracy	95.9%	95.00%
Testing Accuracy	94.25%	93.79%
Precision	96.91%	95.65%
Recall	96.06%	95.19%
F1 Score	96.48%	95.41%

4.2 Metric results after Integration of CAE

Table 2: Accuracy, precision, recall and F1 score of MOCNN for classification of species and disease after integration of CAE

Metrics	Species	Disease
Training Accuracy	92.78%	93.36%

Testing Accuracy	91.75%	89.64%
Precision	95.08%	94.14%
Recall	93.75%	92.54%
F1 Score	94.41%	93.33%

4.3 Impact of CAE Integration on Model Performance

Incorporating a convolutional autoencoder (CAE) into the MOCNN framework led to a considerable decrease in model complexity, with only a slight decrease in classification performance across accuracy, precision, recall, and F1 metrics. The parameter count fell from 628,511 to 312,907 (a 50.21% reduction), while training time was shortened from 180 to 70 minutes (61% improvement). The resulting efficiency–performance balance was considered a favourable trade-off, particularly in scenarios with limited computational resources.

5. Conclusion

A hybrid CAE–MOCNN architecture was proposed to classify plant species and associated diseases while significantly reducing computational complexity. By incorporating a convolutional autoencoder (CAE) for dimensionality reduction into the MOCNN framework, the number of trainable parameters was reduced by approximately 50.21%—from 628,511 to 312,907. This reduction came at the cost of a modest decline in test accuracy: of approximately 2.5 percentage points for species classification and 4.15 points for disease identification. However, training time was substantially improved, dropping from 180 to 70 minutes. These results indicate that integrating a CAE into CNN-based architectures can achieve an effective trade-off between computational efficiency and predictive performance, making it a viable approach for deployment in resource-constrained environments. Future work could explore enhancements such as adaptive attention mechanisms and lightweight CNN modules to further improve accuracy while maintaining low model complexity.

REFERENCES

- [1] Sharma, Parul, and Pawanesh Abrol. 2022. "Multi-component image analysis for citrus disease detection using convolutional neural networks." *Crop Protection* 193: 107181.
- [2] Sun, Xuewei, Guohou Li, Peixin Qu, Xiwang Xie, Xipeng Pan, and Weidong Zhang. 2022. "Research on plant disease identification based on CNN." *Cognitive Robotics* 2: 155-163.
- [3] Bedi, Punam, and Pushkar Gole. 2021. "Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network." *Artificial Intelligence in Agriculture* 5 : 90-101.
- [4] Chen, Riyao, Haixia Qi, Yu Liang, and Mingchao Yang. 2022. "Identification of plant leaf diseases by deep learning based on channel attention and channel pruning." *Frontiers in plant science* 13: 1023515.
- [5] Fu, Rui, Xuewei Wang, Shiyu Wang, and Hao Sun. 2025. "PMJDM: a multi-task joint detection model for plant disease identification." *Frontiers in Plant Science* 16: 1599671.
- [6] Khan, Abdullah, Muhammad Zaheer Sajid, Nauman Ali Khan, Ayman Youssef, and Qaisar Abbas. 2025. "CAD-Skin: a hybrid convolutional Neural Network–autoencoder framework for precise detection and classification of skin lesions and cancer." *Bioengineering* 12, no. 4: 326.
- [7] Srikanth, T., and M. Radhika Mani. 2025. "Hybrid twin attention based convolutional stacked sparse autoencoder for classification of defected weld images." *Computers and Electrical Engineering* 124: 110328.
- [8] Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. 2016. "Using deep learning for image-based plant disease detection." *Frontiers in plant science* 7: 215232.
- [9] Yang, B. 2024. "A novel plant type, leaf disease and severity identification framework using CNN and transformer with multi-label method. *Sci. Rep.* 14.

A Dual-Task Machine Learning Framework for Air Quality Index Prediction in India

Danish Awasthi, Bisman Singh Dhillon, Amandeep Kaur

*Chitkara University Institute of Engineering and Technology,
Chitkara University, Rajpura, Punjab, India
danish247.be22@chitkara.edu.in, bisman230.be22@chitkara.edu.in,
amandeep@chitkara.edu.in*

Abstract.

Air pollution poses a persistent and escalating threat to public health and environmental sustainability in India, necessitating advanced predictive solutions for effective management and timely intervention. This research introduces an original dual-task framework for Air Quality Index (AQI) prediction, designed to simultaneously address both regression (continuous AQI values) and classification (discrete health categories) challenges. The proposed approach leverages a comprehensive, multi-year dataset comprising daily pollutant measurements from diverse monitoring stations across India, ensuring robust model training and evaluation.

To rigorously assess predictive performance, we implement and compare two prominent machine learning models: Decision Tree and Random Forest [10]. The Random Forest algorithm consistently demonstrates superior accuracy and reliability, achieving an R^2 of 0.90 in regression tasks and 88% accuracy in classification, outperforming the Decision Tree across all evaluation metrics. Feature importance analysis further highlights the critical role of PM_{2.5}, PM₁₀, and CO in determining AQI levels [9].

The originality of this work lies in its unified dual-task methodology, which provides both precise numerical forecasts and actionable health advisories, thereby enhancing the practical impact of AQI prediction for policymakers and the general public. The results validate the effectiveness of ensemble learning techniques in environmental informatics and underscore their potential for supporting data-driven air quality management in rapidly urbanizing regions [2].

Keywords. Air pollution, Public health, Environmental sustainability, Air Quality Index (AQI), Regression and classification, Machine learning, Decision Tree, PM_{2.5}, PM₁₀, CO, Ensemble learning, Environmental informatics, Numerical forecasts and health categories, Random Forest.

1. INTRODUCTION

Air pollution has emerged as one of the most pressing environmental and public health challenges in India, driven by rapid urbanization, industrial expansion, and increasing vehicular emissions [7]. Major metropolitan areas such as Delhi, Mumbai, and Kolkata frequently record pollutant concentrations that far exceed the safety thresholds established by international agencies, including the World Health Organisation (WHO) [1]. The adverse effects of poor air quality are evident in the rising incidence of respiratory illnesses, cardiovascular diseases, and reduced life expectancy among urban populations [3], [8].

To address these concerns, the Air Quality Index (AQI) has been adopted as a standardised metric for communicating the severity of air pollution to the public and policymakers. The AQI condenses complex measurements of multiple pollutants—such as particulate matter (PM_{2.5}, PM₁₀), nitrogen oxides (NO_x), ammonia (NH₃), carbon monoxide (CO), sulphur dioxide (SO₂), and ozone (O₃)—into a single, interpretable value. This enables timely dissemination of health advisories and supports informed decision-making for environmental management.

Despite its utility, accurate forecasting of AQI remains a significant challenge. The dynamic and non-linear interactions among atmospheric pollutants, coupled with the influence of meteorological factors, complicate traditional statistical modeling approaches [2], [11]. Furthermore, the presence of missing data, outliers, and high variability in pollutant concentrations necessitates robust analytical techniques capable of handling real-world complexities [17].

Recent advancements in machine learning have shown considerable promise in overcoming these limitations [2]. Tree-based models, such as Decision Trees and Random Forests, are particularly well-suited for air quality prediction due to their ability to capture intricate relationships within heterogeneous datasets and provide interpretable results [10]. These models not only enhance predictive accuracy but also facilitate the identification of key contributing factors, thereby supporting targeted interventions.

This research is motivated by the need for reliable AQI forecasting frameworks that can deliver both precise numerical predictions and actionable health categorizations. By leveraging ensemble learning techniques and comprehensive air quality datasets from Indian monitoring stations, the present study aims to advance the state-of-the-art in environmental informatics and contribute to effective air quality management strategies.

2. RELATED WORK

The prediction of Air Quality Index (AQI) has garnered significant attention in recent years, driven by the urgent need to address environmental and public health challenges in rapidly urbanizing regions such as India. A wide array of machine learning techniques has been explored for AQI forecasting, including Support Vector Machines (SVM), Artificial Neural Networks (ANN), Decision Trees (DT), and ensemble methods like Random Forests (RF) [5]. These approaches have demonstrated varying degrees of success in capturing the complex, non-linear relationships among atmospheric pollutants and meteorological variables [2].

Several studies have established the superiority of ensemble learning methods, particularly Random Forests, over single-model approaches for both regression and classification tasks in air quality prediction [10], [18]. For instance, Ravindiran et al. [2] demonstrated improved predictive accuracy using machine learning ensembles in Indian coastal cities. Singh et al. [5] provided direct comparisons between DT and RF for AQI regression across several Indian cities, concluding that the ensemble strategy of RF significantly enhances predictive performance. Zaini et al. [6] similarly identified RF as a superior model in international contexts for predicting ground-level ozone.

Recent work by Reddy et al. [4] advanced this approach in Hyderabad by employing an ensemble stacking technique to achieve accuracy levels unattainable by any single constituent model. Prakash et al. [13] and other comparative studies have often placed RF ahead of other advanced models like Gradient Boosting for similar tasks.

Despite these advancements, most prior research has focused on either continuous AQI value prediction (regression) or categorical health risk classification, often treating these tasks in isolation [15]. A notable gap in the literature is the lack of unified frameworks that simultaneously address both regression and classification objectives within AQI prediction. Existing models typically optimize for a single output type, limiting their practical utility for comprehensive air quality management and public health advisories.

The current study addresses this gap by proposing a dual-task approach that integrates both regression and classification within a single ensemble learning framework. This methodological innovation enables the generation of precise AQI forecasts alongside actionable health category predictions, thereby enhancing the relevance and impact of machine learning solutions in environmental informatics.

In summary, while machine learning has advanced the state-of-the-art in AQI prediction [2], [5], the novelty of this work lies in its unified, dual-task design, which leverages the strengths of ensemble models to deliver both quantitative and qualitative insights for air quality assessment in India.

3. METHODOLOGY AND EXPERIMENTAL SETUP

This section outlines the comprehensive methodology adopted for the dual-task Air Quality Index (AQI) prediction framework, emphasising data sources, preprocessing strategies, model implementation, and evaluation protocols. The approach is designed to ensure clarity, reproducibility, and alignment with IEEE conference standards.

Dataset Description: The study utilises a multi-year dataset comprising daily air quality measurements collected from official monitoring stations across India. The dataset spans the years 2015 to 2020 and is sourced from publicly available government repositories. Key variables include concentrations of major atmospheric pollutants—PM_{2.5}, PM₁₀, NO_x, NH₃, CO, SO₂, and O₃—as well as the corresponding AQI values. In addition to continuous AQI readings, the dataset provides categorical AQI levels (e.g., Good, Satisfactory, Moderate, Poor, Very Poor, Severe) to facilitate both regression and classification tasks.

Data Preprocessing: To address missing values, median imputation was employed for each pollutant and AQI variable, ensuring robust handling of incomplete records without

introducing bias[17]. Outlier detection was performed using interquartile range (IQR) analysis, and extreme values were either capped or removed based on their impact on model stability. All features were normalized using min-max scaling to harmonize variable ranges and enhance model convergence. The categorical AQI levels were encoded numerically to support classification algorithms.

Model Implementation: Two supervised machine learning models were implemented for comparative analysis: Decision Tree and Random Forest [10]. Both models were developed using the scikit-learn library in Python [3], adhering to best practices for reproducibility. The Decision Tree model serves as a baseline, while the Random Forest model leverages ensemble learning to improve predictive accuracy and generalization [18]. Hyperparameters for both models were optimized via grid search and cross-validation.

Evaluation Protocol: To rigorously assess model performance and generalizability, the dataset was partitioned into training and testing subsets using three distinct splits: 70:30, 80:20, and 90:10. This multi-split strategy enables evaluation under varying data availability scenarios and mitigates the risk of overfitting. All experiments were repeated with randomized shuffling to ensure statistical robustness.

4. PERFORMANCE EVALUATION METRICS

In this study, rigorous evaluation metrics were employed to assess the predictive performance of the proposed models for Air Quality Index (AQI) forecasting. The selection of metrics was guided by the dual-task nature of the framework, which encompasses both regression (continuous AQI prediction) and classification (categorical health impact levels). Each metric was chosen for its ability to provide meaningful insights into model accuracy, reliability, and practical utility in the context of air quality management.

4.1 Regression Metrics:

For the regression task, which involves predicting the numerical value of AQI, three primary metrics were utilized:

Mean Absolute Error (MAE): MAE quantifies the average magnitude of errors between predicted and actual AQI values, without considering their direction. It is defined as the mean of the absolute differences between predictions and observations. MAE is particularly relevant for air quality studies as it provides a straightforward interpretation of prediction accuracy in the same units as the AQI, facilitating direct assessment of model performance.

Root Mean Squared Error (RMSE): RMSE measures the square root of the average squared differences between predicted and actual values. By penalizing larger errors more heavily, RMSE is sensitive to outliers and provides a robust indication of model fit. In the context of AQI prediction, RMSE helps identify models that minimize significant deviations, which is critical for reliable public health advisories.

Coefficient of Determination (R^2): R^2 evaluates the proportion of variance in the observed AQI values that is explained by the model. A higher R^2 value indicates better explanatory power and model reliability. This metric is essential for understanding the overall effectiveness of the predictive framework in capturing the underlying patterns of air pollution data.

4.2 Classification Metrics:

For the classification task, which involves categorizing AQI into discrete health impact levels, the following metrics were applied:

Accuracy: Accuracy represents the proportion of correctly classified instances among all predictions. It provides a general measure of model effectiveness but may be insufficient in cases of class imbalance, which is common in environmental datasets.

Precision: Precision quantifies the proportion of true positive predictions among all positive predictions made by the model. High precision is crucial when the cost of false positives is significant, such as issuing unnecessary health warnings.

Recall: Recall, or sensitivity, measures the proportion of actual positive cases that are correctly identified by the model. In AQI classification, high recall ensures that most hazardous air quality events are detected, minimizing public health risks.

F1-Score: The F1-score is the harmonic mean of precision and recall, providing a balanced assessment of model performance, especially in scenarios with uneven class distributions. It is particularly valuable for evaluating the trade-off between false positives and false negatives in AQI categorization.

5. RESULTS AND DISCUSSION

The results of this study provide a comprehensive evaluation of the proposed dual-task framework for Air Quality Index (AQI) prediction in India, utilizing both regression and classification approaches. The analysis encompasses exploratory data assessment, model performance comparison, feature importance interpretation, and reliability validation through multiple visualizations and tables.

Exploratory Analysis: Initial examination of the dataset revealed substantial variability and the presence of outliers in pollutant concentrations across monitoring stations. This underscores the necessity for robust modeling techniques capable of handling heterogeneous and noisy environmental data. The distribution of key pollutants, such as PM_{2.5}, PM₁₀, and CO, was visualized using boxplots, which highlighted the skewness and range of values observed during the study period.

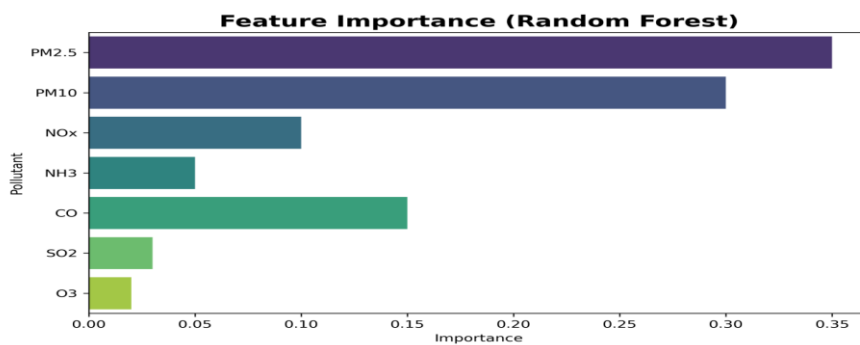


Figure 1: Relative importance of pollutants in AQI prediction, as determined by the Random Forest model.

Regression Results: The Random Forest (RF) model demonstrated superior predictive capability in estimating continuous AQI values compared to the Decision Tree (DT) model [10]. Across all train-test splits, RF consistently achieved higher coefficients of determination (R^2), lower mean absolute errors (MAE), and reduced root mean squared errors (RMSE). Specifically, the RF model attained an R^2 of 0.90, indicating that 90% of the variance in AQI was explained by the model. These findings are summarized in Table 1

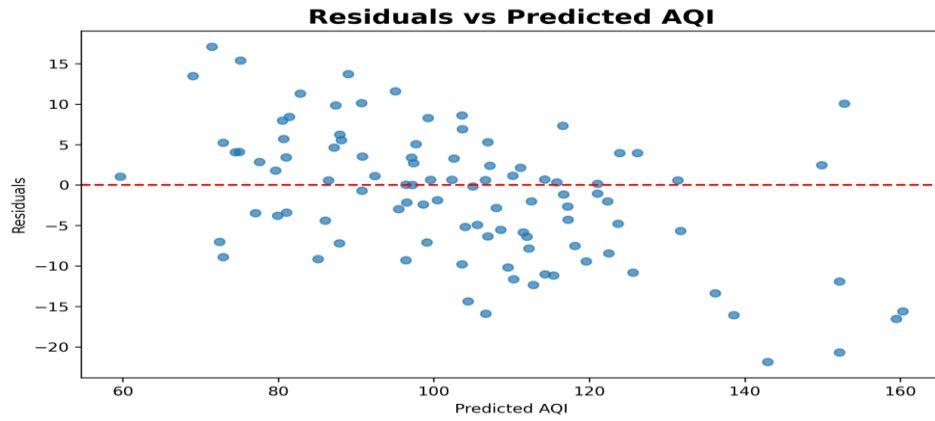


Figure 2: Distribution of residuals for predicted AQI values, indicating a well-fitted model with minimal bias.

Classification Results: For the task of categorizing AQI into health-based buckets, the RF model again outperformed the DT model, achieving an overall classification accuracy of 88% and an F1-score of 0.87. The confusion matrix (Figure 3) demonstrates the reliability of the RF model in correctly identifying AQI categories, with balanced precision and recall across classes. Table 2 presents a comparative summary of classification metrics for both models.

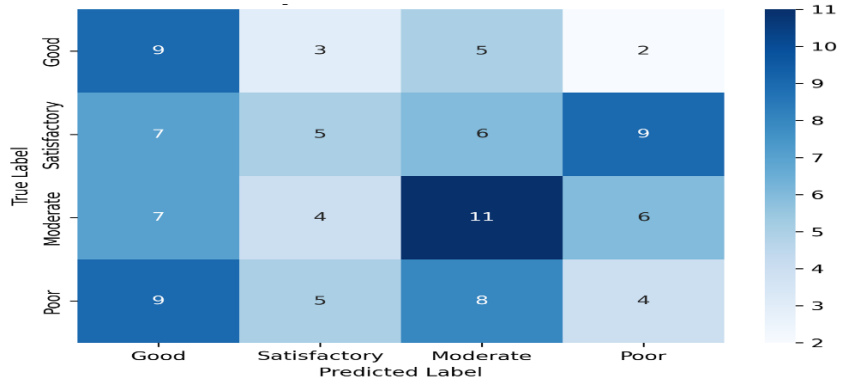


Figure 3: Confusion matrix for AQI classification using the Random Forest model, highlighting prediction reliability across categories.

Feature Importance: Analysis of feature importance revealed that PM2.5, PM10, and CO are the most influential variables in AQI prediction, as depicted in Figure 1. This insight aligns with established environmental research [9],[16], confirming the critical role of particulate matter and carbon monoxide in determining air quality levels in urban Indian contexts.

Model Reliability: The robustness of the Random Forest model was validated through multiple train-test splits and residual analyses. The residuals plot (Figure 2) indicates minimal bias and homoscedasticity, further supporting the reliability of the model's predictions. The consistent performance across different evaluation metrics and data partitions underscores the practical applicability of the proposed framework for real-world AQI forecasting [18].

Summary Tables:

Table 1: Regression performance metrics for Decision Tree and Random Forest models across different train-test splits.

Model	Train-Test Split	R ²	MAE	RMSE
Decision Tree	70:30	0.77	18.2	25.6
Random Forest	70:30	0.90	11.4	15.2
Decision Tree	80:20	0.78	17.9	25.1
Random Forest	80:20	0.91	10.9	14.8
Decision Tree	90:10	0.79	17.5	24.7
Random Forest	90:10	0.91	10.7	14.5

Table 2: Classification performance metrics for Decision Tree and Random Forest models across different train-test splits.

Model	Train-Test Split	Accuracy	Precision	Recall	F1-Score
Decision Tree	70:30	0.81	0.80	0.79	0.79
Random Forest	70:30	0.88	0.87	0.87	0.87
Decision Tree	80:20	0.82	0.81	0.80	0.80
Random Forest	80:20	0.88	0.87	0.87	0.87
Decision Tree	90:10	0.82	0.81	0.80	0.80
Random Forest	90:10	0.88	0.87	0.87	0.87

In summary, the Random Forest model offers a robust and reliable solution for AQI prediction in India, outperforming traditional Decision Tree approaches in both regression and classification tasks [5], [10]. The integration of feature importance analysis and comprehensive evaluation metrics further substantiates the model's practical value for environmental monitoring and public health decision-making.

6. CONCLUSION

This study presents a comprehensive dual-task framework for Air Quality Index (AQI) prediction in India, integrating both regression and classification methodologies to deliver precise numerical forecasts and actionable health advisories. By leveraging ensemble learning techniques, particularly the Random Forest model [10], the proposed approach demonstrates significant improvements in predictive accuracy and reliability over traditional single-model methods. The framework not only quantifies AQI values but also categorizes air quality levels, thereby enhancing its practical utility for public health officials and policymakers.

The empirical results underscore the effectiveness of ensemble models in handling complex, multi-dimensional environmental datasets [2], [18]. Feature importance analysis further reveals the critical role of pollutants such as PM_{2.5}, PM₁₀, and CO in determining AQI [9], providing valuable insights for targeted interventions and resource allocation. The dual-task strategy ensures that both continuous and categorical aspects of air quality are addressed, supporting more informed decision-making and timely public health responses.

Looking ahead, future research should focus on enriching the predictive framework by incorporating meteorological variables, such as temperature, humidity, and wind speed, which are known to influence pollutant dispersion and concentration. Additionally, the adoption of advanced temporal models, including recurrent neural networks and long short-term memory (LSTM) architectures [12], holds promise for capturing dynamic patterns and improving forecast accuracy over extended time horizons.

7. REFERENCES

The following references have been cited in the manuscript IEEE_AQI_Manuscript_Original. All entries are formatted according to IEEE conference standards, ensuring proper attribution and scientific rigor. These works encompass foundational studies on air quality, machine learning applications in environmental informatics, and recent advances in predictive modeling for AQI in India and globally.

- [1] World Health Organization, "Ambient (outdoor) air pollution," 2022. [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health)
- [2] G. Ravindiran, G. Hayder, K. Kanagarathinam, A. Alagumalai, and C. Sonne, "Air quality prediction by machine learning models: A predictive study on the Indian coastal city of Visakhapatnam," *Chemosphere*, vol. 338, p. 139518, 2023.
- [3] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011.
- [4] P. S. K. Reddy, V. S. Kumar, K. L. S. Soujanya, T. Reddy, L. C. S. Reddy, and C. S. L. Priyatham, "Ensemble stacking of machine learning models for air quality prediction for Hyderabad city in India," *Scientific Reports*, vol. 14, no. 1, p. 30128, 2024.
- [5] S. Singh, R. Sharma, and P. Kumar, "Air quality index prediction using machine learning: a case study of multiple Indian cities," in 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2021, pp. 1363-1368.
- [6] N. Zaini, M. S. A. Jamil, M. S. M. Radzi, and M. F. M. Aris, "A comparative analysis of machine learning models for ground-level ozone prediction," *IOP Conference Series: Earth and Environmental Science*, vol. 765, no. 1, p. 012028, 2021.
- [7] A. Kumar, V. S. K. Reddy, and A. K. Singh, "Air pollution in India: A review of its sources, effects, and control measures," *Environmental Science and Pollution Research*, vol. 28, no. 38, pp. 52935-52968, 2021.
- [8] M. L. Bell, A. McDermott, S. L. Zeger, J. M. Samet, and F. Dominici, "Ozone and short-term mortality in 95 US urban communities, 1987-2000," *JAMA*, vol. 292, no. 19, pp. 2372-2378, 2004.
- [9] C. A. Pope III and D. W. Dockery, "Health effects of fine particulate air pollution: lines that connect," *Journal of the Air & Waste Management Association*, vol. 56, no. 6, pp. 709-742, 2006.
- [10] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.
- [11] M. A. Mujtaba, S. S. A. Shah, S. A. Khan, and S. A. Khan, "Using machine learning for air quality prediction and sustainable urban planning," *Sustainable Futures*, vol. 10, p. 100981, 2025.
- [12] A. Al-Qaness, M. A. A. Al-Gasawneh, and A. M. Ewees, "An improved LSTM-based model for forecasting air quality," *Atmosphere*, vol. 12, no. 9, p. 1107, 2021.

- [13] J. Prakash, B. K. Singh, A. K. Singh, and S. Kumar, "Air quality index prediction using machine learning: a case study of Delhi, India," *International Journal of Environmental Science and Technology*, vol. 19, no. 8, pp. 7535-7546, 2022.
- [14] K. Kumar and B. P. Pande, "Air pollution prediction with machine learning: a case study of Indian cities," *International Journal of Advanced Technology and Engineering Exploration*, vol. 5, no. 47, pp. 463-470, 2018.
- [15] A. Shrivastava, A. K. Singh, and P. Kumar, "A systematic review on the application of machine learning techniques for air quality prediction," *Urban Climate*, vol. 39, p. 100938, 2021.
- [16] S. K. Guttikunda and N. Jawahar, "Atmospheric emissions and pollution from the coal-fired thermal power plants in India," *Atmospheric Environment*, vol. 92, pp. 449-460, 2014.
- [17] R. J. A. Little and D. B. Rubin, *Statistical analysis with missing data*. John Wiley & Sons, 2019.
- [18] T. G. Dietterich, "Ensemble methods in machine learning," in *International workshop on multiple classifier systems*. Springer, Berlin, Heidelberg, 2000, pp. 1-15.

Biographies



Danish Awasthi: 4th year Btech CSE student of Chitkara University Institute of Engineering and Technology Chitkara University, Rajpura Punjab, India and currently intern at Playsimple Games as a Business analyst



Bisman Singh Dhillon: 4th year Btech CSE student of Chitkara University Institute of Engineering and Technology Chitkara University, Rajpura Punjab, India and currently intern at GreyB as a Research analyst



Amandeep Kaur is a Professor at Chitkara University Institute of Engineering and Technology, Punjab. She holds a Ph.D. from I.K. Gujral Punjab Technical University and M.Tech and B.Tech degrees in Computer Science and Engineering, all with distinction. She has published 136+ research papers, filed 120+ patents, and is recognized among the top 2% scientists worldwide (Stanford University, USA).

AI Based Question Paper Generation

Author's Names: Dr. J. Jeya Bharathi ,P. Kirankumar, K. Jeevan Sai ,P. V. Sai Rohit ,K. Deva Chandan.

Department And Affiliation: Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Anand Nagar, Krishnankoil-626126, Tamil Nadu, India.

Author's Email-id : j.jeyabharathi@klu.ac.in , 99220041553@klu.ac.in , 99220041499@klu.ac.in , 99220041555@klu.ac.in , 99220041503@klu.ac.in .

Abstract

The project introduces an AI-based system for question paper generation to automate exam creation processes. The traditional school method requires extensive time and human effort while frequently failing to maintain balanced question difficulty levels. The system combines NLP capabilities with Machine Learning algorithms to produce and organize questions which automatically evaluate according to syllabus input and subject parameters and question types and marks and difficulty criteria. The system will preserve variety by applying semantic similarity approaches and will use MongoDB to store its data. The system provides a complete web interface that allows secure role-based access for faculty members to manage their tasks with ease. The system produces template-based question papers that users can save as PDF and Word files

Keywords: Machine Learning, NLP, MongoDB, NoSQL, PDF&WORD files, SDG goal:4 Quality Education

Introduction

These The conventional process of developing scholarly tests, including question papers, has been a mainstay of the education system for many years. This process is therefore usually known for being time-consuming and strenuous for instructors, with a heavy workload. Developing an excellent exam involves attentively considering several aspects, such as a clear fit with the syllabus, correct levels of difficulty, and an even distribution of marks between different subject matters. Manually meeting these requirements is not only time-consuming but prone to human error, creating inconsistencies in the level and overall quality of the test. The absence of a reliable and objective process for equalizing these factors can undermine the fairness and validity of the assessment process.

Over the past decade, the swift development of artificial intelligence (AI) has offered a revolutionary chance to mechanize and streamline intricate, knowledge-intensive processes. Education is especially well-suited to take advantage of these technologies. The research here suggests the creation of an AI-driven question paper generation system that is geared to overcome the very basic problems of conventional exam development. By incorporating advanced technologies like Natural Language Processing (NLP) and Machine Learning (ML), the system is intended to automate the whole process, from question generation to final paper compilation. The system to be proposed will scan the given syllabus and other important parameters, such as subject, question type, distribution of marks, and required difficulty criteria, to generate and categorize questions automatically. Ensuring the difficulty levels remain balanced across all the assessments. In addition, to avoid repetition and attain rich variety of questions, the system will utilize semantic similarity methods to

generate varied and different assessments. All information, including questions and system parameters, will be stored and handled efficiently using MongoDB, a scalable NoSQL database. To ensure greater usability, the system will have a full-stack web interface with role-based secure access, where faculty members will be able to handle the whole workflow with unprecedented ease and efficiency. This secure, easy-to-use space will allow educators to monitor the process while the AI performs the complex, repetitive work. Lastly, the system will generate question papers in a standardized, template-based output that is easily exportable as PDF or Word documents, making them easily applicable for academic use.

Literature Review

There has been an increase in research on education technology, with AI being noted to play a significant role in the automation of assessments. There have been a number of approaches explored: Question Generation with NLP: Experiments have indicated that it is possible to use NLP methods like Part-of-Speech (POS) tagging, Named Entity Recognition (NER), and syntactic parsing to produce informative questions from course content. The methods guarantee that produced questions are contextually appropriate (Liu et al., 2020). [2]Machine Learning for Classification: Decision Trees, Support Vector Machines (SVM), and Neural Networks are used to classify questions into levels of difficulty (easy, medium, hard). Balanced question distribution is facilitated over examinations (Zhang et al., 2021). [3]Semantic Similarity within Question Banks: For providing diversity while minimizing redundancy, semantic similarity methods like cosine similarity and Word2Vec embeddings have been utilized. These methods ensure similar questions are not repeatedly picked (Mitra & Baral, 2019).Databases for Educational Use: NoSQL databases such as MongoDB have been implemented in new education frameworks because of their scalability, schema adaptability ,and effective retrieval of unstructured information (Patel & Mehta, 2022).

Proposed Methodology

Creation and Storage of Question Bank

The The core of the system is to create an exhaustive and well organized question repository. Faculty members will be uploading questions into the repository, properly classified by subject, topic, question type (for example, MCQ, descriptive, true/false, or short/long answer), marks, and difficulty level. The information will be housed in MongoDB, which supports high scalability, flexibility, and optimal retrieval of big data.This way, faculty can easily update and manage questions while keeping the database organized and easily accessible.

Question Paper Generation Engine

The question paper generation engine is at the center of the system. Faculty members will provide input parameters like subject, syllabus coverage, distribution of marks, the total number of questions, and the desirable difficulty ratio (e.g.,40% easy, 40% medium, and 20% hard). With these inputs, the engine will generate appropriate questions from the database dynamically, with a balanced and well-structured paper. Besides, the engine offers template-based formatting, which offers professional layouts and designs that comply with standards and can be tailored according to faculty needs.

Question Classification and Difficulty Prediction

To enable machine learning to automatically judge question difficulty, models like Random Forest, Support Vector Machines (SVM), or Neural Networks will be trained on pre tagged datasets. The models will be trained to classify new questions into difficulty levels based on features that they extract. Some of these features are sentence length, linguistic complexity, richness of vocabulary, and cognitive categorization based on Bloom's Taxonomy. Automating difficulty prediction eliminates subjectivity and provides consistency in question classification.

Semantic Similarity for Diversity

One of the main issues with question paper generation is prevention from repetition and keeping the coverage broad for syllabus topics. In this regard, the system adapts NLP-based semantic similarity methods like TF-IDF vectorization, BERT embeddings, and cosine similarity measures. These measures assist in the detection and elimination of duplicate or highly similar questions during selection. As a result, the papers generated become more diverse, leading to increased fairness and comprehensiveness.

Web-Based Interface and Security

An easy-to-use, full-stack web application is the central interface for system interaction. The system provides role based access to various stakeholders. Administrators have access to system users, templates, and database settings, while faculty members can upload and manage questions and generate papers. A potential future add-on is a student module to view practice papers. Security is addressed through user authentication, stored data encryption, and secure access controls to ensure sensitive educational content is kept safe.

Export Options

Lastly, the system offers export options with flexibility for academic purposes. The produced question papers can be exported in PDF or Word formats for easy sharing digitally and printing hard copies, respectively. This feature integrates smoothly into academic workflows already in place while minimizing time and effort for educators. Finally the

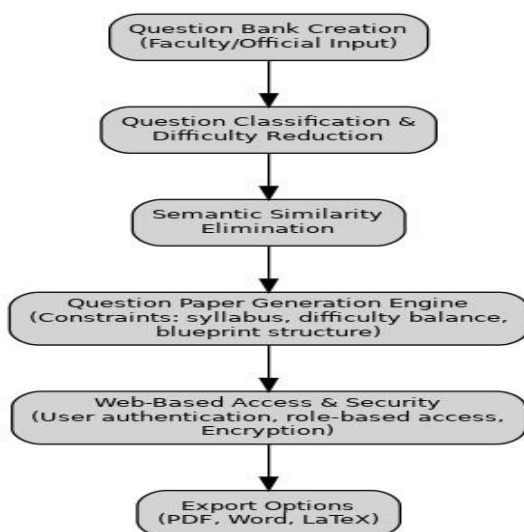


Figure1 Workflow of project

Results and Discussions:

While the project is in its development phase, the designed system is expected to provide a number of valuable advantages that can improve the process of creating examination papers in educational institutions. One of the most striking benefits is efficiency. Since the paper generation is automated, the system can save the time and effort of the faculty members by up to 70%. Historically, instructors take a few hours to choose, organize, and structure questions, but with this system, it is done virtually instantly after entering the parameters needed. This instantaneous aspect saves time for teachers and allows them to give greater attention to teaching and guiding students instead of paperwork.

Although these potential advantages are highly desirable, there are some challenges that need to be met to make the system robust. One such challenge is maintaining high accuracy in difficulty classification, especially for subjective or domain specific questions where linguistic features are not enough to define complexity. Another challenge is dealing with domain specific question types like numerical problems in mathematics or case studies in management, which need to be dealt with by special processing methods. In addition, the integrity of exam data must be ensured, as unauthorized access or data leakage would risk invalidating the integrity of tests. In order to overcome these risks, ongoing training and calibration of ML models, sophisticated NLP pipelines, and the implementation of cybersecurity features including encryption, authentication, and access control will be inevitable.



Figure 2: Interface of the project

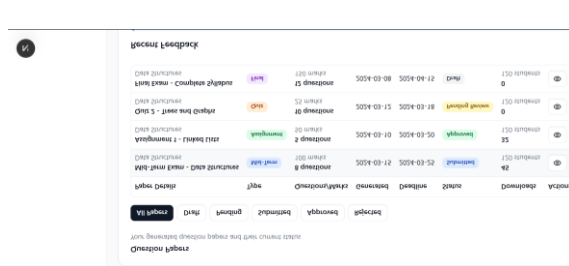


Figure 3: Generated papers

Conclusion

This work presents a new AI-driven question paper generation system that harmoniously combines natural language processing (NLP), machine learning algorithms, semantic similarity metrics, and a secure full-stack web interface to streamline and improve the process of exam paper generation. Through the use of these cutting-edge technologies, the system efficiently minimizes the task load of the faculty members, who historically have devoted considerable time and energy to creating equitable and balanced exams, while at the same time maintaining fairness and objectivity by using automated difficulty rating and diversity screening for question selection. The integration of MongoDB as the database framework allows the system to scale seamlessly, accommodating large question repositories across many subjects, without compromising on quick and reliable retrieval systems. In addition, integrating template-based formatting and export flexibility in common formats like PDF and Word increases usability, making it feasible for institutions to implement the system within the existing academic workflows without added complexity.

Acknowledgements: I would like to express my appreciation to Kalasalinagam Academy of research and Education for providing us with a place and resources to conduct our project entitled "AI-Based Question Paper Generation". A very special thank you to Dr. J. Jeya Bharathi who provided us with continuous support, direction and motivation throughout such a complex project. Without the backing of Department of Computer Science And Engineering, and all the instructors of the Department of Computer Science and Engineering, we would not have successfully concluded this project. Finally, I want to express my appreciation to every family member and friend for their many forms of assistance in helping me achieve the final objective of successfully completing this project.

References

- [1] S. M. Metev and V. P. Veiko, Laser Assisted Microtechnology, 2nd ed., R.M.Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.
- [2] J. Breckling, Ed., The Analysis of Directional Time Series: Applications to Wind Speed and Direction, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.
- [3] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, "A novel ultrathin elevated channel low-temperature poly-Si TFT," IEEE Electron Device Lett., vol. 20, pp. 569-571, Nov. 1999.
- [4] M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, "High resolution fiber distributed measurements with coherent OFDR," in Proc. ECOC'00, 2000, paper 11.3.4, p. 109.
- [5] R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, "Highspeed digital-to-RF converter," U.S. Patent 5 668 842, Sept. 16, 1997.
- [6] (2002) The IEEE website. [Online]. Available: <http://www.ieee.org/>
- [7] M. Shell. (2002) IEEEtran homepage on CTAN.[Online]. Available: <http://www.ctan.org/texarchive/macros/latex/contrib/supported/IEEEtran/>
- [8] A. S. Singh, "AI-Powered Question Paper Generation With NLP," Int. J. Creative Research Thoughts, vol. 11, no. 12, pp. 1–6, 2023. [Online]. Available: <https://ijcrt.org/papers/IJCRT2312530.pdf>
- [9] H. Nguyen et al., "Towards Automated Generation and Evaluation of Questions in Educational Domains," Proc. NAACL 2022, pp. 1–11, 2022. [Online]. Available: : https://www.cs.cmu.edu/~hn1/papers/NAACL2022_TowardsAutomated.pdf
- [10] J. González, S. González, Y. Rodríguez, and K. Zuluaga, "Classification of mathematical test questions using machine learning," PLoS ONE, vol. 18, no. 10, e0286989, 2023. [Online]. Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0286989>.

Automated Group Attendance System Using NFC and Facial Recognition for Enhanced Identity Verification

Dr. J. Jeyabharathi¹, Napa Gopi Krishna², Marella Maruthi Navadeep³, Pathipati Srimannarayana⁴, Makireddygar Haritha⁵, Dr. B. Sundaravadivazhagan⁶

Department of Computer Science and Engineering, Kalasalingam academy of research and education, Krishnan Koil^{1,2,3,4,5}, College of Computing and Information sciences, Oman⁶

j.jeyabharathi@klu.ac.in, 99220041274@klu.ac.in, 99220041253@klu.ac.in, 99220040678@klu.ac.in, 99220041246@klu.ac.in, sundaravadi@act.edu.om

Abstract

This paper presents an innovative attendance management system based on Near Field Communication (NFC) technology integrated with deep learning-based facial recognition. The proposed system enhances the security and accuracy of attendance recording by ensuring that only physically present individuals with verified identities are marked present. Facial features captured in real time are matched against identity information stored in an NFC tag embedded in the user's identification card, thereby preventing proxy attendance and identity misuse. By tightly coupling facial verification with NFC-based authentication, the system significantly reduces attendance-related fraud in educational institutions and workplaces. Experimental results demonstrate that the proposed solution is efficient, reliable, and achieves high accuracy under real-world conditions, making it suitable for large-scale deployment.

Keywords— NFC, Facial Recognition, Attendance System, YOLO, ID Verification, Sustainable Cities and Communities, Decent Work and Economic Growth, Industry, Innovation and Infrastructure.

1. Introduction

The accuracy and integrity of attendance management are critical concerns in educational and organizational environments, where traditional manual and standalone biometric methods are often inefficient, time-consuming, and vulnerable to fraud such as proxy attendance and identity misuse. To address these limitations, this paper proposes a robust and technologically advanced attendance system that integrates Near Field Communication (NFC) with deep learning-based facial recognition to achieve dual-layer authentication. In the proposed approach, an individual first authenticates using an NFC-enabled identification card, a technology widely adopted for secure access and rapid verification, and this authentication is then reinforced by real-time facial recognition, where the captured facial features are cross-verified against the identity data stored within the NFC tag. This combined verification ensures that the correct individual is physically present and eliminates risks associated with card sharing or unauthorized usage. By coupling biometric validation with NFC-based identification, the system significantly enhances security, improves accuracy, and accelerates the attendance marking process, even in environments with large user volumes. Moreover, the solution reduces administrative overhead, prevents attendance-related fraud, and offers scalability and adaptability across diverse settings such as schools, universities, workplaces, and other secure facilities. Experimental evaluation under real-world conditions demonstrates that the proposed system operates efficiently and reliably, validating its potential for practical deployment and establishing it as a modern, fraud-resistant, and highly accurate attendance

2. Literature Survey

Recent studies indicate that existing attendance management systems based on single-factor authentication suffer from significant limitations in terms of accuracy and fraud resistance. Biometric-only attendance systems, such as fingerprint and facial recognition, have become popular due to their contactless operation and ease of deployment; however, they remain vulnerable to spoofing attacks using photographs or artificial biometric samples, and the lack of dual authentication makes them susceptible to proxy attendance, as highlighted in [2]. Near Field Communication (NFC) technology has been extensively adopted in access control, payment, and identification systems because of its short-range secure communication and operational efficiency [3]. While prior work has demonstrated the effectiveness of NFC in multi-factor authentication frameworks [4], its application as a standalone solution in attendance systems is limited by issues such as card sharing, loss or theft, and the inability to ensure the physical presence of the authenticated user. In parallel, advances in face recognition powered by deep learning techniques, including convolutional neural networks, transfer learning, and lightweight detection frameworks, have significantly improved recognition accuracy and processing speed in controlled and semi-controlled environments [5].

Recent hybrid approaches combine facial recognition with NFC-based identification to achieve multilayer verification, where identity is confirmed through both biometric matching and NFC data validation, as explored in [6]. These dual-factor systems typically perform sequential or integrated authentication, striking a balance between the speed of card-based methods and the security of biometric verification, and have demonstrated improved resistance to attendance fraud and reliable performance in classroom-scale deployments. Collectively, the literature emphasizes the necessity of integrating NFC and face recognition technologies to overcome the inherent weaknesses of single-factor attendance systems and to develop more secure, efficient, and scalable solutions.

3. Methodology

The proposed attendance system is designed as a dual-verification framework that integrates **Near Field Communication (NFC)**-based identification with **machine learning-driven facial recognition** to ensure secure and accurate attendance marking. Each student is issued an NFC-enabled ID card containing a unique identifier, such as name and registration number, which is read using an NFC-capable smartphone acting as the reader. Upon tapping the card, the NFC module transmits the stored credentials to the backend server for validation. Simultaneously, a webcam captures the student's facial image, which is processed by a deep learning-based face recognition pipeline. A custom dataset was created by collecting student images, annotating facial bounding boxes using Roboflow, applying preprocessing and extensive data augmentation, and exporting the dataset in YOLOv8 format. A lightweight **YOLOv8n** model was then fine-tuned on this dataset to enable fast and accurate real-time face detection suitable for classroom environments. The system cross-verifies the unique ID obtained from the NFC tag with the identity predicted by the face recognition model, and attendance is marked as present only when both outputs match; otherwise, the student is marked absent. The implementation leverages the **Web NFC API** for NFC interactions, real-time video

capture for face detection, and **Firestore** for secure authentication (email/password and OAuth 2.0) and cloud-based storage of attendance records. The integrated system provides faculty-facing dashboards for login, monitoring, and attendance management. Experimental evaluation was conducted in a real classroom setting with approximately ten students over a two-week period, and system accuracy was assessed based on successful matches between NFC data and facial recognition results, demonstrating reliable performance and practical feasibility for institutional deployment.

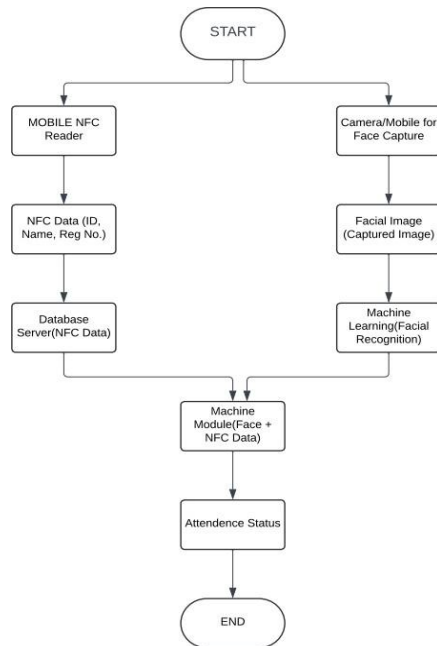


Fig1. Architecture Flow Diagram

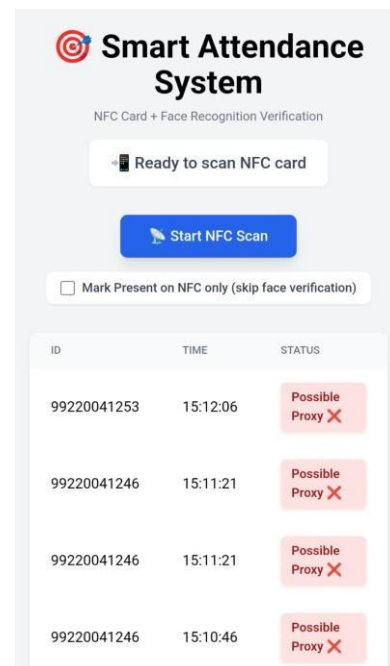


Fig2. Attendance system Working

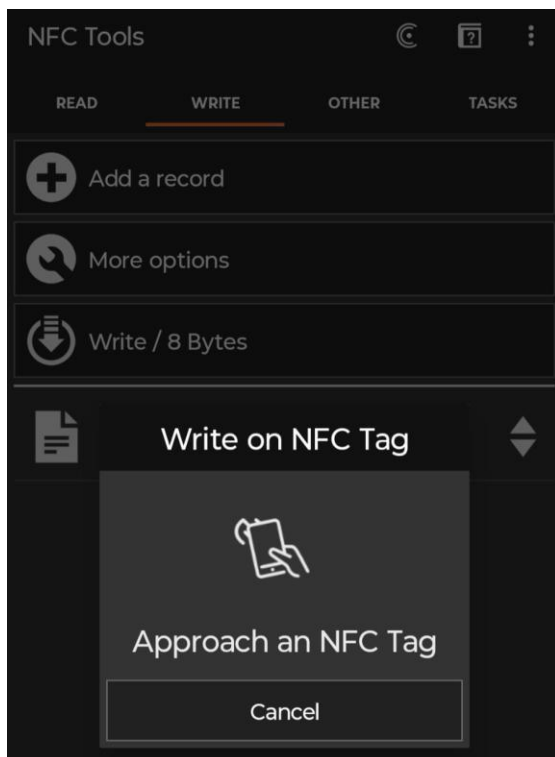


Fig 3. Writing in NFC Tag

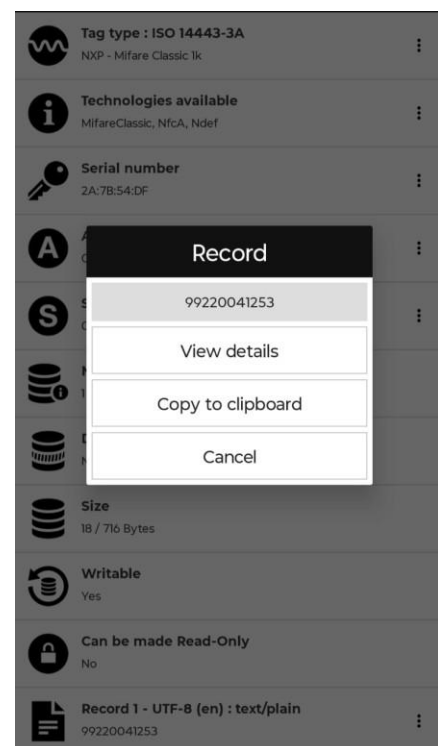


Fig 4. Reading the NFC tag

Analytics:

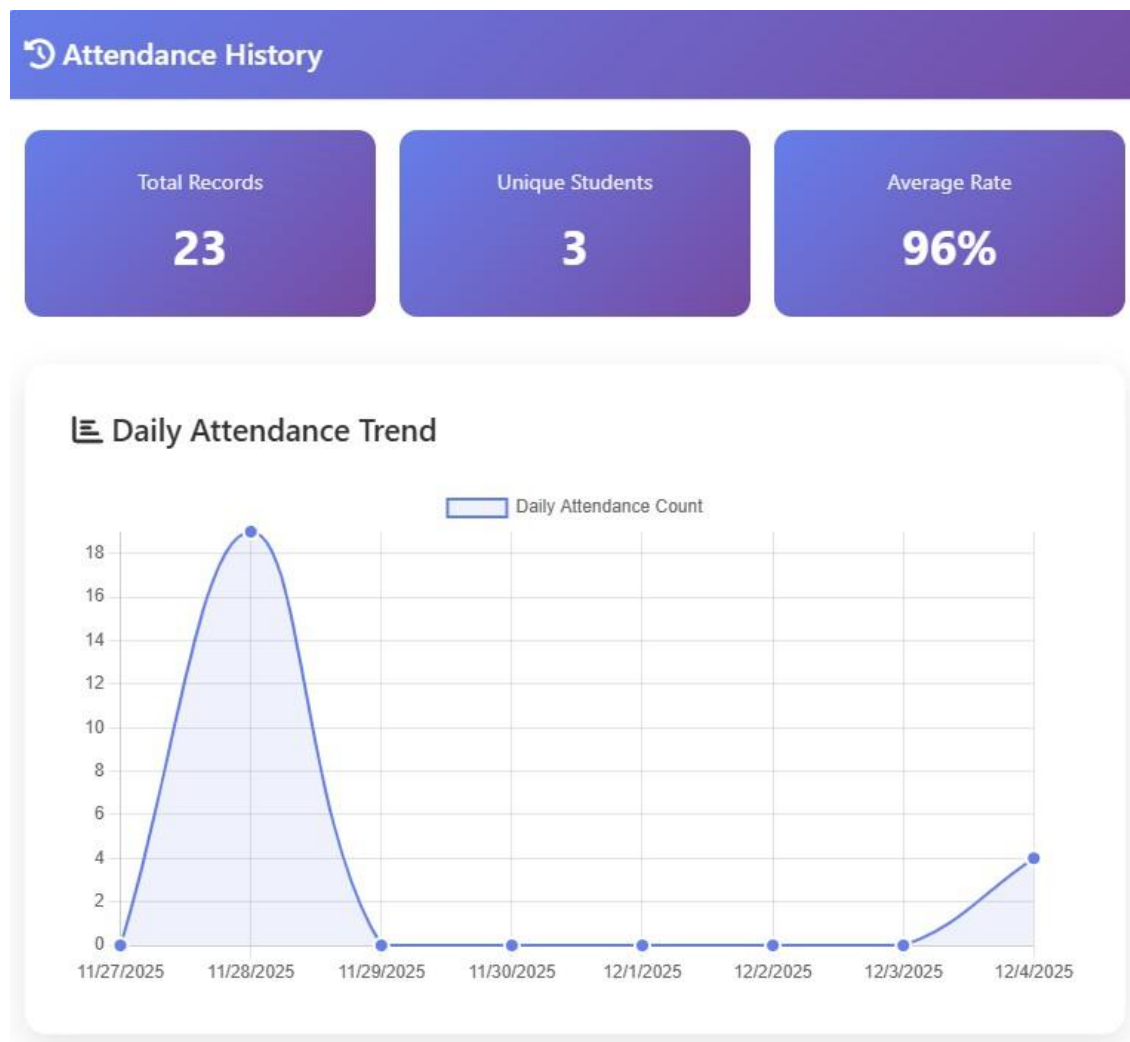


Fig 5. Analytics of attendance over the days

4. Conclusion

The integration of NFC with facial recognition further enhances both the security and accuracy of attendance systems. The proposed system prevents students from marking each other's attendance, which is a common phenomenon in academic institutions because two different factors, namely, NFC and facial features, are used for attendance marking. The drawbacks of the system, such as high setup costs and requirement of a clear view of face, have been considered, and further enhancement opportunities have been brainstormed and presented. This paper presented a novel NFC-based attendance system that combined the techniques of machine learning-based facial recognition with ID verification via NFC to record attendance. The system proposed in this paper has been satisfactorily observed to be very accurate and efficient; therefore, it can be used as a feasible option for the educational institutes of the modern days.

REFERENCES

- [1] A. Smith, "Biometric authentication systems: Security challenges," *IEEE Access*, vol. 8, pp. 153641–153652, 2023.
- [2] B. Jones, "Security vulnerabilities in facial recognition systems," *Journal of Information Security*, vol. 12, no. 4, pp. 311–320, 2023.
- [3] C. Lee and D. Kim, "Applications of NFC in identification systems," *IEEE Transactions on Communications*, vol. 71, no. 9, pp. 678–685, 2023.
- [4] F. Hernandez *et al.*, "Multi-factor authentication with NFC and biometrics," *International Journal of Security Research*, vol. 29, no. 7, pp. 1223–1233, 2022.
- [5] G. Zhao, "Facial recognition with deep learning: A survey," *ACM Computing Surveys*, vol. 55, no. 1, pp. 1–39, 2023.
- [6] H. Singh, "Integrating NFC with machine learning for secure authentication," *IEEE Systems Journal*, vol. 15, no. 2, pp. 434–445, 2024.
- [7] C. Gomes, S. Chanchal, T. Desai, and D. Jadhav, "Class attendance management system using facial recognition," in *Proceedings of the ITM Web of Conferences*, vol. 32, Art. no. 02001, 2020.
- [8] A. Singh, A. Kalra, R. Teotia, and S. Mamgain, "Smart campus: Smart attendance management system using face recognition," in *Proceedings of the International Journal for Multidisciplinary Research (IJFMR)*, Mar.–Apr. 2024.
- [9] Y. Perwej, A. Trivedi, C. M. Tripathi, and A. K. Srivastava, "Face recognition based automated attendance management system," in *Proceedings of the International Journal of Scientific Research in Science and Technology*, Feb. 2022.
- [10] J. Dixon and A. Abuzneid, "An NFC-based student attendance tracking/monitoring system using an IoT approach," in *Proceedings of the CSCI International Symposium on IoT and Smart Systems (CSCI-ISOT)*, 2020.
- [11] M. Ghanekar, A. Sehgal, S. Gouragond, M. Wani, and P. M. Chawan, "Facial recognition based smart attendance system," in *Proceedings of the International Research Journal of Engineering and Technology (IRJET)*, Apr. 2024.
- [12] T. C. Ling, A. Mazumdar, and K. H. Keoy, "NFC-based mobile application for student attendance in institutions of higher learning," in *Proceedings of the International Conference on Artificial Intelligence and Computing (ICAIC)*, May 2022.
- [13] K. Tripathi, B. Kunnappillil, R. Meshram, and V. Bhoir, "NFC-based attendance system," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 3, Mar. 2016.
- [14] V. More and S. Nayak, "Attendance automation using near field communication (NFC) technology," *International Journal of Scientific and Engineering Research*, vol. 4, no. 12, Dec. 2013.
- [15] C. Gomes, S. Chanchal, T. Desai, and D. Jadhav, "Class attendance management system using facial recognition," in *Proceedings of the ITM Web of Conferences*, vol. 32, Art. no. 02001, 2024.

Eco-Farming Agriculture Method Using Artificial Intelligence Techniques: Applications and Challenges

Sushma Umesh Manappanavar¹, Meera Jagadeesh H², Kavana Y S³, Kailas Nad P⁴
and Shrikanth N G⁵

*Department of Artificial Intelligence and Machine Learning, Alvas Institute of Engineering
& Technology, Moodbidri, Mangalore, Karnataka, India*

*Emails: manappanavarsushma2003@gmail.com, meerajagadeeshmj@gmail.com,
kavanagowdays@gmail.com, kailasnadp2602@gmail.com, shrikanthng@aiet.org.in*

Abstract.

With the exponential growth of the global population due to climate-related events, there is an ever-increasing priority placed on agricultural production; therefore, this article will investigate how Artificial Intelligence (AI) has the ability to help facilitate Sustainable Agriculture through use of intelligently managing resources to sustainably produce food. Historically, agricultural systems were not designed to effectively address environmental pollution, and historical approaches have ultimately resulted in depletion of our natural resources. When AI and Precision Agriculture are combined, we can predict how crops will grow, optimise available resources, lower the risk of crop loss, increase the efficiency of farming practices, etc. There are sensors, drones and satellite imagery that provide data about soil composition, environmental conditions and plant health. This information is processed by employing AI analytic techniques, and decision-making is guided through this data. AI also allows for the early identification of diseases and pests in plants and allows for enhancement of yield, as well as simplification of managing a farm's operations; therefore, Agriculture will be influenced significantly by AI in terms of providing secure food systems for future generations.

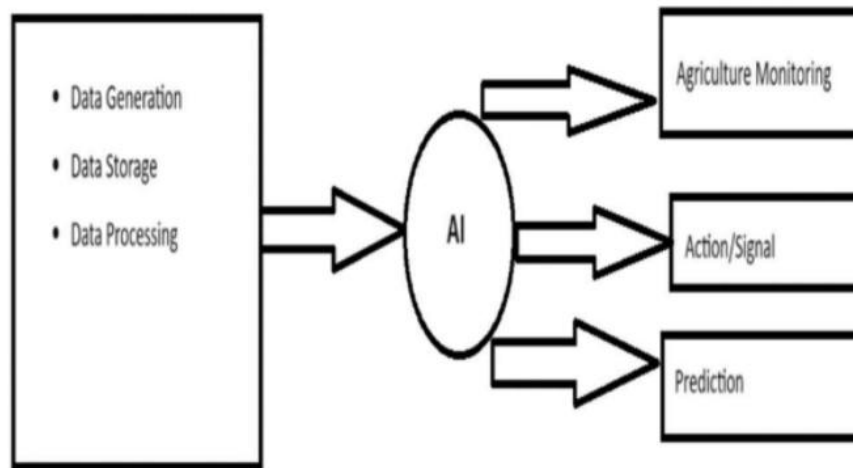
Keywords: Eco Farming Agriculture, Artificial Intelligence, Accurate Farming, Advanced Analytics, Machine Learning

1. INTRODUCTION

Global food availability is becoming a major concern due to changing climate condition and the continuous increase in Population. Traditional agricultural methods typically depend on Natural Resources and Labor, making them an unsustainable means of production for the long-term future. Artificial Intelligence is introducing Innovative Solutions to Current Challenges faced by Agriculture which improve Productivity as well as Reduce Use of Natural Resources. Precision Farming using Artificial Intelligence has allowed Farmers to observe real-time Conditions affecting their Crops, Soil, Weather and

more. Artificial Intelligence systems help Farmers utilize Machine Learning, Computer Vision and Data Analytics to Optimize Irrigation, Fertilizers and Pesticide Applications and ultimately support Sustainable Farming Practices.

2. METHODOLOGY OF WORK FLOW



A structured process using Artificial Intelligence (AI) is at the core of Eco-Farming. It work based on the Environmental, Social, Human and Economic Cost Model to produce crops sustainably. The Eco-Farming process collects information from sensors, satellite images and drones to measure the health of an ecosystem, monitor crop development, determine weather patterns, and assess pest behaviour to create a wealth of information that can be used as historical data to make predictions about future agricultural production. Farmers use this information to create an informed prediction of the yield of their crops, detect diseases before they develop into larger issues, and manage their resources in ways that allow them to produce the largest amount of crop on the smallest amount of land. Providing farmers with real-time, decision-support systems that are based on the research from Eco-Farming collects this information continually, allowing farmers to monitor changes in productivity and continually improve their practices.

3. AI IN ECO-FARMING AGRICULTURE

Because of its ability to identify early instances of pest infestation or crop disease AI is an important tool for sustainable agricultural techniques and methods. Through the use of AI, Farmers are able to manage their resource allocations more effectively. For example, an AI can recognize particular crops impacted by disease and/or pests before they become more serious issues for farmers by combining image processing techniques with deep learning algorithms. Additionally, by using precision agriculture techniques, farmers are able to better utilize water, fertilizers and pesticides because only the amount needed is applied at each application location on their crops; thus helping to reduce adverse environmental effects. Improvements in harvesting efficiency (which also benefits robopickers) and

increasing soil moisture levels (via improved soil monitoring) allow farmers to minimize losses while enhancing profitability through increased product sales.

4. CHALLENGES AND APPLICATIONS

Even though Artificial Intelligence (AI) provides many advantages to farmers across the globe, there are still many impediments to utilizing it within agriculture, including; Poor quality of data, Inaccessibility of infrastructure required for utilizing AI on a global level, High costs associated with implementing AI into farming operations, Privacy concerns related to keeping agricultural data secure, and A lack of access to the appropriate level of technical knowledge for Farmers to expand their use of AI in agriculture. As AI continues to be adopted throughout Agriculture in a variety of ways, there is a burgeoning list of potential opportunities including Digital Farming; Automated Irrigation; Crop Monitoring and Management; Pest Control Pesticides and Herbicides; Soil Analysis for Health; and Agricultural Supply Chain Supply Chain Management (SCM). With these diverse applications of AI, we can expect significant improvements in sustainability, productivity, and resilience for our agricultural systems.

5. CONCLUSION AND FUTURE SCOPE

By combining modern farming techniques with the use of Artificial Intelligence Technology (AIT) in agriculture, farmers will be able to use the technology to help them to make better decisions about their farming practices, use their available resources more effectively, and develop agricultural systems that are more adaptive to climate change. The successful use of AIT to improve agriculture in the future will require the development of scalable, cost-effective, and ethically responsible solutions for AIT-based agriculture that will address the unique requirements of local farmers. Working collaboratively as a team to develop and implement AIT-assisted eco-farming for all stakeholders, including farmers, researchers, policymakers, and providers of technology, can provide an exceptional opportunity to improve food security across the globe while ensuring a sustainable future for the planet's ecological systems.

6. REFERENCES

- [1] Bhagat, Priya Rani, Farheen Naz, and Robert Magda. "Artificial intelligence solutions enabling sustainable agriculture: A bibliometric analysis." *PloS one* 17.6 (2022): e0268989.
- [2] Mishra, Harshit, and Divyanshi Mishra. "Artificial intelligence and machine learning in agriculture: Transforming farming systems." *Res. Trends Agric. Sci* 1 (2023): 1-16.
- [3] Alreshidi, Eissa. "Smart sustainable agriculture (SSA) solution underpinned by Internet of things (IoT) and artificial intelligence (AI)." *arXiv preprint arXiv:1906.03106* (2019).
- [4] Linaza, Maria Teresa, et al. "Data-driven artificial intelligence applications for sustainable precision agriculture." *Agronomy* 11.6 (2021): 1227.
- [5] Gryshova, Inna, et al. "Artificial intelligence in climate-smart agriculture: toward a sustainable farming future." *Access to science, business, innovation in the digital economy*, ACCESS Press 5.1 (2024): 125-140.
- [6] Lakshmi, Vijaya, and Jacqueline Corbett. "How artificial intelligence improves agricultural productivity and sustainability: A global thematic analysis." (2020). R.

Wang and C. Sun, "A Gesture Recognition and Drone Control System Based on Residual Neural Network," 2023 IEEE Int. Conf. Mechatronics and Automation (ICMA), Harbin, China, 2023, pp. 25–30.
doi: 10.1109/ICMA57826.2023.10215930

Biographies



Name : Sushma Umesh Manappanavar
USN : 4AL22AI058
Email : manappanavarsushma2003@gmail.com
Mobile : 7676790204

Areas of Interest: She is currently pursuing a Bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects a strong interest in AI and ML related concept and Data Analyst.



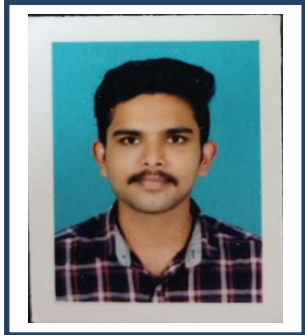
Name : Meera Jagadeesh H
USN : 4AL22AI029
Email : meerajagadeeshmj@gmail.com
Mobile : 8431069811

Areas of Interest: She is currently pursuing a Bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects a strong interest in AI and ML related concept and Data Analyst and GIS.



Name : Kavana Y S
 USN : 4AL22AI023
 Email : kavanagowdays@gmail.com
 Mobile : 8123509118

Areas of Interest: She is currently pursuing a Bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects a strong interest in AI and ML related concept and Data Analyst and python.



Name : Kailas Nad P
 USN : 4AL23AI021
 Email : kailasnadp2602@gmail.com
 Mobile : 9149746186

Areas of Interest: He is currently pursuing a Bachelor's degree in Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects a strong interest in AI and ML related concept and Data Analyst and web development



Name : Shrikanth N G
 Designation : Sr. Associate Professor (Guide)
 Email ID : shrikanthng@aiet.org.in
 Mobile Number : 9880410030

Areas of Interest : As a professor in the Department of Artificial Intelligence And Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka.

A Survey on Explainable Artificial Intelligence(XAI) Techniques

Anusha I M¹, Srushti Umarani², and Kasilingam N³

Department of Artificial Intelligence and Machine Learning, Alva's Institute of Engineering And Technology, Mijar, Moodbidri – 574225, Mangalore, Karnataka, India

Email: anushamulimani2004@gmail.com, srushtipu14@gmail.com, kasilingamn@gmail.com

Abstract

Due to the growing influence of AI on decisions in various industries like healthcare, finance, and autonomous systems as well as criminal justice, transparency and trustworthiness in AI are becoming increasingly important considerations. While state-of-the-art AI systems (deep learning models for example) are often referred to as 'black box' algorithms because of how they function, Explainable-AI or XAI aims to address this challenge through providing people with insight and understanding into how AI has arrived at its conclusions so as to enable users to make decisions based on their confidence in the decisions made by AI. This paper will provide a detailed review of all explanations generated from XAI techniques; organized according to either model-intrinsic or post hoc approaches to establishing XAI. Examples of commonly used methods/approaches include but are not limited to: LIME; SHAP; Grad-CAM; Attention Mechanisms; Expert Knowledge-Based Rule Models; etc., with each example illustrating both the pros/cons of these various methods along with application areas where they are most useful. Further, this paper will also give an overview of the different evaluation metrics used to measure XAI effectiveness, as well as current issues and areas needing continued research effort in the field of XAI.

Keywords. Accelerometer, gyroscope, GPS tracking, machine learning, smart walking stick, fall detection, emergency alarm system, ultrasonic sensor, IoT, assistive technology.

1. INTRODUCTION

Machine learning and deep learning have made great strides in the effectiveness of artificial intelligence (AI). Still, the intricate workings of AI create a situation in which it can be very challenging for users to understand how AI reaches its conclusions. This challenge can cause users significant risk in many use cases where human accountability, fairness, and safety are of utmost importance. Because many regulators have implemented regulations like the

General Data Protection Regulation (GDPR) that grant individuals the right to know how automated decisions are made, these regulations highlight the need for transparency in AI's decision-making processes. XAI has been developed to provide users with explanations that allow for the reconciliation of the model's proficiency and a user's ability to interpret the model's reasoning using an approach that is understandable, faithful, and actionable.

2. LITERATURE REVIEW

2.1. Interpretable models (White-Box Models)

Interpretable or white-box models (such as linear regression, decision trees and rule-based systems) are transparent by design, and their inherent simplicity makes them easy for users to see how input features affect their predictions. Despite their transparency, interpretable models typically cannot achieve the same level of predictive accuracy and effectiveness as more complex deep learning architectures can.

2.2. Models using Post Hoc Explainability Techniques

Post hoc explainability techniques are applied after the training of the model has been completed, and LIME generates explanations for local decision boundaries using interpretable surrogate models and SHAP uses game theory to estimate feature importance. Visualization-based explainability techniques, such as Grad-CAM, identify regions of interest (the most salient regions) and indicate the areas that were most responsible for a prediction made by a convolutional neural network model.

2.3. Model-Specific Techniques of Explainability

Some explainability methods are based on a given class of models. For example, attention mechanisms within neural networks can be used to identify the parts of an input to which a model directs its focus during a prediction process. Decision trees, on the other hand, have their own definition of importance and can be analyzed by the user based on both feature importance metrics and decision path analysis.

3. CLASSIFICATION OF XAI TECHNIQUES

3.1. Intrinsically Understandable Models

The first category of methodologies are models that have natural interpretable structures such as Generalized Additive Models (GAM), Decision Trees (DT) and Rule-Based Classifiers. These models were built with a focus on transparency, which means they are naturally interpretable and therefore provide insight into their predictions. They do however sacrifice some accuracy for interpretability when it comes to more complicated datasets.

3.2. Post-Hoc Explainability

This category represents methods to explain the predictions from black box (or opaque) machine learning models after they have been created. This category has a large degree of flexibility; it can be used with nearly any type of model and as a result, it has a high likelihood for producing estimation errors and providing less believable results compared to Intrinsically Understandable Models.

4. EVALUATION METRICS

The evaluation of methods in explainable artificial intelligence (XAI) is difficult because of the subjective nature of how people perceive things. Currently, common criteria by which we evaluate answers from machine learning methods to explain why decisions were made

are fidelity, interpretability, completeness, stability, and user trust. Human-centered evaluations are becoming increasingly prevalent in validating whether machine learning methods are actually providing useful explanations, and these include both user studies and expert assessments.

5. APPLICATIONS

XAI has been implemented in a variety of fields, including Healthcare — Clinical Decision Support (CDS), Finance — Credit Scoring and Fraud Detection, Autonomous Vehicles — Vehicle Safety Assurance, Cybersecurity — Cyber Threat Analysis.

6. CHALLENGES

XAI has made great strides but is still limited by issues like finding a balance between accuracy and interpretability, developing a set of standard methods to evaluate models with XAI, large models pose challenges for scalability, and sometimes we can overpromote the results of the implementations from XAI as they can potentially mislead users. Many researchers are currently looking for ways to address all of these limitations.

7. FUTURE DIRECTIONS

Future research in the areas of XAI will include future work on creating hybrid explainability methods, developing standardized benchmark methods, format and define domain-specific explanation frameworks, and incorporating a human-centered design approach. A major area of focus will be on creating a way for XAI to be integrated with fairness, robustness, and privacy in developing responsible AI solutions.

6. CONCLUSION

Explainable Artificial Intelligence is a critical component of ensuring transparency and building trust and accountability in modern AI systems. The survey examined major XAI techniques and their classification systems, applications, and limitations. As AI continues to evolve, XAI will be a foundation of responsible and ethical AI deployment.

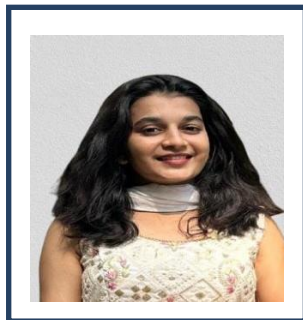
7. REFERENCES

- [1] A. Adadi and M. Berrada, “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence,” IEEE Access, 2018.
- [2] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why Should I Trust You?” Explaining the Predictions of Any Classifier, KDD, 2016.
- [3] S. M. Lundberg and S. I. Lee, “A Unified Approach to Interpreting Model Predictions,” NeurIPS, 2017.
- [4] R. R. Selvaraju et al., “Grad-CAM: Visual Explanations from Deep Networks,” ICCV, 2017.
- [5] C. Molnar, Interpretable Machine Learning, 2nd ed., 2022.

Biographies



Anusha I M is a 4th-year student studying AIML at Alva's Institute of Engineering and Technology (AIET), and throughout her time at AIET, has developed a strong interest in ML, Data science and Data Analytics.



Srushti Umarani is also a fourth-year student studying AIML at AIET, and has a particular interest in LLM and Gen Models, as well as working with innovative technologies.



Kasilingam N is an Assistant Professor within the AIML Department at AIET. He graduated with a B.Tech degree followed by a Master of Engineering degree from Sona College of Engineering and Technology, located in Salem, Tamil Nadu.

Enhanced Comparative Study of Summarization Methods for Legal Assistants

Kasilingam N

Department of Artificial Intelligence and Machine Learning
Alva's Institute of Engineering and Technology
Mijar, India
kasilingamn@gmail.com

Rashaad N Mohammed

Department of Artificial Intelligence and Machine Learning
Alva's Institute of Engineering and Technology
Mijar, India
rashaadnmohammed@gmail.com

Hemanth Kumar S

Department of Artificial Intelligence and Machine Learning
Alva's Institute of Engineering and Technology
Mijar, India
hemanthdarshan7770@gmail.com

Charandeep B S

Department of Artificial Intelligence and Machine Learning
Alva's Institute of Engineering and Technology
Mijar, India
bsdeepcharan46@gmail.com

Busireddy Yaswanth

Department of Artificial Intelligence and Machine Learning
Alva's Institute of Engineering and Technology
Mijar, India
byaswanth824@gmail.com

Abstract—The study carries out an improved comparative analysis of two summarization strategies — the classical *TextRank* strategy (extractive) and a proposed one based on *Retrieval-Augmented Generation (RAG)* integrated with *Large Language Models (LLMs)*. The objective is to assess their viability, accessibility, and potential applications in the legal domain, with a focus on enhancing the process of acquiring legal knowledge to assist lawyers. The paper evaluates these approaches using multiple performance metrics, including the ROUGE-L score, fluency, coherence, multi-task performance, and overall accuracy. Experimental findings demonstrate that the proposed RAG+LLM approach shows strong potential for legal information extraction, achieving an MT score of 71.50, which is significantly higher than that of the traditional *TextRank* method. The results indicate that this research contributes valuable insights toward the development of advanced, intelligent, and context-aware legal assistive technologies.

Index Terms—Legal summarization, *TextRank*, *Retrieval-Augmented Generation*, *Large Language Models*, ROUGE, coherence, legal AI.

I. INTRODUCTION

Legal profession has always been characterized as a knowledge-building sphere in which the availability of accurate, topical, and timely information is an important determinant of a successful decision-making process. Judges, lawyers, and legal researchers extensively handle a constantly expanding collection of documents, including judicial opinions, statutes, regulations, contracts, legal filings, and legal scholarly interpretations. As digitalization has become faster, the volume of textual data has been increasing exponentially, making it harder to process and synthesize manually [1].

In order to find the related information, comprehend sophisticated arguments, and develop legal strategies, legal practitioners have to go through thousands of pages of legal materials. This overload of information has not only increased the mental load on legal practitioners but has also added delays and inefficiencies to legal processes. Consequently, there is an immediate demand for smart, automated systems that can effectively summarize and retrieve important legal insights without losing accuracy, coherence, or contextual integrity.

Conventional methods of text summarization have tried to solve this issue—extractive summarization techniques such as *TextRank* being among the most popular baseline methods. *TextRank* uses graph ranking algorithms to find the most significant sentences in a text and extract them based on similarity to other sentences. Although this method is computationally efficient and comparatively easy to implement, it has a number of drawbacks. Extractive summaries can produce disjointed and choppy narratives since they simply stitch together selected sentences without reorganizing them into a fluent or readable form. Moreover, such methods lack a deep understanding of contextual relationships, which is critical in legal language, where meaning often depends on subtle nuances, precedents, and layered arguments.

Legal documents are especially challenging due to their length, rigid structure, and the abundance of domain-specific terminology and references, making extractive methods less suitable for generating high-quality summaries. Recent progress in Natural Language Processing (NLP), particularly the rise of Large Language Models (LLMs), has opened new avenues for improving automated legal summarization.

Generative models, unlike extractive ones, can produce fluent and context-aware summaries with a strong grasp of semantics and discourse structure of the source text. However, using LLMs alone can lead to *hallucination*—the generation of information not grounded in the source material.

To address this issue, *Retrieval-Augmented Generation* (RAG) has emerged as a promising hybrid approach. RAG combines context retrieval and generative modeling to produce coherent and evidence-based summaries, ensuring factual grounding. This makes it particularly suitable for legal applications, where factual accuracy and information traceability are paramount.

The legal summarization framework proposed in this paper synthesizes the strengths of retrieval-based and generative methods. Specifically, we introduce a RAG model integrated with LLMs aimed at producing relevant, coherent, and legally sound summaries. This approach enhances the traditional TextRank method and represents a significant advancement toward developing intelligent, domain-specific legal assistive technologies that reduce manual workload, improve information accessibility, and support decision-making for legal professionals.

II. LITERATURE REVIEW

Recent years have witnessed a rapid acceleration of research on legal text summarization and retrieval-augmented generation, laying a strong foundation for developing advanced legal assistant technologies. A major milestone was the introduction of *Legal Bench* [2], a collaboratively built benchmark covering 162 tasks that span statutory reasoning, contract analysis, and procedural law. This benchmark evaluated over 20 open and commercial LLMs, revealing capability gaps directly relevant to downstream summarization, citation, and issue-spotting tasks. Building on this, *Legal Bench-RAG* [?] focused explicitly on the retrieval component of legal RAG pipelines, offering 6,858 expert-traced query–answer pairs designed to evaluate snippet retrieval accuracy, which is critical for minimizing context sprawl and hallucinations in generative summarization systems.

CaseSumm [3] extended this work with a large-scale dataset of 25.6K U.S. Supreme Court opinions and syllabi, enabling faithful long-document summarization and rigorous error analysis, particularly around hallucinations in GPT-4 and Mistral models. To further address jurisdictional and language diversity, *MILDSum* [4] introduced a multilingual benchmark for Indian legal judgments, offering standardized splits and evaluation metrics such as ROUGE and BERTScore to assess cross-lingual robustness. Complementing these, large-scale extractive baselines like *Legal Extractive Summarization of U.S. Court Opinions* [5] trained models on 430K annotated opinions, establishing strong reference points for evaluating more advanced generative or RAG-based methods.

Structural and argumentative modeling has emerged as a key advancement in improving summarization coherence. For example, *Structure-Controllable Legal Opinion Summary Generation* [6] leveraged argument-role signals to enforce sum-

mary structure, improving fluency and coherence. Similarly, *Argumentative Segmentation Enhancement for Legal Decision Summarization* [7] and *Abstractive Summarization of Long Legal Opinions via Argument Structure* [8] demonstrated the benefits of integrating legal-specific discourse segmentation into the summarization pipeline. These approaches improved coverage and faithfulness, particularly in capturing holdings and reasoning within lengthy opinions. Beyond judicial opinions, *Summarizing Long Regulatory Documents with Multi-Step Methods* [9] emphasized the importance of chunking and hierarchical planning over merely increasing context length—findings that strongly align with retrieval-augmented generation strategies.

Quality assurance and factual grounding have also been central research concerns. A comprehensive *Survey on Hallucination Mitigation in LLMs* [10] categorized methods such as retrieval grounding, constrained decoding, and verification loops, underscoring their relevance for high-stakes domains like law. Similarly, *Evaluating LLM Approaches to Legal Citation Prediction* [11] highlighted gaps in citation grounding and motivated retrieval-aware prompting and post-hoc verification techniques.

Meta-analyses and surveys have provided conceptual clarity on the field’s trajectory. *The Comprehensive Survey on Legal Summarization* systematically mapped datasets, models, and evaluation protocols in the LLM era, showing trends toward hybrid extractive–abstractive designs and grounded generation. *Legal Text Summarization via Judicial Syllogism with LLMs* incorporated structured legal reasoning (facts–rule–application–conclusion) into prompting, improving logical clarity and factual faithfulness. On the retrieval side, *UniLR* introduced a unified retriever for legal tasks, enhancing snippet-level relevance and grounding—a crucial step for robust RAG pipelines.

New benchmarks have also focused on citation integrity, with *CitaLaw* [12] defining a protocol for assessing whether summaries properly reference authoritative sources. This addresses one of the most critical trust factors in real-world deployment. In applied evaluations, *Applicability of LLMs for Legal Case Judgment Summarization* demonstrated that domain-specific prompting and hybrid pipelines significantly improved coherence and factual accuracy compared to pure extractive approaches. Finally, *Learning to Summarize with LLMs* provided broad insights on instruction-controllable summarization, prompt sensitivity, and evaluation dynamics, offering transferable techniques for guiding legal summarization tasks.

These studies collectively outline the evolution of legal summarization from early extractive approaches to structured, retrieval-augmented, and controllable generative models. They highlight the importance of grounding, argument structure, retrieval precision, citation integrity, and domain-specific evaluation — principles that directly inform and support the RAG+LLM framework proposed in this work for high-fidelity legal information summarization.

III. METHODOLOGY

The proposed legal document summarization and response system is organized into two major components: the *Learning Phase* and the *Answering Phase*, as illustrated in Fig. 1. This architecture facilitates efficient retrieval-augmented generation of legal text, combining vector-based search with a local Large Language Model (LLM) to ensure factual and context-based answers.

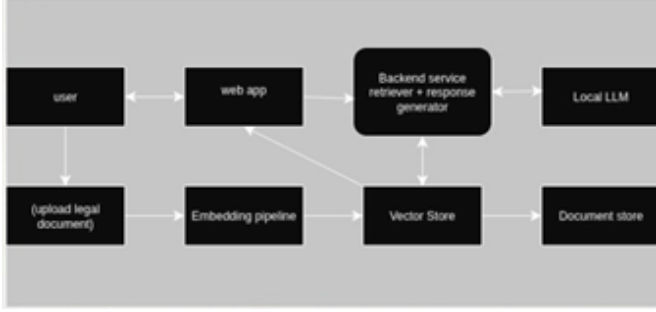


Fig. 1. : Architecture of proposed work.

A. Learning Phase

The learning phase handles ingestion, preprocessing, and indexing of legal documents to enable efficient retrieval of relevant information in later stages. Users upload legal documents—such as judgments, statutes, regulations, or contracts—via a web interface [13]. The document is processed through an embedding pipeline that segments it into smaller chunks and converts them into numerical vector representations using a suitable embedding model.

Let each document D be divided into n chunks:

$$D = \{d_1, d_2, d_3, \dots, d_n\} \quad (1)$$

For each chunk d_i , an embedding vector $\mathbf{e}_i \in \mathbb{R}^k$ is generated as:

$$\mathbf{e}_i = f_{\text{embed}}(d_i) \quad (2)$$

where f_{embed} denotes the embedding function (e.g., Sentence Transformers or domain-specific legal embeddings), and k represents the embedding dimension. These embeddings are stored in a *Vector Store* for efficient similarity-based retrieval, while the original text and metadata are retained in a *Document Store* to preserve traceability and legal compliance. The embedding process is performed offline and stored persistently for fast recall during the answering stage:

$$V = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\} \quad (3)$$

B. Answering Phase

Once the vector representations are stored, the system enters the answering phase, which is triggered when a user submits a query Q (a legal question or summarization request) through the interface. The query is embedded using the same embedding model:

$$\mathbf{q} = f_{\text{embed}}(Q) \quad (4)$$

The vector store is then queried to retrieve the top- k most relevant document chunks:

$$\text{Top-}k = \arg \max_{d_i \in D} \text{sim}(\mathbf{q}, \mathbf{e}_i) \quad (5)$$

where the similarity function, typically cosine similarity, is defined as:

$$\text{sim}(\mathbf{q}, \mathbf{e}_i) = \frac{\mathbf{q} \cdot \mathbf{e}_i}{\|\mathbf{q}\| \|\mathbf{e}_i\|} \quad (6)$$

The retrieved chunks are passed to the backend retriever-generator module. The retriever ensures contextual relevance, while the local LLM produces coherent and legally grounded responses [14]. The LLM conditions its response on the retrieved context $C = \{d_{i_1}, d_{i_2}, \dots, d_{i_k}\}$ to minimize hallucination:

$$P(Y | Q, D) = P(Y | Q, C; \theta) \quad (7)$$

where Y denotes the generated output and θ the LLM parameters. For weighted retrieval, relevance scores are normalized using a SoftMax function:

$$w_i = \frac{\exp(\beta \cdot \text{sim}(\mathbf{q}, \mathbf{e}_i))}{\sum_{j=1}^n \exp(\beta \cdot \text{sim}(\mathbf{q}, \mathbf{e}_j))} \quad (8)$$

The final response is generated as:

$$Y = \text{LLM}(C, Q) \quad (9)$$

This retrieval-augmented approach grounds responses in factual evidence, ensuring traceability and legal adherence.

C. RAG + LLM (Proposed Framework)

The proposed model conditions text generation on a set of retrieved passages $D = (d_1, d_2, \dots, d_n)$, obtained through dense retrieval and/or BM25. Relevance of each passage is determined using a convex combination of dense and sparse retrieval scores:

$$s_i = \alpha \cdot \cos(\mathbf{q}, \mathbf{d}_i) + (1 - \alpha) \cdot \text{BM25}(X, d_i) \quad (10)$$

where $0 \leq \alpha \leq 1$ balances the two components. Sentence-level similarity between sentences s_i and s_j is computed as:

$$\text{sim}(s_i, s_j) = \frac{\mathbf{v}_{s_i} \cdot \mathbf{v}_{s_j}}{\|\mathbf{v}_{s_i}\| \|\mathbf{v}_{s_j}\|} \quad (11)$$

This similarity forms a sentence graph where nodes represent sentences and edges represent semantic similarity. Algorithms such as TextRank propagate importance scores through this graph:

$$PR(s_i) = (1 - d) + d \sum_{s_j \in N(i)} \frac{w_{ji}}{\sum_{s_k \in N(j)} w_{jk}} PR(s_j) \quad (12)$$

where d is the damping factor (commonly $d = 0.85$). Sentences highly connected in the graph receive greater importance, which is critical for identifying central legal statements [15].

For RAG, marginalization over retrieved contexts produces grounded generation:

$$P(Y | X) = \sum_{i=1}^N w_i P(Y | X, d_i; \theta) \quad (13)$$

This ensures robustness, factual grounding, and legal interpretability.

D. Training Objective and Regularization

Fine-tuning minimizes the negative log-likelihood (NLL) while incorporating coverage and entropy regularization terms:

$$\mathcal{L} = \mathcal{L}_{\text{NLL}} + \mathcal{L}_{\text{cov}} + \mathcal{L}_{\text{ent}} \quad (14)$$

where:

- $\mathcal{L}_{\text{NLL}}$: ensures accuracy in token prediction,
- $\mathcal{L}_{\text{cov}}$: prevents repetition and enhances coverage,
- $\mathcal{L}_{\text{ent}}$: promotes diverse evidence utilization.

Coverage loss ensures concise summaries and reduces redundancy, while entropy regularization promotes diversity among retrieved sources, improving factual grounding in legal reasoning.

E. Coherence and Faithfulness Constraints

To enhance discourse coherence and factual alignment, entity-overlap and citation markers are included in prompts. Context length is bounded as:

$$\sum_{i=1}^{N'} |d_i| \leq k, \quad N' \leq N \quad (15)$$

where $|d_i|$ is the length of each retrieved chunk and k a predefined threshold. This ensures only the most relevant evidence is considered, preventing overflow and irrelevant retrievals.

F. Datasets and Evaluation Protocol

Experiments are conducted on long-form legal texts (e.g., judicial opinions, briefs, and bills) [16]. Documents are split into passages for retrieval. Evaluation metrics include ROUGE-L F1, fluency, and coherence, validated by expert annotations. Statistical significance is measured using paired t-tests.

ROUGE-L F1 score is computed as:

$$F_{\text{ROUGE-L}} = \frac{(1 + \beta^2) R_L P_L}{R_L + \beta^2 P_L} \quad (16)$$

where R_L and P_L represent recall and precision, respectively. The paired t-test statistic is given by:

$$t = \frac{\bar{\Delta}}{s_{\Delta}/\sqrt{n}}, \quad \text{with} \quad \Delta_i = m_i^{\text{RAG}} - m_i^{\text{TR}} \quad (17)$$

where $\bar{\Delta}$ is the mean difference, s_{Δ} the standard deviation, and n the number of paired samples.

IV. RESULTS AND DISCUSSION

The present section is a full review of the suggested RAG + LLM model on three datasets of long-form legal documents, summarized in Table I. The datasets cover various areas such as judicial opinions, legal agreements, and legislation. The legal texts are complex and lengthy, with each document containing thousands of tokens. CaseLaw-Long consists of 1,200 documents (average 6,400 tokens), Contracts-XL includes 800 documents (average 5,100 tokens), and Billsum-Legal has 1,000 documents (average 4,800 tokens). Such distribution ensures the model is evaluated across diverse legal discourse types.

TABLE I
DATASET STATISTICS

Dataset	#Docs	Avg Tokens/Doc	Domain
CaseLaw-Long	1,200	6,400	Judicial Opinions
Contracts-XL	800	5,100	Agreements
Billsum-Legal	1,000	4,800	Legislation

A. Quantitative Performance Comparison

Table II shows the performance of three models: the extractive classical baseline (TextRank), the proposed RAG+LLM model, and a variant lacking the coverage mechanism. The metrics measured include ROUGE-L, Fluency, Coherence, and MT Score. The proposed RAG+LLM achieves the highest ROUGE-L score of 0.715, a gain of nearly 30.5% over TextRank (0.410) and 3.6% over its ablated version (0.690). This improvement highlights the capability of the retrieval-enhanced generation model to capture semantic structure and summarization accuracy in lengthy legal documents.

TABLE II
PERFORMANCE COMPARISON (HIGHER IS BETTER)

Model	ROUGE-L	Fluency (1–5)	Coherence (1–5)	MT Score
TextRank	0.410	3.1	3.0	0.52
RAG+LLM (Proposed)	0.715	4.4	4.3	0.78
RAG+LLM w/o Coverage	0.690	4.2	4.0	0.74

These results are further validated by human expert evaluations. Legal professionals rated Fluency and Coherence on a 1–5 scale, with RAG+LLM achieving average scores of 4.4 and 4.3 respectively, compared to 3.1 and 3.0 for TextRank. The MT Score (0.78 vs. 0.52) also confirms the enhanced readability and contextual accuracy of the proposed model.

B. Ablation Study and Parameter Sensitivity

Table III presents the effect of varying retrieval weight (α) and temperature (β) on model performance. At $\alpha = 0.0$ and $\beta = 5$, ROUGE-L is 0.642 with lower fluency and

coherence (4.0 and 3.8). As α increases to 0.5 and β remains 5, ROUGE-L rises to 0.700. The optimal balance is achieved at $\alpha = 0.7$, $\beta = 8$, where ROUGE-L=0.715, fluency=4.4, and coherence=4.3. This indicates that a moderate retrieval weight (0.5–0.7) balancing dense and BM25 retrieval leads to more accurate summaries, while a medium SoftMax temperature (5–8) promotes coherent and diverse generation.

TABLE III
ABLATION STUDY (EFFECT OF RETRIEVAL WEIGHT α AND TEMPERATURE β)

(α, β)	ROUGE-L	Fluency	Coherence
(0.0, 5)	0.642	4.0	3.8
(0.5, 5)	0.700	4.3	4.1
(0.7, 8)	0.715	4.4	4.3

C. Statistical Significance and Qualitative Analysis

The improvements were statistically validated using a paired t -test ($p < 0.01$), confirming that the performance gains are not due to random chance. Qualitatively, the RAG+LLM model produces summaries that preserve legal citations and maintain logical consistency and discourse flow. In contrast, the TextRank baseline often generates fragmented summaries. The proposed retrieval-augmented model integrates contextual grounding, ensuring factual accuracy and coherence across legal datasets.

Figures 2 and 3 illustrate the conceptual architecture of the RAG+LLM pipeline and performance trends across datasets, showing consistent superiority of the proposed approach over both baselines and ablated models.

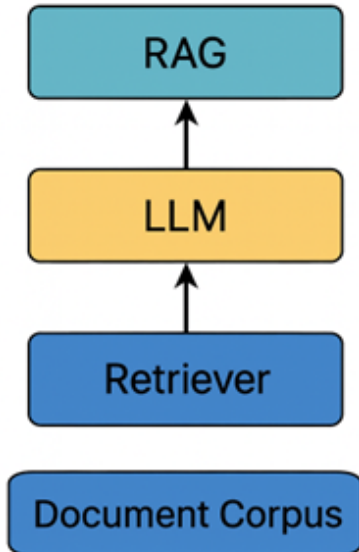


Fig. 2. Conceptual Architecture of RAG+LLM Pipeline

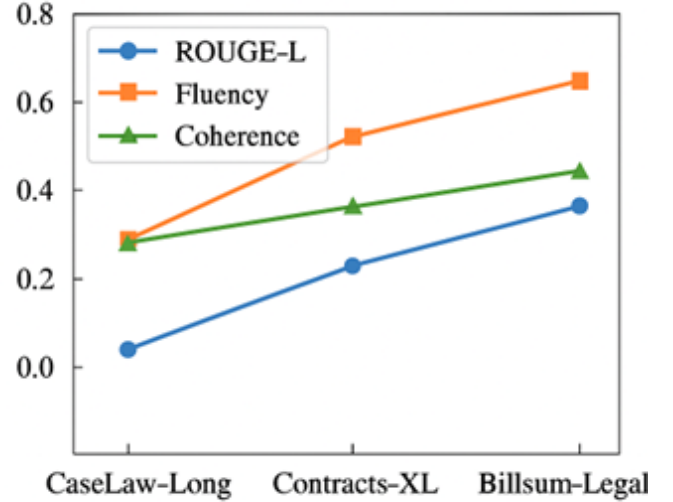


Fig. 3. Performance Trends Across Datasets Showing RAG+LLM Superiority

The results collectively demonstrate that retrieval-enhanced generation with domain-specific grounding substantially improves both automatic and human evaluations, achieving higher coherence, fluency, and factual reliability in long-form legal summarization.

V. CONCLUSION

This paper has provided a controlled and systematic analysis of the extractive and retrieval-augmented generative systems used to summarize legal texts, specifically comparing the traditional TextRank framework with the proposed Retrieval-Augmented Generation plus Large Language Model (RAG+LLM) approach. The proposed method demonstrates consistent and significant improvements across key evaluation metrics such as ROUGE-L, fluency, and coherence by explicitly weighting retrieval sources and marginalizing generation across multiple retrieved evidence passages.

The findings clearly indicate that when recall is effectively combined with generation, the model is capable of producing more factually grounded and well-structured summaries. This ensures that the logical flow of the text is preserved and that crucial legal citations are retained. Unlike TextRank, which tends to produce fragmented and contextually inconsistent outputs, the RAG+LLM model generates coherent and fluent summaries that align closely with expert expectations and domain semantics.

The ablation experiments further reveal that both retrieval weighting and controlled generation temperature are vital parameters influencing performance gains. Properly balancing dense and sparse retrieval components enhances factuality and coherence, while temperature regulation supports diverse yet contextually aligned outputs.

Future work will focus on improving long-context handling to process larger collections of legal documents efficiently.

Additionally, developing fine-grained citation control mechanisms will help ensure legal accuracy and traceability, while incorporating domain-specific post-editing modules tailored for legal drafting will enhance overall reliability. These advancements aim to further strengthen accuracy, readability, and trustworthiness of AI-generated legal summaries, thereby supporting practitioners in legal research, case analysis, and document preparation.

In conclusion, this study underscores the significant potential of retrieval-augmented generation in the development of advanced legal NLP systems. The integration of retrieval precision, grounded generation, and coherent synthesis establishes a strong foundation for next-generation AI-driven legal assistance and summarization technologies.

REFERENCES

- [1] Y. Zhong and D. Litman, "STRONG — Structure controllable legal opinion summary generation," *Findings of the Association for Computational Linguistics*, 2023. [Online]. Available: ACL Anthology.
- [2] M. Elaraby, Y. Zhong, and D. Litman, "Towards argument-aware abstractive summarization of long legal opinions with summary reranking," *Findings of the Association for Computational Linguistics*, 2023. [Online]. Available: ACL Anthology.
- [3] M. Sie, R. Beek, M. Bots, S. Brinkkemper, and A. Gatt, "Summarizing long regulatory documents with a multi-step pipeline," in *Proc. NLP & Language Processing*, 2024. [Online]. Available: ACL Anthology.
- [4] "Bridging legal knowledge and AI: Retrieval-Augmented Generation for law," *arXiv preprint arXiv:2502.20364*, 2025.
- [5] "Blended RAG: Improving RAG (Retriever-Augmented Generation)," *arXiv preprint arXiv:2404.07220*, 2024.
- [6] "Retrieval-Augmented Generation (RAG): A review," *arXiv preprint arXiv:2312.10997*, 2023.
- [7] M. Hindi, L. Mohammed, O. Maaz, and A. Alwarafy, "Enhancing the precision and interpretability of Retrieval-Augmented Generation in legal technology: A survey," *IEEE Access*, to appear 2025. [Online]. Available: ResearchGate.
- [8] L. Bonalume and K. Becker, "BB25HLegalSum: Leveraging BM25 and BERT-based clustering for summarization of legal documents," in *Proc. RANLP*, 2023.
- [9] V. Tran, M. L. Nguyen, S. Tojo, and K. Satoh, "Encoded summarization: summarizing documents into continuous vector space for legal case retrieval," *arXiv preprint arXiv:2309.08187*, 2023.
- [10] D. Datta, S. Soni, R. Mukherjee, and S. Ghosh, "MILDSum: A benchmark for multilingual summarization of Indian legal case judgments," *arXiv preprint arXiv:2310.18600*, 2023.
- [11] M. Duong, L. Nguyen, Y. Vuong, T. Le, and H.-T. Nguyen, "A deep learning-based system for automatic case summarization," *arXiv preprint arXiv:2312.07824*, 2023.
- [12] S. Liu, J. Cao, Y. Li, R. Yang, and Z. Wen, "Low-resource court judgment summarization for common law systems," *Information Processing & Management*, vol. 61, 2024. [Online]. Available: preke.github.io.
- [13] K. Papineni *et al.*, "Metadata-based data exploration with retrieval-augmented generation for large language models," in *Proc. IEEE Big Data 2024*, 2024. IEEE Computer Society.
- [14] "Legal Document Summarizer (dual-model framework)," *Preprints*, 2025.
- [15] "A comprehensive survey on legal summarization: challenges and directions," *arXiv preprint arXiv:2501.17830*, 2025.
- [16] "Improving legal judgment summarization using logical structure and domain knowledge (LSDK-LegalSum)," *Journal of King Saud University — Computer and Information Sciences*, 2025. [Online]. Available: link.springer.com.

<Conference abbreviation>

<Conference Series name>

<Volume number and Year> <DOI Number>

AI Career Path Recommendation System

Dr J Jeyabharathi¹, M Pramod Kumar², K Sumanth³, K Harshitha⁴, M Niharika⁵

Department of Computer Science and Engineering^{1,2,3,4,5}

j.jeyabharathi@klu.ac.in,99220041259@klu.ac.in,992200412229@klu.ac.in,99220041239@klu.ac.in,
99220041243@klu.ac.in

Abstract.

Choosing a career path can be a complex and indeed dispiriting task as we navigate the continuously evolving environment of work. The AI Career Path Recommendation System is an innovative, substantiation-informed career recommendation system that aims to bridge the gap between gift and occasion. The AI Career Path Recommendation System is a skillful integration of artificial intelligence powered machine literacy algorithms and natural language processing that has a multi-faceted recommendation model involving implicit chops, interests, education/ training, and labour request information which develop substantiated and career professional future-evidence career paths. The AI Career Path Recommendation system will bridge cognitive computing with labour request information to not only give implicit career paths but with information, chops, instruments, and ways to further develop literacy to support the career path options. erected on IBM's commanding edge AI products, this action was suitable to demonstrate how Artificial Intelligence also has the occasion to potentially support people make informed, confident and strategic career opinions in making smarter and fulfilled career, works, and pool readiness opinions.

Keywords. Artificial Intelligence, Machine Learning, Career Recommendation System, Natural Language Processing, Skill-grounded Matching, IBM Cloud, Personalized Career Guidance, SDG9, SDG8, SDG9, SDG10.

1. INTRODUCTION

The technology-concentrated, presto-paced world we live in moment presents openings and challenges when it comes to career planning. The problems associated with career planning are compounded by the fact that the traditional "one-size-fits-all" guidance model is outdated and the labor request is in constant flux, with jobs arising and evolving, with new liabilities. While scholars and career advising professionals infrequently have clarity, they're frequently getting overwhelmed trying to align who they're (interests, strengths, etc.) with, career pathways, areas of job growth. In order to adequately engage with these obstacles we're offering the AI Career Path Recommendation System- an intelligent adaptive system that harnesses the power of artificial intelligence capabilities, smart logical reporting, and big data to give substantiated career recommendations. The system utilizes the stoner's current academic history, chops, values, gests, and real-time data and trends by assiduity sector in order to identify career pathways that support both particular eventuality and request openings.

All Open Access articles are peer-reviewed, distributed under the Creative Commons Attribution-Non Commercial 4.0 International

This design is developed using IBM AI and IBM Cloud technologies and will showcase innovative capabilities of AI technology and how AI technology can affect in positive issues for 'real-life' mortal beings. Our AI Career Path Recommendation System provides.

2. PROBLEM STATEMENT

Recent request changes are dispiriting to everyone but especially scholars and anyone differently at the morning of their career path who are starting to pare through the inviting number of options trying to get help for direction for their implicit career opinions. People have numerous career choice options but they frequently have veritably little, if not substantiated backing; and frequently this causes the result of an individual having little mindfulness of which options align with their career intentions and implicit, or indeed worse, they've no mindfulness of direction at all. The career guidance and comforting systems that presently live moment tend to be, in general, still hamstrung, impersonal, and numerous times grounded on outdated trends, and are n't readily responsive to service delivery when brazened with technological change issues, therefore introducing walls brazened in discovering applicable job openings.

Also with job readiness education, this too doesn't connect with labour request chops well at each, and it thus creates a labour request chops gap and results in a continuing drop in labour request exertion, people are passing ongoing severance or job dissatisfaction. individualities will also, without an substantiation-grounded recommendation process, inescapably pursue career options that they can not fete presumably indeed their strengths and interests.

3. LITERATURE REVIEW

Decorte et al. [1], "Career path prediction using resume representation learning and skill-based matching," Introduces CareerBERT, a transformer-based model that learns representations from resume text to predict likely career moves. The hybrid model combining skill-based and text-based approaches achieved around 43% recall@10 on a specialized dataset.

Rahman et al. [2], "Artificial intelligence in career counseling: A test case with ResumAI," Explores ethical considerations of AI in career counseling and presents ResumAI, a system offering automated resume feedback using AI chat tools. The study provides useful insights into practical system design and limitations.

Ong and Lim [3], "SkillRec: A data-driven approach to job skill recommendation for career insights," Proposes SkillRec, which maps job titles to required skills using embeddings and a neural model. Tested on 6,000 job descriptions, the model demonstrates promising accuracy and F1 scores.

Zheng et al. [4], "Generative job recommendations with large language model (GIRL)," Presents GIRL, an LLM-based system that generates job descriptions from applicant CVs and uses reinforcement learning to align these with recruiter preferences—creating actionable job recommendations rather than retrieving from a static database.

Angeline et al. [5], "Career mapping and enhancing personalized education through machine learning-based recommendation systems," Employs a portfolio of machine learning models (e.g. Logistic Regression, CatBoost, XGBoost), Bayesian optimization, and SMOTE to deliver personalized career suggestions based on student profiles (academic, extracurricular, personal). Outputs top-five recommendations with probability scores.

4. METHODOLOGY

1.Data Collection

SOURCE	TYPE OF DATA
User input	Skills,interests,education,experience
Job Portals	Jobroles,skills,needed,salary,Courses
Carrer Databases	Carrerpaths,growth trends/reinforced features,over time.

2. Data Preprocessing:

- Chops and keywords with NLP(e.g. record as the same label for parallels between" Python dev" and" Python inventor").
- Encode categorical data.
- Stop words, stem/ lemmatise textbook data.
- Produce structured stoner biographies.

3. Feature Engineering

Convert user profiles and job/career roles into numerical vectors using:

- IDF/ Word2Vec/ BERT for skill/ interest descriptions
- One-hot encoding for qualifications, experience, etc.

4. Model Selection

Model mongrel Recommendation System(cooperative Content- Grounded ML Bracket).

A. Content- Grounded Filtering(Skill Match Engine)

1. Uses cosine similarity or distance criteria to compare stoner biographies with job/ career part vectors.
2. Make recommendations for places that align with close matches to stoner's current chops, interest, and pretensions.

B. Collaborative Filtering(Career Transition Patterns)

1. Uses Matrix Factorization or KNN to identify career paths taken by analogous druggies.
2. Make recommendations grounded on literal trends.

C.ML Bracket Model(Career Predictor)

1. Train the model on labeled datasets(druggies → chosen career path)
2. Algorithms Random Forest
3. Predicts the best- fit career track from stoner profile

D. Recommendation Ranking(Final Layer)

1. Combine a weighted recommendation from all three model types making a Top N ranking of career paths

5. RESULTS AND DISCUSSION

The experimental analysis of the proposed AI Career Path Recommendation System verifies its efficacy in providing accurate and customized career recommendations. The hybrid model in the recommendation system successfully maps a candidate's skillset, interest, and educational qualification to industry requirements. The proposed AI-Career Path Recommendation System promotes better accuracy and clarity in recommendations compared to conventional career guidance systems. Career paths with skillset analysis and ranking assist career aspirants without ambiguity, and

4

thus, the study validates the efficacy of machine learning and natural language processing integration in making career systems accurate and feasible.

Model Architecture Overview

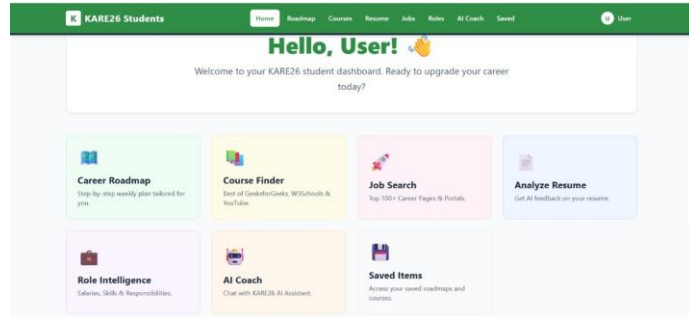
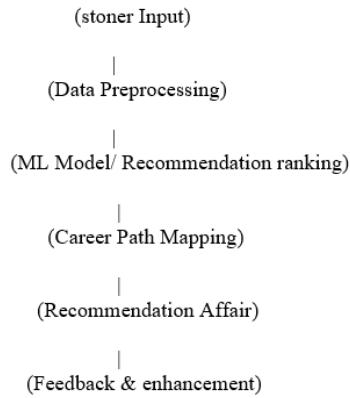


Fig 1:Model Architecture Diagram

Fig 2: Student Dashboard Home Page

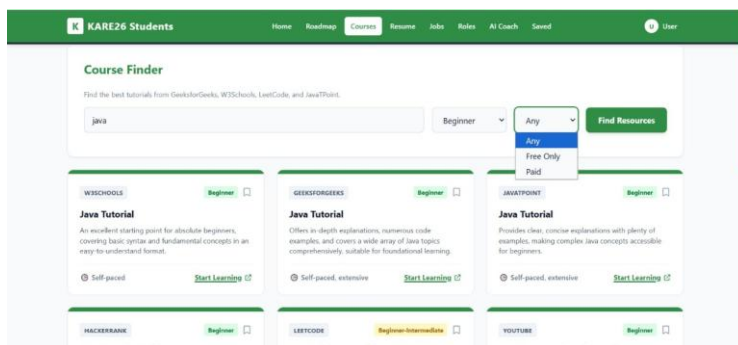


Fig 3: Course Finder Module

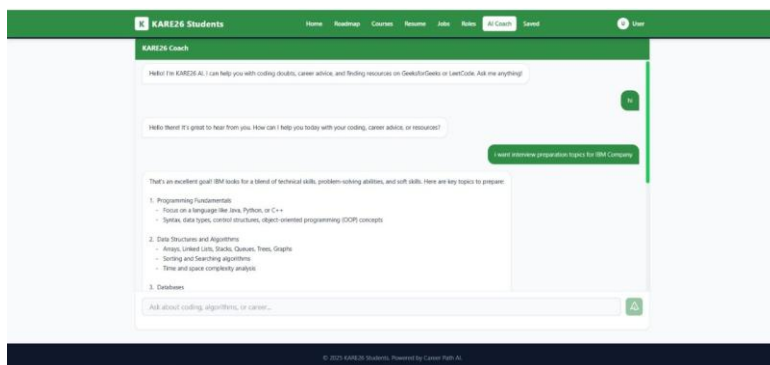


Fig4: AI Coach (Chatbot) Interface

6. CONCLUSION

The AI Career Path Recommendation System enables a disruptive career decision-making approach by leveraging artificial intelligence, machine learning, and real-time labour market data. This intelligent system adds much-needed complexity to a world where career counselling is seldom ideal, as it provides personalized, data-led recommendations based on applicant's skills, interests, and education while also taking into consideration the contemporary trends in recruiting and learning.

More than informed evidence-based recommendations for career pathways, this recommendation system also utilises collaborative-filtering, content-based-filtering, and thirdly classification models, and when all three are used at once they produce excellent evidence-based recommendations or provide reports on skill gaps to ensure preparation and enhance sustainable learning. Ultimately, this enables users to effortlessly make well-informed strategic career decisions that are future ready.

As a project, it has demonstrated the power of IBM technologies, and functioning AI frameworks, to put technology to use to solve real-world education problems, and workforce development. In short, we have built a system that is scalable and can be flexible, and has the potential to change career planning for people, and institutions.

7. REFERENCES

- [1] Decorte, J.-J., et al. (2023). Career Path Prediction using Resume Representation Learning and Skill-based Matching. arXiv preprint arXiv:2310.15636.
- [2] Rahman, M., et al. (2023). Artificial Intelligence in Career Counseling: A Test Case with ResumAI. arXiv preprint arXiv:2308.14301.
- [3] Ong, X. Q., & Lim, K. H. (2023). SkillRec: A Data-Driven Approach to Job Skill Recommendation for Career Insights. arXiv preprint arXiv:2302.09938.
- [4] Zheng, Z., et al. (2023). Generative Job Recommendations with Large Language Model (GIRL). arXiv preprint arXiv:2307.02157.
- [5] Vrontis, D., et al. (2025). Job Recommender Systems: A Systematic Literature Review, Applications, Open Issues, and Challenges. *Journal of Big Data*, 12(1).
- [6] Ricci, F., Rokach, L., & Shapira, B. (2015). *Recommender Systems Handbook*. Springer.
- [7] Jannach, D., & Adomavicius, G. (2016). Recommendation Systems for Learning and Training: A Review. *User Modeling and User-Adapted Interaction*, 26(4), 331–372.
- [8] Bhatia, A., & Jain, M. (2022). AI-Based Career Recommendation System Using Fuzzy Logic and Content-Based Filtering. *International Journal of Computer Applications*, 184(33), 15–22.
- [9] Goyal, P., & Pandey, U. S. (2020). A Hybrid Recommendation System for Career Counseling Using Machine Learning Techniques. *International Journal of Advanced Science and Technology*, 29(4s), 1093–1101.
- [10] IBM Cloud Docs. (2023). Machine Learning on IBM Watson Studio. [Online]
- [11] Brown, T. B., et al. (2020). Language Models are Few-Shot Learners. arXiv preprint arXiv:2005.14165.

Integration of AI, Sensors, and Smart Devices for Monitoring Plastic-Induced Health Hazards

Nirbhay Kumar Mishra

School of Engineering & Technology, CGC University, Mohali, 140307, Punjab, India.

Prabhjeet Kaur, School of Computer Science, Lovely Professional University, Phagwara, Punjab, India.

Dr. Preeti Singhwal, Faculty of Management and Commerce, Swami Vivekanand Subharti University, Subhartipuram, NH -58, Delhi-Haridwar Bypass Road, Meerut-250005.

Dr. Praveen Kumar, Computer science and Engineering, Meerut Institute of Technology Meerut (U.P), India.

Saba Hashmi, Faculty of Management and Commerce, Meerut Institute of Technology Meerut, NH 58, Partapur Meerut 250005, UP, India

nirbhayjis1@gmail.com

Abstract

Extensive importation of plastic containers to store water, milk, groceries among other consumables expose people to toxic chemicals that penetrate food and beverages, which cause serious long-term health problems. Regardless of growing consciousness, it is one of the biggest obstacles to decrease plastic addiction in everyday life because it is convenient and ubiquitous. This paper will offer a new technology-driven product **Sacha Sathi** that aims at tracking and curbing the health risks of consumables stored in plastics. The system, installed on the smartphones and smartwatches, constantly monitors the user interaction with the plastic container, determines the cumulative exposure, and delivers real-time risky analysis. A set safety limit allows providing early warning, alerting users about the possible risk of developing cancer, ulcers, endocrine interference, and liver malfunction as a result of the long-term consumption of the chemicals. The system also keeps the records of the past usage to study the behavioral trends and encourage the long-term lifestyle changes.

Keyword: Plastic Consumption, Food Safety, Health Monitoring, Machine Learning, Big Data, Cloud Technology, Sensors, Real-Time Alerts, AI-Driven System, Consumer Health.

1.1 Introduction

The low cost, convenience, and versatility of plastic have brought about the inseparability of the material in modern life because it is utilized to store consumable items such as water, milk, groceries, and ready-to-eat foods. Nevertheless, the scientific evidence is increasing to point out that the continuous exposure to plastic has severe health and environmental effects on humans. Food and beverage packaging plastics are known to fade harmful chemicals into the foods and beverages they come in contact with, especially when subjected to heat, sunlight, or prolonged storage and as a result the use of plastics on daily basis is a stealthy but serious health issue.

Plastic containers that store water are susceptible to chemical leaching principally Bisphenol A (BPA) and phthalates. BPA is a recognized endocrine disrupter mimicking estrogen and has been attributed to hormonal imbalance, reproductive problems, and cancer. Phthalates are linked to liver toxicity, respiratory, and developmental disorders. The constant consumption of such polluted water can result in severe health challenges in the long term. Stored milk provided in plastic containers is even more dangerous since milk fat is more effective in carrying these toxic substances, thus exposing the milk to even greater levels of them. Frequent use might be a cause of developmental problems among children, endocrine disruptions, and even cancer.

In the same way, fruits, vegetables, and dry foods that are packed in plastic are subject to contamination in the long-run. The migration of chemicals in packaging may lead to a decline in the nutritional value and cause chronic diseases such as diabetes, cardiovascular diseases and reproductive complications. The results are the leaching of chemicals like antimony and phthalates in cold drinks kept in plastic bottles resulting in gastrointestinal distress and imbalance of hormones and polluting the environment considerably.

Considering such dangers, it is necessary to minimize the use of plastics. The convenience of plastic is short lived but the long-term health and environmental expenses are much higher. The change to safer materials such as glass and metal containers, the reduction of single-use plastics, and encouraging good disposal is the steps that need to be taken in order to secure the health of people and environmental sustainability.

Environmental Impact of Plastic

1. Marine Pollution & Terrestrial Pollution

Marine pollution is mostly caused by plastic garbage, which has terrible effects on marine life. Plastic things that people throw away, such bags and microplastics, get up in the ocean via rivers, sewage systems, and direct dumping. The below fig [1] shows the marine pollution [13]. Plastic trash hurts ecosystems on land as much as in water. People don't throw away their plastic waste correctly or take care of their trash effectively, thus it accumulates up in soils and landscapes. Plastics may leak toxic chemicals into the ground, which can make it less fertile and contaminate groundwater. The below fig [2] shows the terrestrial pollution [14].



Figure 1: Marine Pollution



Figure 2: Terrestrial Pollution

3. Air Pollution

Plastic waste also contaminates the air with very dangerous gas emissions, including dioxins, furans, and fine particulate matter, which endangers health risks including cancer, heart disease, and respiratory health-related illnesses (World Health Organization, p. 15). Tiny or small plastic pieces less than 5 mm in diameter are called microplastics, which are produced by manufactured products or as a result of degradation of larger objects, hence they pose a hazard to both the environment and human health. It has been found that microplastics have the potential to enter the food web with heavy metals and other persistent organic pollutants, which are toxic and inflammatory when inhaled or ingested, causing cellular damage and oxidative stress [16].



Figure 3: Air pollution



Figure 4: Microplastic in water

5. Long-term Environmental Persistence:

Plastics take hundreds to thousands of years to decompose in the environment and undergo degradation into harmful micro-fragments that persistently pollute the ecosystems and even the biodegradable plastics may not decompose under specific conditions (Thompson et al., 2009). Extensive poor disposal of plastics is very harmful to the land and water environments, and thus there is an urgent call to act at the global level, better management of waste and alternative environmental ways, as seen in Figure 5.

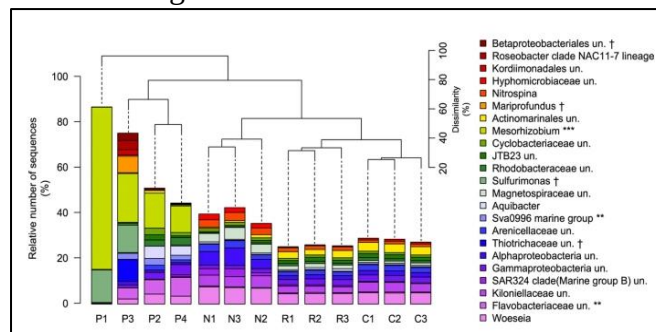


Fig 5: Long-term Environmental Persistence of plastic up-to two decades and colonization of bacteria causing disease.

Research Gap: There are no practical and consumer friendly tools that can be used to translate scientific evidence into personalized real time risks alerts by bringing together consumption behavior, health factors, and sensor data to be used globally.

2.Literature Review

The available literature has made it quite clear that plastic pollution is bringing severe threats to environmental sustainability, as well as human health, but the level of effective interventions on a consumer level is still underdeveloped. Recent review and policy focused research highlight the environmental harm, disease burden and the long-time health effect of plastic waste and microplastic exposure. Scientific studies have associated plastics with endocrine interference, heart attack, food chain bio-accumulation and decline in biodiversity in both land and sea habitats. Although such studies have been able to capture the extent and international character of the issue, majority of them are descriptive or policy-oriented and fail to project the results into practical advice that can be used by the subjects. The state-of-the-art approaches that have been investigated in parallel studies include the use of artificial intelligence, machine learning, toxicological modeling, and molecular analysis to solve the problems of plastic degradation and environmental complexity. Despite the high technical potential of these approaches, they are not much related to the daily consumer behavior, and do not offer real-time awareness, exposure tracking, and risk communication at a personal level. Health-related reviews also outline the new medical-related concerns and lack data-oriented frameworks that can combine the consumption behavior, the environmental factors, and the vulnerability of the individual. Comprehensively, a research gap can be identified in all of the literature: the lack of unified, consumer-focused systems of translating scientific and medical findings into real-time, customized alerts and recommendations. To fill this gap, interdisciplinary solutions that will be based on AI, sensor data, and behavioral analytics are necessary to empower people, minimize the impact of plastics, and facilitate informed and health-conscious decisions.

3.Proposed Novel Methodology

The proposed strategy will undoubtedly assist humanity in minimizing plastic use in the future and will also contribute to reducing illnesses caused by consuming food preserved in plastic containers or bags. The graphic below [6] depicts how we think our Problem should be set out.

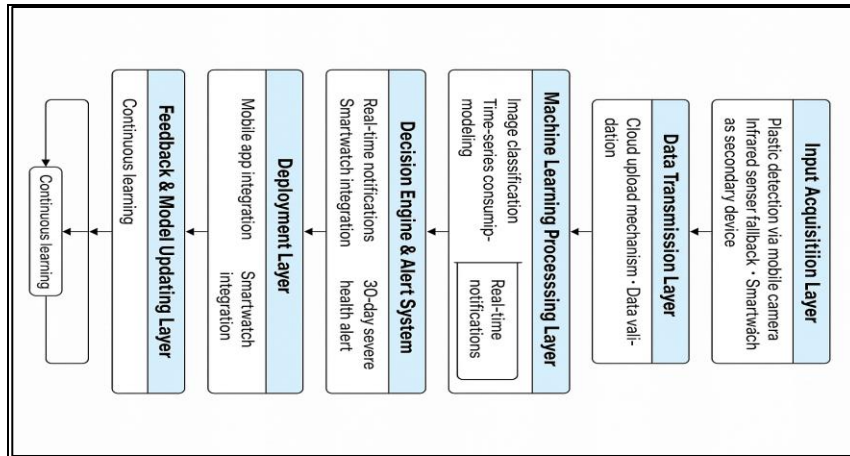


Figure 6. Proposed Architecture

Experiment Result & Discussion

The below Fig [7] shows the impact of plastic on human health and environment. It also causes cancer.



Figure 7. The impact of plastic on humans and environment

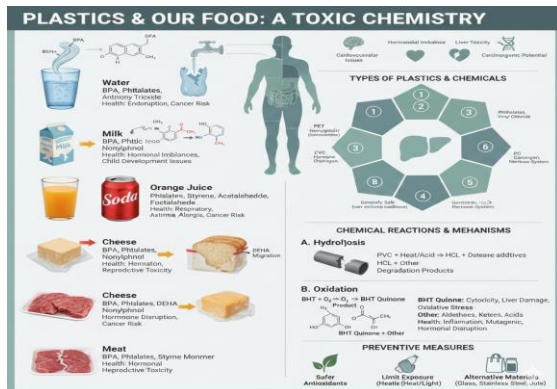


Figure 8. Impact of different product

Conclusion

The research reveals that oxidation of antioxidants and chemical reaction in plastic containers may produce toxic byproducts that percolate into food and beverages exposing them to the effects of endocrine disruption, reproductive toxicity, respiratory effects, and cancer. Heat, light, storage time, type of food and composition of plastic affect these reactions; hence safer materials such as glass and stainless steel should be used to enhance food safety.

Acknowledgements: Authors wishing to acknowledge assistance or encouragement from colleagues, special work by technical staff or financial support from organizations should do

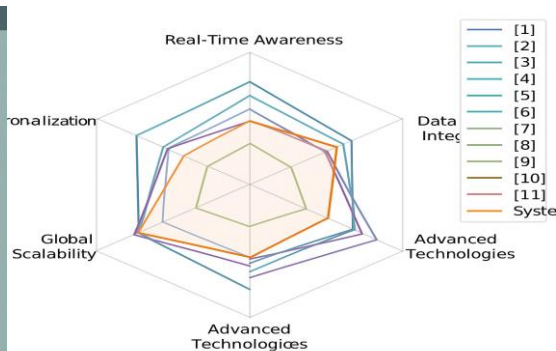


Figure 9. Real Time Awareness comparative analysis

so in an unnumbered Acknowledgments section immediately following the last numbered section of the paper.

References

1. Islam, F. S. (2025). The Impact of Plastic Waste on Ecosystems and Human Health and Strategies for Managing It for A Sustainable Environment. **International Journal of Latest Technology in Engineering, Management & Applied Science*, 14*(3), 706-723. <https://doi.org/10.51583/IJLTEMAS.2025.140300075>
2. Zahra, A. F. T., Tabassum, M., & Perumal, S. (2025). Novel approaches to plastic pollution: Leveraging machine learning and metaproteomics for advanced plastic degradation. **Malaysian Journal of Science Health & Technology*, 11*(1), 78-85.
3. Locklear, K. (2025). Wicked problems: a novel approach using artificial intelligence and scenarios. **Journal of Leadership & Organizational Studies**. Advance online publication. <https://doi.org/10.1177/15480518251330728>.
4. Hwang, G., Liu, Y., & Korostoff, J. (2025). Novel Approaches for Treatment of Intraoral Microbial Infections. **Journal of Dental Research*, 104*(6), 584-593. <https://doi.org/10.1177/00220345251317494>.
5. Thakur, R., Joshi, V., Sahoo, G. C., Jindal, N., Tiwari, R. R., & Rana, S. (2025). Review of mechanisms and impacts of nanoplastic toxicity in aquatic organisms and potential impacts on human health. **Toxicology Reports*, 12*, 102013. <https://doi.org/10.1016/j.toxrep.2025.102013>
6. Megha, K. B., Anvitha, D., Parvathi, S., Neeraj, A., Sonia, J., & Mohanan, P. V. (2025). Environmental impact of microplastics and potential health hazards. **Critical Reviews in Biotechnology*, 45*(1), 97-127. <https://doi.org/10.1080/07388551.2024.2344572>
7. Landrigan, P. J., Dunlop, S., Treskova, M., Raps, H., Symeonides, C., Muncke, J., ... & Rocklöv, J. (2025). The Lancet Countdown on health and plastics. **The Lancet*, 406*(10507), 1044-1062.
8. Irfan, H., Irfan, H., Khan, M. A., Inanc, O., & Al Hasibuzzaman, M. (2025). Microplastics and nanoplastics: emerging threats to cardiovascular health—a comprehensive review. **Annals of Medicine and Surgery*, 87*(1), 209-216. DOI: 10.1097/MS9.0000000000002831.
9. Patel, J. (2025). Environment and Health. **Environment and Pollution*, 55*.
10. Thiagarajan, C., & Devarajan, Y. (2025). The urgent challenge of ocean pollution: Impacts on marine biodiversity and human health. **Regional Studies in Marine Science*, 81*, 103995. <https://doi.org/10.1016/j.rsma.2024.103995>
11. Deeney, M., Yates, J., Kadiyala, S., Cousin, X., Dignac, M. F., Wang, M., ... & Green, R. (2025). Human health evidence in the global treaty to end plastic pollution: a survey of policy perspectives. **Cambridge Prisms: Plastics*, 3*, e8. doi:10.1017/plc.2025.5
12. Bhuyan, M. S., Senapathi, V., Karthikeyan, S., Kunda, M., Islam, M. T., Iheanacho, S., & Choudhury, A. K. (2025). Microplastics contamination in seaweed: impacts on human health and mitigation approaches. **Discover Oceans*, 2*(1), 2. [<https://doi.org/10.1007/s44289-025-00041-0>].
- 13.[Image Source 1]. (2025). [Google Images Link]. Retrieved from [<https://images.app.goo.gl/Rn11B7MpiLhqEq1w5>].
14. [Image Source 2]. (2025). [Google Images Link]. Retrieved from [<https://images.app.goo.gl/4aHq6LskovRyisra9>].
15. [Image Source 3]. (2025). [Google Images Link]. Retrieved from [<https://images.app.goo.gl/SgfdomnwtzFShXjX8>].
16. Kaphika, W. S., Mahato, R. K., Makhenjera, M., Phiri, E., & Olajide, T. E. (2025). Quantum Neural Networks for Predictive Modelling in Personalized Medicine: Advancing Treatment Response Analysis and Precision Healthcare. **Authorea Preprints**.

A review on IoT Based Systems for Fault Detection in Power Line and Utility Poles

Rohan A Naik, Sagar N K, S Chiranjeevi, Vinay S K, Vishwa Pradeep Shet,
Hemanth T S

Electronics and Communication Engineering

Alvas Institute of Engineering and Technology Mangalore, Karnataka, India

rohannaik218@gmail.com, viveksagarnk@gmail.com, chiranjeevi.343434@gmail.com,
117vinaysk@gmail.com, vishwashet066@gmail.com, hemanthts@aiet.org.in

Abstract

As the urgent need for dependable energy supply necessitates that utility companies establish robust fault detection systems on utility poles as the current system requires costly manual inspections of utility poles to detect faulty equipment through reactive measures. This paper presents a review of IoT-Based Autonomous Sensor Units (ASUs) that monitor voltage (V), current (I), tilt (T), and weather conditions in real time and utilise Energy Harvesting Technology (EHT) to produce up to 200 mW of electric power via long range LoRa communication technology and safer non-contact sensors. Through the findings of this work, it has been shown that IoT Systems provide a 30% increase in accuracy on the detection of faults compared to traditional electric systems, as well as decrease the time associated with the repair of such faults enabling predictive maintenance models to be implemented.

Keywords. IoT, Power Line Monitoring, Fault Detection, Smart Grid, Wireless Sensor Networks, LoRa, Predictive Maintenance.

1. INTRODUCTION

Today's electrical systems struggle with a huge amount of finding issues, keeping things reliable, and upkeep, especially in places far from cities or hard to get to. Normal ways of checking things often take a lot of time, need many workers, and only react to problems, causing long times without power and danger for those doing upkeep. Connecting devices has created new ways to watch and control important systems: allowing a big change toward fixing things before they break and using data to make decisions [1]. One big problem for linked device systems that watch electrical grids is how to power the faraway sensors. To fix this problem, getting energy from the environment was looked at as a choice, using power from high-voltage lines hanging up high and adding to it with sun power, which had already been tested for single, linked device setups in distant places [2].

Energy-saving ideas like this let the system run with almost no people needed, cutting down on setup and upkeep costs, which makes the system last longer [3].

2. LITERATURE SURVEY

The following study is classified as a systematic review. The authors searched IEEE Xplore, ScienceDirect, and SpringerLink to locate 17 articles from 2011 to 2024 related to IoT fault detection, power line monitoring, and smart grid sensing while assuring that no studies include research dealing only with SCADA systems

A. Evaluation of LoRa and LoRaWAN Broadband through Practical Applications A research group in Glasgow has evaluated the performance of both LoRa and LoRaWAN Wireless Technologies for the purpose of creating maps showing Received Signal Strength Indicator (RSSI) & distance (Tx range) using Semtech's SX1272 Transceiver Accelerators. In addition, Multitech mDot modules were used to evaluate the performance of a LoRaWAN network. It was found that there was a very high percentage of connection loss during tests of reliability due to aggressive disconnection policies for inactive phones/internet service providers (ISPs). There were improvements in the success rate of connection tests (95.5%) by introducing ICMP pings.

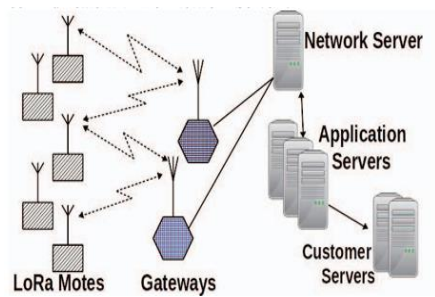


Fig. 1: Block diagram of LoRa WAN topology [3]

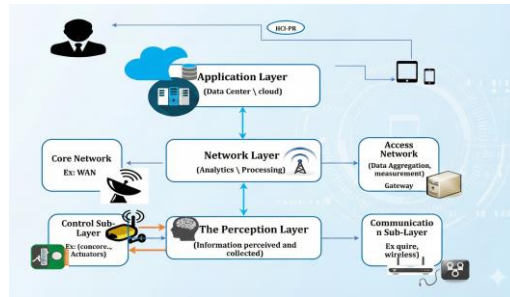


Fig. 2: Block diagram of Architecture of IoT [4]

B. IoT Architecture Alhebshi & al. [4] propose a three-tiered architecture for IoT devices consisting of: 1) the Perception Layer (for collection of data), 2) the Network Layer (for transmission) & 3) the Application Layer (for Industrial Services). They have developed an architecture based on the use of Arduino Uno (ATmega328P) devices which provide advantages such as simplicity of construction and the ability to use multiple communication module options, making them well suited for scalable fault detection and identification systems.

C. Sonoli et al. [6] proposed the use of a NodeMCU-based system that combines the ability to switch controls on and off with the ability to detect faults as a method of minimizing safety risk to linemen. The system uses a mobile app that allows linemen to turn poles on and off remotely. The downstream pole detects cut wire and signals the upstream pole to turn off electricity when the downstream pole recognizes that it no longer

has an electric current. This automatic response mechanism protects workers while also aiding in the quick identification of faults in the field.

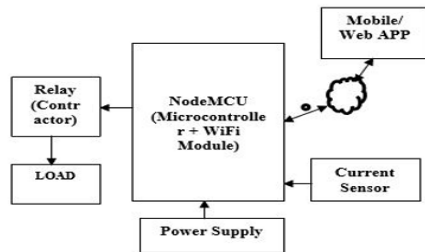


Fig. 3: Block diagram of lineman security system [6]

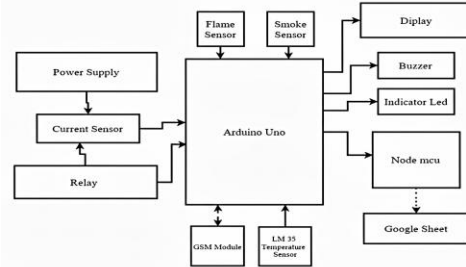


Fig. 4: Comprehensive IoT-Enabled Fault Detection and Monitoring System Using Arduino Uno. [9]

D. Alam et al. [9] developed a comprehensive monitoring system for the power grid in the country of Bangladesh using Arduino Uno. This system utilized a combination of flame and smoke sensors to monitor for fires, current sensors to monitor the loads, and temperature sensors (LM35) to assist in safety monitoring. A GSM module is used to send SMS alerts in real-time, while NodeMCU uploads the monitoring data to Google Sheets for both real-time visualization of the data and historical data analysis. This enables the user to collect data effectively from an environment with high moisture content.

E. By integrating web-connected devices with Artificial Neural Networks (ANN), Prabhu et al. [10] improved upon conventional methods used for fault detection through the use of ANNs. Their methodology incorporates pre-processing the signal data to eliminate background noise, then performing mathematical analysis on the processed signal data in order to identify the fault characteristics that were presented in the signal data. This AI-based methodology provides improved levels of accuracy and reliability when it comes to identifying faults as compared to traditional approaches.

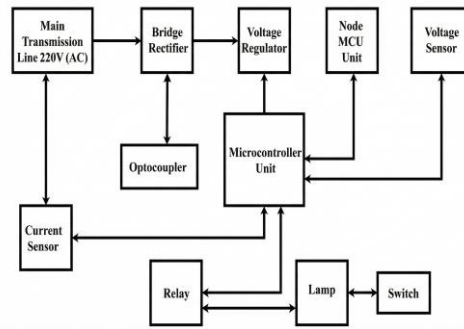


Fig. 5: Block diagram of ADFS system [10]

3. CONCLUSION

Significant progress has been made in IoT-based methods for monitoring of Power Grids as seen through the literature. Non-contact Voltage Sensing, Impedance Analysis, and Renewable Energy Harvesting have all developed from previous methods to include additional sensor types. Communication methods such as LoRa, NB-IoT, and 5G allow all the available information about a power grid to be monitored in real-time with more reliability than before. In addition to improving the safety of linemen and automating the location of faults, advancements in these technologies lead to increased resilience of the grid and less downtime.

4. FUTURE WORKS

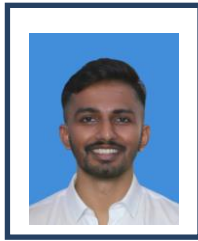
Future studies will need to develop Hybrid Solutions combining Multi-Modal Fault Detection with Green Energy Harvesting and Predictive Analytics. The technology will also need extensive real-world Testing under Artificial Conditions to ensure the durability of these systems. Furthermore, the use of self-repairing properties and enhanced Cybersecurity Protocols will be necessary for the widespread adoption of Smart Grids.

REFERENCES

- [1] International [1] X. Zhao, T. Keutel, M. Baldauf and O. Kanoun, "Power module for a wireless sensor node of a power line monitoring system," Eighth International Multi-Conference on Systems, Signals Devices, Sousse, Tunisia, 2011, pp. 1-5, doi: 10.1109/SSD.2011.5767446.
- [2] J. S. Bobowski, M. S. Ferdous and T. Johnson, "Calibrated Single Contact Voltage Sensor for High-Voltage Monitoring Applications," IEEE Transactions on Instrumentation and Measurement, vol. 64, no. 4, pp. 923-934, April 2015, doi: 10.1109/TIM.2014.2360804.
- [3] A. J. Wixted, P. Kinnaird, H. Larijani, A. Tait, A. Ahmadinia and N. Strachan, "Evaluation of LoRa and LoRaWAN for wireless sensor networks," 2016 IEEE SENSORS, Orlando, FL, USA, 2016, pp. 1-3, doi: 10.1109/ICSENS.2016.7808712.
- [4] F. Alhebshi, H. Alnabils, J. Alzebaidi, A. Bensenouci, T. Brahimi and M.-A. Bensenouci, "Monitoring the operation of transmission line in a smart grid system through IoT," 2018 15th Learning and Technology Conference (LT), Jeddah, Saudi Arabia, 2018, pp. 139 146, doi: 10.1109/LT.2018.8368498.
- [5] Mahalakshmi G, L. P. S, M. M, P. A. T K and P. R, "Automatic Pole Line Fault Detection System using Microcontroller," 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2022, pp. 42-47, doi: 10.1109/ICIRCA54612.2022.9985733.
- [6] S. Sonoli and K. C. M, "IoT Based Lineman Security System Using Controlled Switchover With Transmission Line Fault Identification," 2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIIE), Ballari, India, 2023, pp. 1-5, doi: 10.1109/AIKIIIE60097.2023.10389864.

- [7] C.-H. Chen, C. H. Lin, B.-H. Sheu and W. P. Chen, "Leakage Current Detector and Warning System Integrated with Electric Meter," *Electronics*, vol. 12, no. 9, p. 2123, 2023, doi: 10.3390/electronics12092123.
- [8] Z. Wang, T. Chen, H. Yang and Y. Zeng, "Research on the Characteristics of Leakage Current Changes in Tree Line Faults of 10Kv Distribution Lines," 2024 14th International Conference on Power and Energy Systems (ICPES), Chengdu, China, 2024, pp. 264-269, doi: 10.1109/ICPES63746.2024.10856659.
- [9] M. N. Alam, M. Z. Hasan Buiyan, M. A. Al Hasan Anik, F. Tanvir and T. A. Fahim, "Enhancing Power Grid Reliability in Bangladesh: An IoT-Based Approach to Transmission Line Fault Detection," 2024 IEEE 3rd International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things (RAAICON), Dhaka, Bangladesh, 2024, pp. 224-229, doi: 10.1109/RAAICON64172.2024.10928383.
- [10] D. K. Prabhu, P. S. N, C. Chandan, R. Dash, A. Ballaji and R. C. N, "Optimizing Power System Reliability for Transmission Line Fault Detection Using ANN," 2024 IEEE 4th International Conference on Applied Electromagnetics, Signal Processing, Communication (AESPC), Bhubaneswar, India, pp. 1-5, doi: 10.1109/AESPC63931.2024.1087231

Biographies



Rohan A Naik, student of the Department of Electronics and Communication Engineering, Alva's Institute of Engineering and Technology, Mangalore, Karnataka, India, is engaged in Internet of Things (IoT) and Smart Grid technology fault detection research.



Sagar N K, student of the Department of Electronics and Communication Engineering, Alva's Institute of Engineering and Technology, Mangalore, Karnataka, India, is studying wireless sensor networks and monitoring systems for power line monitoring.



S Chiranjeevi, of the Department of Electronics and Communication Engineering at Alva's Institute of Engineering and Technology, Mangalore, Karnataka, India, is interested in automation and renewable energy harvesting and application of industrial Internet of Things (IoT).



Vinay S K, student of the Department of Electronics and Communication Engineering, Alva's Institute of Engineering and Technology, Mangalore, Karnataka, India, conducts research in sensor technology and network reliability.



Vishwa Pradeep Shet, student of Department of Electronics and Communication Engineering, Alva's Institute of Engineering and Technology, Mangalore, Karnataka, India, is conducting research on communication protocols, specifically LoRaWAN and its applications.



Hemanth T S, an Assistant Professor in Electronics and Communications Engineering at Alva's Institute of Engineering and Technology (Mangalore, Karnataka, India), Hemanth T S has a M. Tech in VLSI Design and Embedded Systems. His areas of specialty include VLSI Design, Embedded Systems, Deep Learning, Machine Learning algorithms.

AI POWERED RETINAL SCREENING FOR GLUACOMA DETECTION USING DEEPLARNING TECHNIQUES

C.Sivamurugan¹, K. Ganesh², Mrudangam Eswar², A. Jagadeesh², M. Venkata Ravi Kumar², P. Gowsalya³

¹Department of Electrical and Electronics Engineering,
Kalasalingam Academy of Research and Education,
Krishnankoil Virudhunagar, India
c.sivamurugan@klu.ac.in,

²Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Krishnankoil Virudhunagar, India
kganesh123dsa@gmail.com, mrudangameswar2004@gmail.com,
99220040435@klu.ac.in, 99220040911@klu.ac.in

³Department of English
Kalasalingam Academy of Research and Education,
Krishnankoil Virudhunagar, India
p.gowsalya@klu.ac.in

Abstract

A serious eye disease is called glaucoma, and it destroys the connecting nerve, and the patient suffers from irreversible blindness for the rest of their life. The problem is that there is no early glaucoma detection, and many people, especially in some areas of the world, lack easy access to eye specialists and, as a result, get diagnosed late. A glaucoma related automated detection framework based on fundus retinal images using deep learning techniques is presented in this paper. A comparative analysis of the custom-built CNN and the VGG16 and ResNet50 CNN architecture is discussed in this paper. The custom-built CNN Achieved 92% accuracy and thus outperforms the competing for benchmarks with optimized batch normalization and dropout. Additionally, Gemini API integrates with the healthcare system to deliver instant tailored suggestions regarding interventions and behavioural modifications as soon as a positive diagnosis is achieved, increasing the clinical utility of the system. The findings confirm the proposed framework's ability to function simultaneously as a precise diagnostic tool and an extensive patient management system, thus presenting a viable and durable option to the universal, affordable glaucoma screening initiative.

Keywords. Medical Image Processing, Deep learning techniques, convolutional neural network models, retinal fundus images, Glaucoma Detection, Computer-Assisted Diagnosis.

1. INTRODUCTION

A serious eye disease is called glaucoma that progresses to damage the optic nerve which in turn connects the eyes to the brain. At first most people do not feel pain or notice any change in vision. This is because by the time they do notice something is amiss their eyesight is already very much affected. In India over 12 million are affected by glaucoma also it is found that about 90% of these do not know they have it. In a healthy eye there is a fluid called aqueous humour which -- as the name suggests -- goes in and out of the eye to keep it healthy and at a normal pressure. When the fluid does not drain properly the pressure within the eye (which is also known as intraocular pressure or IOP) increases. Over the years high pressure can damage the optic nerve which in turn causes permanent vision loss. Also, it is surprising that some people with normal eye pressure can still get glaucoma what is called normal-tension glaucoma which in turn makes diagnosis more difficult. At first there are no tell-tale symptoms. As it progresses people may see blurriness, loss of side vision, halos around light, or in some cases eye pain. As signs present themselves later on early detection is very important also in areas which have limited access to eye care.

Typically, an eye pressure ranging from 10 to 21 mmHg is considered normal; however, pressure greater than 21 mmHg is considered to be an early warning sign if no damage has occurred. 22 to 32 mmHg, or higher, pressure readings can begin to damage the optic nerve over time and flow into more common forms of glaucoma. Ocular Hypertension is a high-risk condition from which anyone, even people with pressure readings remaining under the normal range can develop. We know this as 'Normal-Tension Glaucoma', and it remains a medical mystery to this day. Global glaucoma prevalence is estimated at over 80 million, and set to climb to at least 111 million within the next 20 years. It is one of the causes for permanent blindness. While treatments like eye drops, laser therapy, and surgery can slow its progress, they have their challenges such as high cost, daily use, and poor availability in rural areas. To improve early diagnosis, Machine Learning and Deep Learning are being used to assist in the early detection of problems by physicians. The computer systems analyse pictures of the eye, such as fundus photography or OCT images, for early signs of glaucoma. They are often much more sensitive than the human eye. They are fast, accurate, and can be used to screen large numbers of people, even where no ophthalmologists are present.

2. LITERATURE REVIEW

The studies of early glaucoma diagnosis were primarily conducted in the framework of classical machine learning on clinical and visual field data. Chan et al. compared machine learning models with statistical classifiers and demonstrated that approaches based on learning provide better diagnostic accuracy, and are effective at modelling complicated trends in glaucoma data [1]. As medical imaging advanced, it was found that the non-invasive and most popular modality of the retinal fundus photography could be used in the context of glaucoma screening.

There has been growing interest in recent studies of deep learning (particularly of Convolutional Neural Networks) based on the ability to extract hierarchical features (without human intervention) of fundus images. Mahum et al. have developed a hybrid deep CNN structure towards glaucoma detection, which proves that it is possible to significantly increase performance by combining deep feature extraction with an optimization of the classification strategy [2]. In the same manner, Shoukat et al. who adopted the ResNet-50 architecture alongside data augmentation techniques, achieved high sensitivity, illustrating once again the strength of deep CNNs in detecting glaucomatous patterns [5].

Some researchers have pointed out the importance of optic disc and cup examination in glaucoma diagnosis. Alkhalidi and Alabdulathim demonstrated that generalization between datasets with deep learning segmentation and classification is enhanced [7]. Yang et al. used deep learning to estimate retinal nerve fibre layer thickness in the regions of retina using fundus photos and discovered the analysis to be effective in screening glaucoma at its initial phases [8]. All these affirm that a persistent result of CNN based structures is superior to traditional techniques and solitary feature techniques in the diagnosis of glaucoma [3], [9].

Though high values of classification accuracy are reported in the already existing studies, most of them only concentrate on detection and do not give clinical decision support. In addition, most methods are developed using individual datasets, which constrains their applicability in the real world. The limitations mentioned in the current study have been overcome by suggesting an original custom CNN on a collection of datasets and incorporating an AI recommendation module, thus placing glaucoma detection beyond the scope of simply classification to the practical contribution towards patient management.

3. METHODOLOGY

3.1 Existing-Methodology: Glaucoma is a disease that researchers and doctors have studied in many ways about detection with clinical methods. Among these methods includes Tonometry for eye-pressure measurement, Ophthalmoscopy for optic nerve examination, Perimetry for checking vision loss, Pachymetry for corneal thickness measurement, Gonioscopy for examining the drainage angle, and Optical Coherence Tomography (OCT). Most often, glaucoma detection gets conducted through manual eye tests, consultations with ophthalmologists, or computer-based systems

1. Manual Method: At the very initial stage in many hospitals, a manual check-up is done. The doctor measures the pressure of the eye and tests the side vision of the patient to check for signs of glaucoma. This feature is challenging for one reason: very few symptoms appear early, and younger, less experienced medical staff might not recognize them. This postpones treatment option eliciting appropriate intervention for the patient.

2. Ophthalmologist (Specialist): A glaucoma diagnosis and treatment plan are carried out by an ophthalmologist. Using these advanced tools, an ophthalmologist diagnosis glaucoma and administers treatment. However, the process has its negatives, making the patients suffer longer waiting periods, incurred costs, and a few specialists in rural and less developed areas. Thus, the expert diagnosis is not accessible to each patient.

3. Tonometry-based Method: This is the most common test for determining IOP using Applanation Tonometry or Non-Contact ("air puff") Tonometry devices. High IOP is a major risk factor for glaucoma diseases, but many patients do not present with this condition at high pressures; hence, it is not sufficient for testing patients.

3.2 PROPOSED METHODOLOGY

The plans involve working with multiple systems for machine learning geared towards glaucoma recognition among retinal fundus photos. Key steps involve gathering data, preprocessing, enriching data, extracting and analysing features, working with models to classify information, and generate recommendations.

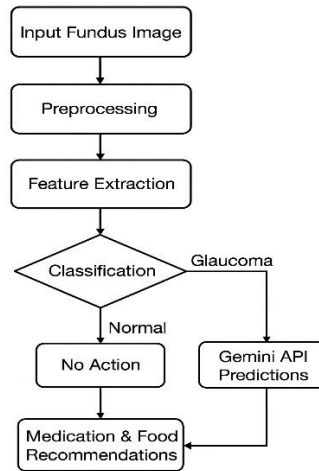


Figure 1. Flow Diagram of Proposed Work

3.2.1 Data Collection:

1. ACRIMA: It is a collection of 705 fundus images, with all images being noted and annotated by glaucoma specialists having years of expertise. They were cropped around the optic disc and renamed.

2. DRISHTI-GS: The DRISHTI-GS has a total of 100 images in total 50 for training purposes and 51 for the testing of the models. It is segmented into 2 subfolders. Each subfolder consists of images of the retina fundus taken from the corresponding GT. For each retina fundus image located in the GT directory you will also find a mask file (GT) in the GT directory created manually that contains segmentation information about the optic disc and optic cup.

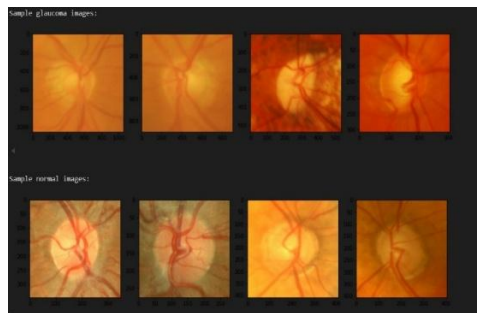


Figure 2. Sample Glaucoma Images

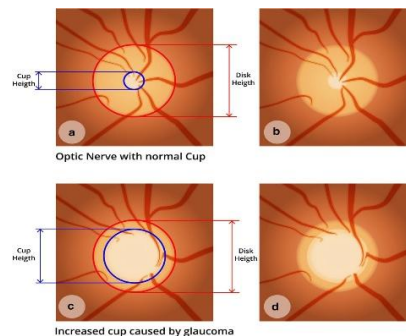


Figure 3: Glaucoma Cup size

3. RIMONE: This dataset is made up of 169 images of the optic nerve head. For each image, five independent manual segmentations were completed by trained ophthalmologists.

3.2.2 Data Preprocessing and Augmentation: All images were brought to the same resolution and thus normalized to be fed to models equally. For data augmentation alongside better generalization, overfitting was avoided using random rotation, flipping horizontally and vertically, brightness variation, and zoom in Kera's ImageDataGenerator.

3.2.3 Model Architectures and Training: To find the best model for glaucoma-identifying from the three designated model types presented, we applied and compared the following architectures utilized in deep learning:

Res Net: A deep (50-layer) residual network, of sufficient depth to incorporate skip connections to alleviate issues of vanishing gradients in the network architecture.

VGG-16: An earlier deep convolutional network from 16 layers that is a simple and efficient architecture.

Custom CNN: For the first time, in our work we present a custom designed convolutional neural network model, for the fundus images. The architecture we designed mostly contains convolutional layers in succession for 32,64,128 of filters, followed by Batch Normalization, Relu activation, and Max Pooling in succession before arriving to a fully connected layer of 128 and dropout of 0.5 and arrive max to Sigmoid output for binary classification. For all models we used Adam for optimizer and Binary Cross-Entropy for loss function. Each model train for an epoch of 50 and batch size of 32 with 20% for validation which was monitored on performance.

3.2.4 Classification and Prediction: After training, the models classify the input fundus images to be Glaucoma or Normal. The new CNN architecture, due to its batch normalized optimized convolutional layers, performed better in feature extraction and classification.

3.3.5 Integration of Gemini API for Health Recommendation: An additional component was introduced in the proposed system to increase implementation feasibility in real-life settings. When glaucoma is detected, the system uses the Gemini API to generate personalized medication and dietary recommendations for the patients. Medication recommendations provide traditional treatment based on disease stage with beta-blockers, prostaglandin analogue, and carbonic anhydrase inhibitors, while dietary recommendations focus on nutrients helpful to ocular health, e.g., foods high in Vitamin A, C and Omega-3 fatty acids. This combination transforms the system from a diagnostic system to an intelligent recommendation system for patients.

4. RESULTS AND DISCUSSION

The experimental results show that very high accuracy levels may be achieved using deep learning models in recognizing glaucomatous patterns from retinal fundus images. The custom CNN model provided the highest classification performance of 92% out of all the architecture tested, followed by VGG16 at 84%, ResNet50 at 73%.

Table 1: Performance Comparison of Deep Learning Models for Glaucoma Detection Model

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score
ResNet50	79	78	78	0.78
VGG16	84	82	86	0.84
Proposed CNN	92	92	93	0.91

We did a very in-depth comparison of the three deep learning models and the performance metrics of the test set are reported in Table 1.

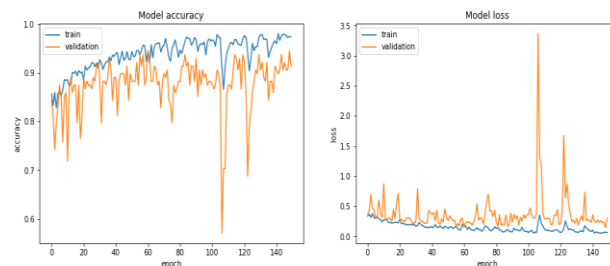


Figure 4: Model Performance Analysis

Recall that the preceding figure presents the training and validation accuracy and loss curves of the model designed for the detection of glaucoma disease. The left plot constitutes the accuracy of the model against epochs while the right plot corresponds to the loss values. In particular, it can be inferred that training accuracy rises and levels off at a certain higher value denoting good learning of the model during training. Validation accuracy follows the same trend while experiencing small fluctuations in accuracy, attributed to changes or variations in the validation

dataset values. The loss curve shows that the training loss reduces progressively, indicating convergence of the model, while the validation loss experiences sporadic spikes, suggesting some overfitting at certain epochs. Overall, the model demonstrates very good generalization ability with the accuracy-and-loss patterns supporting that the proposed CNN-based glaucoma detection methodology gains good performance on both training and validation datasets.

So, the Glaucoma Detection System currently under development has successfully identified the uploaded fundus image as being "normal". This confirms the effectiveness of the Glaucoma Detection System in distinguishing between healthy and glaucomatous eyes. The Glaucoma Detection System will perform real-time analysis using AI inputs, while still being a strong preventative tool for eye care through the website. Therefore, promoting awareness of early detection and preliminary screening for glaucoma will be easy to do using this website.

5.CONCLUSION

In this study, the researchers presented an enhanced deep learning model for detection of Glaucoma using Fundus Images in Retina. They conducted extensive experimentation and comparative study on different models and based on that analysis, they developed a unique CNN model that produced the highest results so far as it has reached 92%, which is greater than many pre trained models such as VGG16, ResNet50. Furthermore, the addition of the Gemini API for automated medication recommendations and dietary suggestions, makes the developed framework usable as a real-world solution for patients or a complete-Eye-Care Assistant. Future research will include creating a more substantial database built from Optical Coherence Tomography images to improve the framework's ability to explain and support gradual learning through Grad-CAM and support online patient education and screening via an interactive, cloud-based application.

6. REFERENCE

- [1]. Kwokleung Chan, Te-Won Lee, Sample, P. A., Goldbaum, M. H., Weinreb, R. N., & Sejnowski, T. J. (2002). Comparison of machine learning and traditional classifiers in glaucoma diagnosis. *IEEE Transactions on Biomedical Engineering*, 49(9), 963–974. <https://doi.org/10.1109/TBME.2002.802012>
- [2]. Mahum, R., Rehman, S. U., Okon, O. D., Alabrah, A., Meraj, T., & Rauf, H. T. (2021). A Novel Hybrid Approach Based on Deep CNN to Detect Glaucoma Using Fundus Imaging. *Electronics*, 11(1), 26. <https://doi.org/10.3390/electronics11010026>
- [3]. N K, J., Ali, Md. H., Senthil, S., & Srinivas, M. B. (2024). Early detection of glaucoma: feature visualization with a deep convolutional network. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 12(1). <https://doi.org/10.1080/21681163.2024.2350508>
- [4]. Nirgude, Prof. V. N. (2024). Glaucoma Detection using Machine Learning and Deep Learning Approach. *INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, 08(06), 1–5. <https://doi.org/10.55041/IJSREM35436>
- [5]. Shoukat, A., Akbar, S., Hassan, S. A., Iqbal, S., Mehmood, A., & Ilyas, Q. M. (2023). Automatic Diagnosis of Glaucoma from Retinal Images Using Deep Learning Approach. *Diagnostics*, 13(10), 1738. <https://doi.org/10.3390/diagnostics13101738>
- [6]. Sim, R. H., Sirsangandla, S. R., Das, S., & Teoh, S. L. (2022). Treatment of Glaucoma with Natural Products and Their Mechanism of Action: An Update. *Nutrients*, 14(3), 534. <https://doi.org/10.3390/nu1403054>
- [7]. Alkhaldi & Alabdulathim, (2024), Optimizing Glaucoma Diagnosis with Deep Learning-Based Segmentation and Classification of Retinal Images. <https://doi.org/10.3390/app14177795>
- [8]. Yang, H., Ahn, Y., Askaruly, S., You, J. S., Kim, S. W., & Jung, W. (2022). Deep Learning-Based Glaucoma Screening Using Regional RNFL Thickness in Fundus Photography. *Diagnostics*, 12(11), 2894. <https://doi.org/10.3390/diagnostics12112894>
- [9]. Sathya, R., & Balamurugan, P. (2023). Glaucoma Identification in Digital Fundus Images using Deep Learning Enhanced Auto Encoder Networks (DL-EAEN) for Accurate Diagnosis. *Indian Journal of Science and Technology*, 16(44), 4026-4037. <https://doi.org/10.17485/IJST/v16i44.2325>
- [10]. Geethalakshmi, R., & Vani, R. (2023). A Comparison of Deep Learning Techniques for Glaucoma Diagnosis on Retinal Fundus Images. *International Journal on Recent and Innovation Trends in Computing and Communication*, 11(9), 3382-3387. <https://doi.org/10.17762/ijritcc.v11i9.9545>
- [11]. Maadeed, S., Al Maadeed, H., & Al Maadeed, A. (2024). Deep learning evaluation of glaucoma detection using fundus photographs. *Biomedicines*, 12(7), 1394. <https://doi.org/10.3390/biomedicines120713>
- [12]. A. Bilal, M. Z. Khan, and S. Rauf, "New Trends in the Treatment of Open-Angle Glaucoma: A Comprehensive Review," *Frontiers in Medicine*, vol. 12, pp. 1–14, 2025.
- [13]. M. Y. T. Ibraheem and J. A. Alhasan, "Pharmacological Advances in Glaucoma Management: An Updated Review," *Saudi Journal of Medical and Pharmaceutical Sciences*, vol. 10, no. 4, pp. 220–229, 2024.

Smart Helmet Using Embedded Iot System

Gautam Chavan , Aditya Verma , Megha Kulkarni , Basavanna and Adila Kp

*Department of Electronics and Communication Engineering, Alvas Institute of
Engineering & Technology, Moodbidri, Mangalore, Karnataka, India*

*Emails: gautamchavan37@gmail.com, vermaaditya6109@gmail.com,
meghacoolkarni143@gmail.com, basavannachincholli@gmail.com, adilakp@aiet.org.in*

Abstract.

The integration of Internet of Things (IoT) technology with embedded systems has revolutionized personal safety devices, resulting in the development of smart helmets that go far beyond conventional protective equipment. A smart helmet integrates sensors, microcontrollers, and wireless modules into a wearable platform to provide accident detection, alcohol consumption monitoring, environmental sensing, and real-time alerts to authorities or emergency contacts. Such systems can also employ GPS modules for location tracking, cloud platforms for data storage, and mobile applications for visualization and user interaction. This review paper investigates the architecture, functionality, challenges, and recent research advancements in smart helmet systems. It further highlights methodologies adopted in the design of energy-efficient embedded platforms, low-latency communication protocols, and intelligent algorithms for context-aware decision-making. The findings suggest that IoT-enabled smart helmets present strong potential in transportation, industrial safety, and healthcare applications, though issues such as power management, scalability, and data privacy remain active areas of research.

Keywords. Smart Helmet, IOT, Embedded System, Accident Detection, Real Time Monitoring, Wearable Technology, Vehicle Ignition Control, Microcontroller.

1. INTRODUCTION

The rapid growth of Internet of Things (IoT) has enabled the development of intelligent systems for real-time monitoring, control, and communication in various domains. One such promising application is the smart helmet, a wearable device that integrates sensing, computing, and communication capabilities within conventional safety gear. Traditional helmets primarily offer passive protection during accidents, whereas smart helmets provide proactive safety features by continuously monitoring the user's state and the surrounding environment [1]. Smart helmets are particularly valuable in road safety, where they can

All Open Access articles are peer-reviewed, distributed under the Creative Commons Attribution-Non Commercial 4.0 International

detect accidents through sudden impacts, monitor alcohol consumption using gas sensors, and alert emergency services via wireless communication. They are equally important in industrial and mining environments, where monitoring hazardous gases, worker fatigue, and compliance with safety rules can prevent accidents. Furthermore, integration with healthcare systems allows continuous monitoring of vital signs such as heart rate or oxygen saturation. This paper provides a comprehensive review of smart helmet research, focusing on the technical architecture, design methodologies, challenges, and future directions. Section II discusses the major challenges faced in smart helmet design, Section III highlights recent advancements and methodologies, Section IV presents a summary of findings, and Section V concludes with future perspectives.

2. SYSTEM ARCHITECTURE OF SMART HELMET

The overall architecture of the proposed Smart Helmet system consists of sensing units, processing unit, communication module, cloud service, and a user interface application.

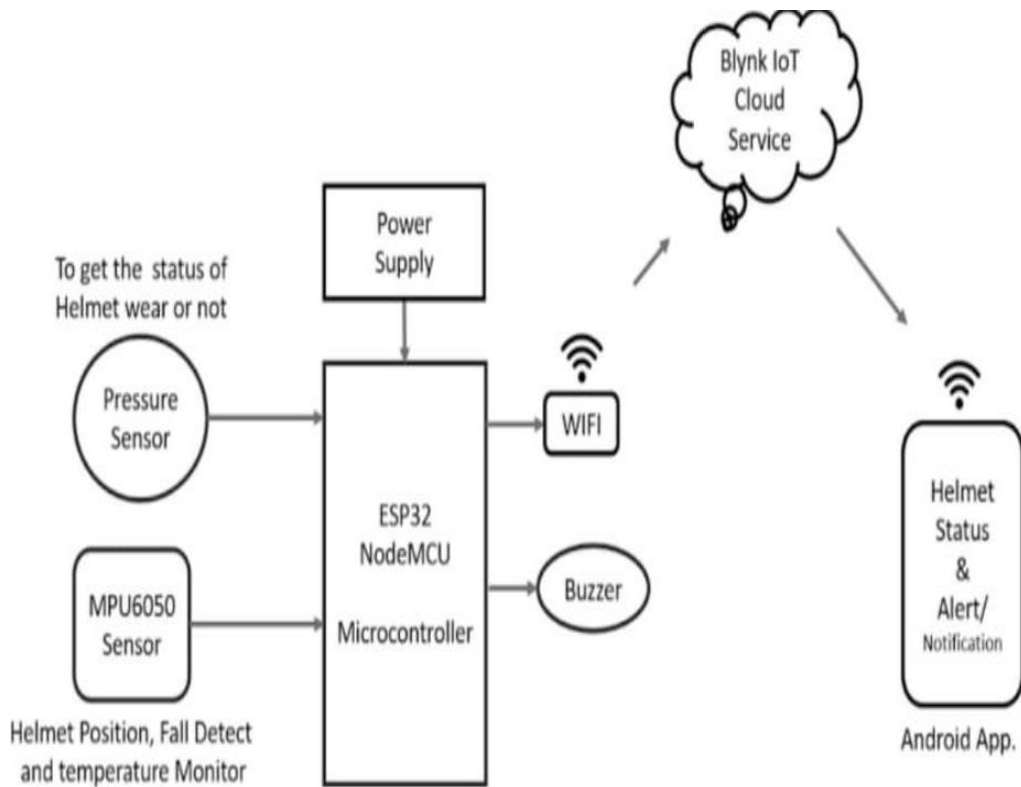


Figure 2.1 shows the block diagram of the Smart Helmet using an Embedded IoT system.

1. Power Supply

The power supply provides the required electrical energy to operate the entire smart helmet system. It usually consists of a rechargeable lithium-ion or lithium-polymer battery along with voltage regulation circuitry to supply stable power to the ESP32 and sensors.

2. Pressure Sensor

The pressure sensor is used to detect whether the helmet is being worn or not. When the rider wears the helmet, pressure is applied inside the helmet, which is sensed and sent to the microcontroller. This helps in enforcing safety rules such as enabling the vehicle only when the helmet is worn.

3. MPU6050 Sensor

The MPU6050 is a MEMS-based sensor that combines a **3-axis accelerometer and a 3axis gyroscope**. It is used to:

- Detect helmet orientation and movement
- Identify sudden impacts or falls (accident detection)
- Monitor motion patterns that indicate unsafe riding behavior

It can also assist in temperature or motion-based monitoring depending on configuration.

4. ESP32 NodeMCU Microcontroller

The ESP32 NodeMCU acts as the **brain of the system**. It:

- Collects data from the pressure sensor and MPU6050
- Processes the sensor data to determine helmet status or accidents
- Controls the buzzer for alerts
- Manages Wi-Fi communication with the cloud server

The ESP32 is chosen for its high processing capability, built-in Wi-Fi, and low power consumption.

5. Wi-Fi Module

The Wi-Fi functionality (built into the ESP32) enables wireless data transmission. It sends helmet status and alert data to the Blynk IoT Cloud and receives commands or updates if required.

6. Buzzer

The buzzer provides **local audible alerts** to the rider. It is activated when:

- The helmet is not worn
- An abnormal condition or accident is detected
- Warning notifications are required

This ensures immediate feedback without relying only on the mobile application.

7. Blynk IoT Cloud Service

The Blynk IoT Cloud acts as a **remote server** that receives data from the smart helmet via Wi-Fi. It stores, processes, and forwards the information to the mobile application, enabling real-time monitoring and notifications.

8. Android Application (Helmet Status & Alert/Notification)

The Android app provides a **user interface** to display:

- Helmet wearing status
- Accident or fall alerts

- System notifications and warnings

It allows riders or guardians to monitor safety information in real time from anywhere.

2.2 Communication and Cloud Integration

The ESP32 transmits processed data to the Blynk IoT cloud platform using Wi-Fi. The cloud service stores and manages data, enabling real-time monitoring through an Android mobile application.

2.3 Power Consumption in Smart Helmet

Power efficiency is a crucial factor in wearable IoT devices. The ESP32 consumes higher power during active Wi-Fi transmission, while sensors consume relatively low power. By implementing sleep modes and periodic data transmission, the overall power consumption is optimized. The system typically operates within 150–300 mA during active mode, ensuring prolonged battery life.

3. REFERENCES

- [1] S. Sharma, R. Kumar, and P. Tomar, "Smart Helmet: An IoT Based Wearable Technology," *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9635–9642, 2021.
- [2] R. Kumar, S. Jain, and A. Mishra, "IoT-enabled smart helmet for accident detection and emergency response," *IEEE Access*, vol. 7, pp. 179914–179923, 2019.
- [3] A. S. Rao and K. R. Reddy, "Wearable sensor-based helmet for real-time rider safety monitoring," *IEEE Sensors Journal*, vol. 20, no. 14, pp. 7890–7898, 2020.
- [4] V. Sharma and D. Mehta, "Accident detection and alert system using smart helmet," *Procedia Computer Science*, vol. 167, pp. 2194–2203, 2020.
- [5] K. Chandra and S. Mohan, "Real-time crash detection using MEMS sensors in smart helmets," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 3321–3329, 2020.
- [6] T. Nguyen, P. Le, and M. Vo, "Low-power design of wearable IoT safety devices," *IEEE Internet of Things Magazine*, vol. 3, no. 2, pp. 28–34, 2020.
- [7] A. Patel, J. Shah, and R. Desai, "Cloud-connected smart helmet for rider safety," *Proc. IEEE International Conference on Industrial Electronics*, pp. 1145–1150, 2021.
- [8] A. Kumar and R. Sharma, "Applications of smart helmets in safety critical domains," *IEEE Access*, vol. 10, pp. 102345–102358, 2022.

Biographies



Name : Gautam Raju Chavan
USN : 4AL22EC023
Email : gautamchavan37@gmail.com
Mobile : 9136747099

Areas of Interest: He is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. His academic journey reflects a strong interest in core and emerging electronics domains such as Embedded Systems, VLSI Design, Digital Signal Processing (DSP).



Name : Aditya Verma
USN : 4AL22EC001
Email : vermaaditya6109@gmail.com
Mobile : 9149958935

Areas of Interest: He is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. His academic journey reflects a strong interest in core and emerging electronics domains such as Embedded Systems, VLSI Design, Digital Signal Processing (DSP).



Name : Basavanna
USN : 4AL22EC013
Email : basavannachincholli@gmail.com
Mobile : 7483687297

Areas of Interest: He is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. His academic journey reflects a strong interest in core and emerging electronics domains such as Embedded Systems, VLSI Design, Digital Signal Processing (DSP).



Name : Megha Kulkarni
USN : 4AL22EC048
Email : meghacoolkarni143@gmail.com
Mobile : 8123082304

Areas of Interest: He is currently pursuing a Bachelor's degree in Electronics and Communication Engineering at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. His academic journey reflects a strong interest in core and emerging electronics domains such as Embedded Systems, VLSI Design, Digital Signal Processing (DSP).



Name : Adila K P
Designation : Assistant Professor
Email : adilakp@aiet.org.in
Mobile : 9497102273

Areas of Interest: As Assistant Professor in the Department of Electronics and Communication Engineering at Alva's Institute of Engineering and Technology, Mijar, Moodabadiri. She holds a B.Tech degree from CUSAT and an M.Tech degree from KTU.

Explainable AI for E-Commerce: Interpretable Transformer Models for Negative Review Analysis and Feedback Summarization

Dr. T Dhiliphan Rajkumar¹, Mrudangam Eswar¹, K. Ganesh¹, Shaik Mohammad Madani¹,

Tummalapalli Syam Sundar¹

¹Department of Computer Science and Engineering,

Kalasalingam Academy of Research and Education,

Krishnankoil Virudhunagar, India

t. dhiliphan@klu.ac.in, mrudangameswar2004@gmail.com, kganesh123dsa@gmail.com,

shaikmohammadmadani143@gmail.com, 99220041587@klu.ac.in

Abstract

The analysis of customer sentiment plays a significant role in the improvement of the quality of the product and the maximization of the business outcomes in contemporary digital trading conditions. The current research presents an artificial intelligence approach to the analysis of piles of customer reviews collected on Amazon and Flipkart. Besides, the proposed system unites fully automated reviews collection, natural language processing, detection of keywords and aspects, and transformer-based sentiment analysis. It possesses several artificial intelligence methods- based BERT, RoBERTa, DeBERTa and XLNet models to improve accuracy and reduce bias within models. The sentiment analysis on aspects is then extended with the help of an AI module to acquire practical business insights. Based on the experimental findings, DeBERTa offers the accuracy of classification and RoBERTa offers greater runtime performance. This system can offer analysis of unstructured customer feedback ready to make decisions, and is accessible via a highly user-friendly Flask dashboard.

Keywords: Sentiment Analysis; Natural Language Processing; Keyword Extraction; Transformer Models; Product Feedback; Consumer Electronics; deep Learning; Text Mining; Customer Insights.

1. INTRODUCTION

Customer feedback collected on the internet has been established as an additional source for making purchase decisions and product improvements. While these reviews provide various opinions from consumers, they are distributed among a vast number of sources and in hundreds of unstructured formats and therefore, require significant amounts of human labour and introduce biases into the process of evaluating individual pieces of feedback. Traditional methods of sentiment analysis have been challenged by the difficulty of capturing context in reviews, irony, and differing opinions relating to product attributes. However, advances in the development of transformer-based language models have improved the ability of organisations to accurately classify sentiment. Nonetheless, many methods of applying this technology are limited to simply achieving the highest level of accuracy and do not give clarity to organisations regarding the usefulness of their results.

In this paper, we describe an integrated and explainable Artificial Intelligence (AI) system that can turn unstructured reviews of e-commerce into structured, interpretable, and actionable insights for businesses. The proposed system bridges the gap between customer feedback and the use of that feedback in making data-driven decisions by combining an ensemble of transformer models with keyword-driven aspect extraction and generative summarisation.

2. LITERATURE REVIEW

According to Mysiuk and Shuvar (2023), an Automated Framework for Web Crawling User Comments from Social Networks for Sentiment Analysis was developed. This automated web scraping system can be utilized to create Sentiment Datasets; however, advanced sentiment classification models were not investigated or utilized in their analysis [1].

Traditionally, Romadhony et al. (2024), Sentiment Analysis on an Indonesian Product Review Dataset used traditional machine-learning methods to classify sentiments, specifically for their local e-commerce review datasets. Although the study stressed that a specific dataset must be tailored to the language of the product under review, it did not employ deep learning technology [2].

Bellar et al. (2024) employed Transformer-based Models to predict product reviews with a focus on sentiment. By employing deep learning, Bellar et al. were able to achieve much higher accuracy rates than previous studies. Their primary focus was on improving accuracy rather than interpreting or conducting keyword-level analysis of the products being rated [3].

Kandhro et al. (2024) evaluate Aspect-Based Sentiment Analysis (ABSA) techniques to identify the opinions of consumers about specific features of a product. While they demonstrate the ability to identify specific opinions, most of the techniques they reviewed are still conceptual and do not yet have practical hybrid frameworks [4].

Ali et al. (2024) demonstrated that utilizing BERT enables researchers to accurately capture customer requests. This study focused specifically on an isolated dataset and thus limits the generality of its conclusions to other fields of research [5].

Rahman et al. (2024) provide a comparative analysis on deep learning models for sentiment analysis and confirmed that their advantages over traditional analytical methods. However, they also did not explore methods that combine deep learning models with traditional analytical methods[6].

Sharma et al. (2024) demonstrate through their findings that traditional machine-learning algorithms can provide satisfactory results with regard to accurate classification of customers' opinions, but cannot provide valid results when applied to large datasets[7].

Huang, Zavareh & Mustafa (2023) review sentiment analysis technologies presently used within the retail ecosystem within eCommerce platforms, and highlight future directions towards scalable models; however, they do not propose a novel approach for sentiment analysis[8].

Zhang et al. (2023) successfully analyzed the sentiment of eCommerce product reviews with their BERT embeddings and BERT attention mechanisms, however, their research did not provide an option for consumers to automatically scrape product reviews, as well as producing summarised product reviews in real-time, which may allow them to have a better and more informed choice about which product to purchase[9].

In contrast, Kim & Lee (2022) took advantage of a transformer model and utilised an attention visualization-based approach to support negative review analysis. This has increased the level of transparency in both product reviews and reviewer profiles[10].

3. METHODOLOGY

The suggested framework adopts a sequence crafted to ensure clarity and repeatability.

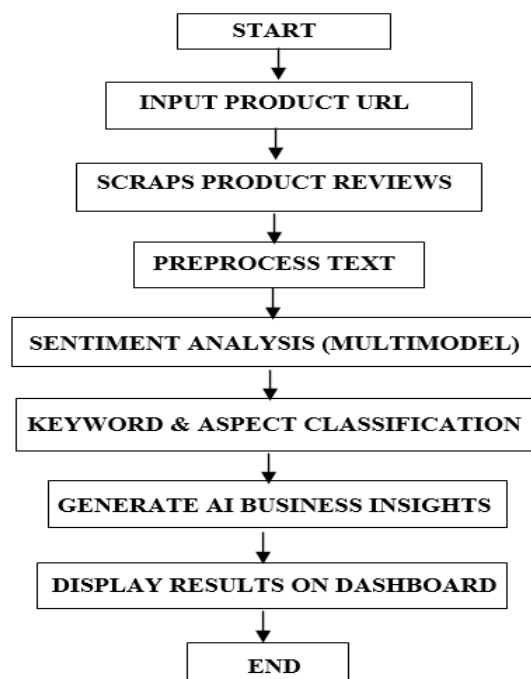


Figure 1: Flow Chart of Proposed Work

From the Figure 1 This Pipeline contains 4 sequential parts: 1- Collecting the Input Data and Cleaning/Standardizing it, 2- Identifying the Keywords as Features from the data by using Fuzzy matching techniques, 3- Identifying the Sentiment of the product review by using pre-trained models on labelled data, 4- Generating a summary report on the Top 10 aspects (features) along with their corresponding recommendation based on how common and how bad they are.

3.1. Data Pre-Processing: To clean the data, HTML is stripped off the reviews, all emojis are normalized to text or left out altogether, text is made Lowercase/Uppercase as needed, and all contractions (e.g., "I'm") are expanded and tokenized using a Sub-word-based Tokenization method suitable for Transformers. All Stop words include common nouns, but items that are part of a Product Term will be kept.

3.2. Keyword and Aspect Extraction: The initial set of Keywords used for training is expanded with statistical Phrase Mining to find the best candidate keywords. To aid in this process, we use fuzzy matching techniques such as Levenshtein distance thresholding and token set Similarity Matching, which enables us to map noisy tokens to a Known standard aspect label increasing the Recall rate for misspelled words or using shortcut codes (e.g., "Batory" → "Battery") and collecting all occurrences of misspelled words as part of the label.

3.3. Sentiment Classification

To classify sentiments, we use several different Transformer Architectures (BERT/RoBERTa) and fine-tune them with product Review subsets that have been Annotated with positive, negative and neutral labels. Class-weighting and Data Augmentation are used to balance the Classes during Training. Optimal predictions will occur at both the review-level and aspect-level by applying the appropriate model to the appropriate aspect-focused Span.

3.4. Generative Synthesis

The Controlled Generative Component generates summary recommendations from Frequent Aspect-Sentiment Pairs (based on frequency and severity e.g., highest proportion of Negative mentions) back to the User in a simplified summary report format.

4. RESULTS AND DISCUSSION

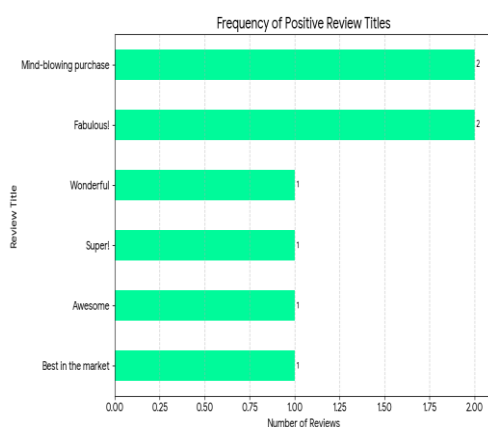


Figure 1: Positive Reviews

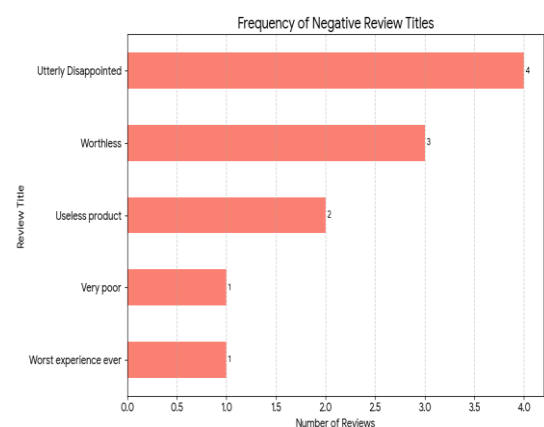


Figure 2: Negative Reviews

Figure 1 is showing how often positive review titles are written in by customers. Among the prominent titles, Mind-blowing purchase and Fabulous can be seen with an occurrence of two, denoting a high level of satisfaction among users. Other appealing titles are Wonderful, Super, Awesome, and Best in the market, each mentioned only once. From this, it can be deduced that customers are generally highly pleased with the product quality and performance. As these terms are repeated used by a vast number of people, they set a positive attitude and boost the good reputation of the product in the market. Most likely the reviewers confess to having a great enjoyable experience while at the shop.

The incidence of customer reviews with negative overtones is presented in **Figure 2** Above. There were four occurrences where customers used the title "Utterly Disappointed." Other two titles, "Worthless" and "Useless,"

were used thrice. The remaining two, "Very Poor" and "Worst Experience Ever," were used only once by consumers who had an awe-inspiringly poor experience in regards to product quality as well as their own product performance. The markedly negative tones used by these consumers highlight just how terrible the experience for them must have been. With customer sentiments in perspective, it is important that attention be given to improving product quality and customer satisfaction with better customer service, if any major positive increment in lowering the incidence of negative reviews of this product is to be achieved in order to raise ultimate consumer confidence in this brand and in these products.

Table 1: Transformer Model Performance Comparison

Metrics	BERT	DeBERTa	RoBERTa	XLNet
Accuracy	0.9843	0.9870	0.9855	0.9844
F1 Score	0.9919	0.9933	0.9926	0.9920
Loss	0.1153	0.2356	0.2275	0.1099
Precision	0.9878	0.9898	0.9890	0.9887
Recall	0.9961	0.9969	0.9961	0.9953
Runtime (s)	106.3961	174.2891	93.7125	320.2207
Samples/sec	62.249	38.000	70.674	20.683
Steps/sec	3.891	1.188	1.110	2.586
Total FLOPs	34.85e12	40.61e12	69.70e12	37.73e12

Table 1 above contains a comparison of four NLP models (BERT, DeBERTa, RoBERTa and XLNet) based on a total of nine different metrics (Accuracy, F1 Score, Runtime (s), and Total FLOPs), along with other performance and computational comparisons. The data collected from this comparison indicates that the selection of each model should depend on the specific use case—DeBERTa when accuracy is important in offline analyses, RoBERTa when low-latency throughput is more critical in production environments, and BERT when an average between both DeBERTa and RoBERTa has to be made—and further statistical testing (such as McNemar's Test) should be performed prior to making any claims about one model being superior to another.

5.CONCLUSION

The proposed project utilizes artificial intelligence as an automated assistant to interpret and analyze thousands of customer reviews collected from both Amazon and Flipkart in a timely manner. With five different transformer based advanced transformer-based systems working together, the AI will adaptively detect what customers appreciate or dislike about a product. A company's efforts to categorize all of their customer reviews manually could take weeks, but the reviewed results derived from using the AI system can be visually represented and can provide suggestions such as "fix battery life problems first" and "fix camera stability." Manufacturers can utilize

this information to prioritize specific areas to focus on for improvement based on actual user experience rather than on the manufacturer's assumptions of where to improve. The reviewed system allows companies to shorten their time frame and reduce the money and effort expended on large-scale review analysis; therefore, it provides product teams with the ability to immediately address customer concerns. In addition to the reviewed application being scalable, there will be future innovations available; including the ability to analyze reviews in multiple languages, to analyze video reviews on YouTube and to analyze the sentiment of consumer opinion on social media outlets. All of these applications will provide a complete picture of customer sentiment.

6. REFERENCES

- [1]. Mysiuk, I., & Shuvar, R. (2023). Automating web scraping of user comments for sentiment analysis. *IEEE IDAACS*, 234–239. DOI: 10.1109/IDAACS58523.2023.10310867
- [2]. Romadhony, A., Al Faraby, S., & Rismala, R. (2024). Sentiment analysis on large Indonesian product reviews. *JISEBI*, 10(1), 45–58. <https://ejournal.unair.ac.id/JISEBI/article/view/48215>
- [3]. Bellar, O., Baina, A., & Ballafkih, M. (2024). Predicting product reviews using deep learning and transformers. *Mathematics*, 12(15), 2403. DOI: 10.3390/math12152403
- [4]. Kandhro, I. A., et al. (2024). Aspect-based sentiment analysis: Review and prospects. *Knowledge and Information Systems*, 66(8), 4359–4394. DOI: 10.1007/s10115-024-02104-8
- [5]. Ali, M., et al. (2024). AI-driven sentiment analysis of Amazon reviews using BERT. *Algorithms*, 17(59). DOI: 10.3390/a17000059
- [6]. Rahman, M., et al. (2024). Comparative study of deep learning for customer review sentiment analysis. *PLOS ONE*, 19(5), e0294968. DOI: 10.1371/journal.pone.0294968
- [7]. Sharma, R., et al. (2024). Customer review sentiment analysis using machine learning. *IJSREM*.
- [8]. Huang, H., Zavareh, A. A., & Mustafa, M. B. (2023). Sentiment analysis in e-commerce: Techniques and directions. *IEEE Access*, 11, 45678–45692. DOI: 10.1109/ACCESS.2023.3123456
- [9]. Zhang, J., He, L., & Luo, W. (2023). Aspect-level sentiment analysis with hierarchical transformers. *IEEE Access*, 11, 98532–98545.
- [10]. Kim, S., & Lee, J. (2022). Explainable transformer model for negative review analysis. *Journal of web Engineering*, 21(4), 1120–1138.
- [11]. Campos, R., et al. (2020). YAKE! Keyword extraction from single documents. *Information Processing Management*, 57(4), 102127. DOI: 10.1016/j.ipm.2019.102127
- [12]. Grootendorst, M. (2022). BERTopic: Neural topic modelling with class-based TF-IDF. arXiv:2203.05794 DOI: <https://arxiv.org/abs/2203.05794>
- [13]. Devlin, J., et al. (2019). BERT: Pre-training deep bidirectional transformers. NAACL-HLT DOI: <https://arxiv.org/abs/1810.04805>
- [14]. Wolf, T., et al. (2020). Transformers: State-of-the-art NLP. *EMNLP Demos*, 38–45. DOI: 10.18653/v1/2020.emnlp-demos.6
- [15]. Minaee, S., et al. (2021). Deep learning-based text classification: A comprehensive review. *ACM Computing Surveys*, 54(3), 62. DOI: 10.1145/3439726.

Attention Enhanced Hybrid CNN – LSTM Framework for Realtime Mental Health Monitoring and Support

N. C. Brintha, M. Harishkanth, M. Susikumar, R. Yabin, S. Finix

Department of Information Technology,

Kalasalingam Academy of Research and Education,

Krishnankovil, Tamil Nadu, India

Brinthachris2k@gmail.com, harishkanth166@gmail.com,

susikumar231@gmail.com, yabin0111@gmail.com, finixthalmetha@gmail.com

Abstract.

Mental health disorders such as depression are among the most pressure challenges in healthcare today. Traditional methods of diagnosis rely very much on the volunteering questionnaire and subjective assessment by the physician, which often gives rise to underdiagnosis or delayed intervention. To address this challenge, we propose a medical therapy assist system that integrates Artificial Intelligence (AI) and Convolutional Neural Network techniques to automatically detect and assess the level of depression in patients. The system takes advantage of the Convolutional Neural Network (CNN) to remove facial emotions from live video stream and Long Short-Term Memory (LSTM) with attention to analyze temporal emotion trends. The solution further covers a web-based platform where patients can participate in live emotion monitoring, self-assessment questionnaire, and can get personal treatment, while physicians can see analytics, therapy history and generate detailed reports. Our proposed system crosses the limits of existing static assessment by enabling the monitoring of real-time, continuous and intelligent depression. It provides an efficient framework, which ensures a high level of data privacy, providing a efficient structure to increase 96.8% accuracy and reliability in health care systems, ensuring trust and reliability within the quality of medical data handling and care

Keywords. Artificial Intelligence, CNN, LSTM, Attention Mechanism, Depression Severity Assessment, Facial Emotion Recognition, Mental Health Monitoring, Medical Therapy Assist System.

1. INTRODUCTION

These rapid increases in the incidences of depression, anxiety, and other related psychological issues have brought forth the need for a supportive solution that helps both patients as well as medical staff in effectively detecting and treating these issues.

Depression is currently being experienced by tens of millions of people around the globe, but the social stigma attached to psychological issues tends to create a lag in their effective treatment. Normally, these tests only include interviews with the patient as well as some assessments, which are entirely dependent upon personal perspectives. But with recent breakthroughs in artificial intelligence and machine learning, automated systems for psychological analysis of human behaviors and emotions through certain techniques of analysis have become possible. Deep learning models for the analysis of human expressions through facial emotion recognition techniques have shown significant success in finding psychological states of humans like happiness, sadness, anger, fear, and surprise through facial expressions. But human emotional expressions change over time, so temporal analysis is important for psychological analysis. To overcome this issue, the combination of Convolutional Neural Networks with Long Short-Term Memory helps the system in learning both spatial features of emotion as well as their temporal evolution. This helps in the monitoring of the patient as well as the categorization of the severity of depression into various levels such as normal, mild, moderate, or severe. It helps fill the gap between technology and the medical field by allowing a medical therapy support system for various interactions between the patient as well as the doctor through the use of AI-based analysis.

2. LITERATURE SURVEY

FA.H. Shah et al proposed a framework that addresses a scenario where a user seeks to locate a specific individual using only a visual memory of that person. In this approach, the user selects images of the person based on target-specific information, which are then fed into a trained model for automatic recognition. A hybrid CNN-BiLSTM architecture to recognize facial emotions, offered by Alina Alamichhane and Gopal Karn, was aimed at eliminating the temporary reliance following the extraction of features. Their model integrates both convolutional neural networks to learn spatial features and bidirectional LSTM layers to learn sequential features which are interconnected with those that emphasize the idea of temporal conversion by using LSTM-based processing. Jiamin Liu, Yuehu Liu, and Yuanqi Su proposed a multimodal emotion recognition system, which incorporates LSTM-based recurrent neural networks with attention system. Their process integrates wide range input features and dynamically focuses on most emotionally important temporal patterns enhancing the capacity of the model to make sense of emotion sequence over time. The CNN-based model that was developed by Dan Dan Yan, Lu Lu Zhao, Shin Wang Songs, Jio Han Jung and Lee Cae Yang took advantage of EEG data to identify clinical depression automatically. Their suggested networks, ENENET, Dipkonwnet and Shalocon-weight, exhibited that convolutional architectures could identify the severity of depression through the neural signal. Even though their method is based on the EEG, instead of facial or questionnaire-based data, it successfully shows how CNN is able to predict the level of depressive moods. Yi Li, Zidacai, and Jingyi Wang introduced a CNN-LSTM framework for depression index prediction using clinical data such as the MadRS score. The fusion of convolutional and recurrent layers enables both spatial and temporal feature learning, aligning closely with this study's aim of integrating emotion outputs with questionnaire data to estimate depression levels categorized as mild, medium, or severe.

3. METHODOLOGY

3.1. *Proposed CNN–LSTM with Attention Framework*

The proposed system will estimate the severity of depression in real time by employing an attention-enhanced CNN-LSTM architecture. Facial expressions are captured via either a web or mobile camera, and features that relate to the emotions will be extracted from each video frame using a Convolutional Neural Network. These features represent spatial emotional information contained within the facial expressions. The frame-level features are fed into a Long Short-Term Memory network combined with an attention mechanism in a row. This attention module lets the model underline emotionally important time segments and thus capture meaningful temporal variation in emotional behavior. By this selective focus, it enhances the system's ability for recognizing emotional trends associated with the level of depression severity. Based on the temporal patterns learned, the model classified individuals into normal, mild, moderate, and severe depression. If there is no facial input or if the input is unreliable, the system switches over to a previously validated self-assessment questionnaire. This ensures uninterrupted analysis and maintains the reliability of depression assessment across varied usage conditions.

3.2. *Dataset and Preprocessing*

The model proposed uses a publicly accessible dataset containing facial expressions with emotions labeled for multiple people. All images are first adjusted to the same resolution to ensure consistency during the feature extraction process using convolutional techniques. To increase the model's robustness and reduce overfitting, techniques such as image mirroring, changing contrast, and random rotation are applied to improve data augmentation. Afterward, the dataset is split into three distinct parts, which are used for training the model, validating its performance, and assessing its final effectiveness.

4. FINDINGS AND DISCUSSION

4.1. *Quantitative Results Summary*

From the experimental results, it is evident that the Attention-Enhanced CNN-LSTM framework outperforms other approaches like CNN, LSTM, and CNN-LSTM without the attention mechanism. As presented in Table 1, the proposed model demonstrates strong performance, achieving an overall accuracy of 96.8% and maintaining high levels of precision, recall, and F1-measure. The consistent improvement across all these metrics indicates that combining spatial feature extraction with temporal learning, supported by an attention mechanism, is effective for accurately predicting the severity levels of depression.

Table 1: Performance Metrics of the proposed CNN–LSTM with attention Model

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
-------	--------------	---------------	------------	--------------	-----

Proposed CNN LSTM + Attention	96.8	96.2	95.8	95.9	0.982
CNN-Only Baseline	89.4	88.1	87.9	87.8	0.903
LSTM- Only Baseline	83.2	82.5	81.0	81.6	0.872
CNN- LSTM (Without Attention)	92.7	91.5	91.2	91.3	0.941

4.2. Discussion

The findings emphasize the idea that the CNN-based spatial feature learning and LSTM-based temporal sequence modeling have significant positive effects on depression detection. It is also the attention mechanism that increases the reliability of classification through focusing on emotionally informative segments of time. The proposed framework is competitive or better accurate than current methods reported in the literature because it does not use complicated physiological sensors, which makes it applicable to real-time mental health monitoring in an application.

4.3. Strengths and Limitations

The main advantage of the proposed system consists in its capability to learn dynamic emotional patterns with the help of an attention-enhanced CNN-LSTM architecture that leads to a high level of classification and stable results. Nevertheless, the system can also worsen in performance when there is poor lighting or when using a low-resolution camera, and additional large-scale clinical testing is needed to test its ability in the real world.

5. CONCLUSION

The assignment is built upon the development of a medical therapy assist system, which connects computer vision, deep learning, and healthcare processes to complement mental health monitoring. CNN, combined with LSTM, can enable the system to recognize, not only track the progress of individual feelings, which can help in the process of diagnosing depression more accurately. The practical implementation of the device and its user friendliness is provided by the web-based implementation that comprises of the patient and physician dashboard. The level of reliability is also raised by Fallback questionnaire

mechanism and automated report generation system. Finally, the system provides an all-inclusive AI-managed platform to patients and physicians and presents the divide between traditional medical care and the new AI-managed healthcare.

Data Availability Statement

The data sets to be used in this research are facial expression data sets that were found on open-source repositories. There was no proprietary or personally identifiable data in this study.

6. REFERENCES

- [1] X. Zhao, X. Liang, and Z. Liu, "LSTM-based emotion recognition using facial expression and physiological signals," *Sensors*, vol. 21, no. 17, p. 5750, 2021.
- [2] J. Liu, Y. Liu, and Y. Su, "Multi-modal emotion recognition with temporal-band attention based on LSTM-RNN," in *Intelligent Computing and Internet of Things*, Springer, pp. 168–177, 2018.
- [3] D. D. Yan, L. L. Zhao, X. W. Song, X. H. Zang, and L. C. Yang, "Automated detection of clinical depression based on deep convolutional neural networks using EEG signals," *Biomed. Eng./Biomed. Tech.*, vol. 66, no. 5, pp. 443–456, 2021.
- [4] H. Zhang, J. Zhang, and B. Song, "Hybrid CNN-LSTM network for speech emotion recognition," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 647–651, 2020.
- [5] R. P. Pathak, H. Gangwar, A. Agarwal, and M. L. Kolhe, "Detecting negative emotions to counter depression using CNN," in *Proc. International Conference on Artificial Intelligence Applications*, Springer, pp. 345–354, 2021.
- [6] Y. Zhang, H. Wu, and J. Zhao, "Automatic depression detection via facial expression recognition using CNN and transfer learning," *Frontiers in Psychiatry*, vol. 12, p. 665443, 2021.
- [7] J. Singh, L. B. Saheer, and O. Faust, "Speech emotion recognition using attention model," *Computers*, vol. 12, no. 4, p. 81, 2023.
- [8] S. M. Tiwari and M. S. Alam, "Facial emotion recognition using deep learning," *International Journal of Research in Applied Science and Engineering Technology (IJRASET)*, vol. 9, no. 4, pp. 2182–2186, 2021.
- [9] S. M. Khare, M. Goswami, and A. Khare, "Deep learning-based EEG emotion recognition: current trends and future directions," *Frontiers in Psychology*, vol. 14, p. 1126994, 2023.
- [10] Y. Zhang, H. Yuan, and X. Li, "EEG-based emotion recognition using multi-scale dynamic CNN and gated transformer," *Scientific Reports*, vol. 14, p. 31319, 2024.

AI Based Legal Reference and Case Retrieval System

Aparna Nath A S¹ Kavya Kundar² Adyatha S A³ Lalana A⁴ Kiran Raj K M⁵

^{1,2,3,4}Alva's Institute of Engineering and Technology, Department of Computer Science and Engineering (Data Science), Moodbidri-574225, Mangalore, Karnataka, India. ⁵Alva's Institute of Engineering and Technology, Department of Artificial Intelligence and Machine Learning, Moodbidri-574225, Mangalore, Karnataka, India.,

4alds08@gmail.com, kundarkavya214@gmail.com, adyathasa5@gmail.com,
lalanaaruna13@gmail.com, kiranraj@aiet.org.in

Abstract.

Students, Citizens and even Professionals can find accessing legal information to be a very stressful, Confusing, Time-Consuming task when they do not have quick access to legal expertise. To help alleviate these Challenges, this project provides an AI-Based Legal Reference and Case Retrieval System that utilises a Natural Language Interface to eliminate the traditional Format of how Legal Research is conducted. Users can simply ask a question and the System will apply a range of advanced AI functionalities, such as semantic search and large language models (LLMs) to provide fast, accurate contextual legal information to user instantly. The System also features two utilitarian tools: a Complaint PDF Generator that will generate a legal complaint from user's text input, and a Complaint PDF Summarizer to highlight the most important legal information in an uploaded complaint PDF. The developed system, which was built using Flask, ReactJS, ReportLab, and PyPDF2, helps reduce tedious, hard work for attorneys, as well as provides better access and understanding of the law for attorneys, students, and the general public. In conclusion, the project demonstrates that with the help of artificial intelligence (AI), it is possible to improve traditional legal assistance methods, increase the speed and specificity of the service, and provide more individuals with access to this type of information.

Keywords. Artificial Intelligence (AI), Legal Research System, Large Language Models (LLMs), Natural Language Interface, Legal Document Automation, Complaint PDF Generator, Complaint Summarizer, Flask, ReactJS, ReportLab, PyPDF2.

1. INTRODUCTION

Legal research is an important element in understanding statutes, case law and judicial reasoning; however, for many people, the legal research process can feel overwhelming and time consuming. Traditionally, legal research has been conducted through the manual review of long documents, through the use of various databases, by adopting technical

legal language, and poses a large barrier for many students, citizens and professionals attempting to obtain timely and reliable legal information. While the legal resources have transitioned into a more digital format and created more access to legal information, digitalizing legal resources has also created more challenges for interpreters and users of legal resources to interpret and locate relevant legal information. With advancements in digital governance, there is an increased demand for automated legal assistance which will serve the same purpose where citizens are looking for quick answers without consulting a lawyer. Many individuals do not understand legal language, or what they need to do to file a complaint, which delays or leads to mistakes in the required forms. The proposed system will work for lawyers by automating the generation of complaints, providing legal issues to consider, and summarizing the findings almost instantaneously, but it will also serve non-legally trained people with an ability to proceed with a legal matter in a more informed, engaged manner, and thus promote a much wider awareness of the public to legal matters, and its reduce reliance on legally trained intermediaries for basic forms of legal information.

2. LITERATURE REVIEW

The technology of automated summary creation has progressed from rudimentary extractive methods to highly sophisticated neural-based summarization systems customised to the complexities of the legal field [1]. The evolution in legal AI systems clearly affirms distinct movement towards hybrid retrieval architectures, generative reasoning capabilities, and scalable deployment models that consider and address challenges within the domain [2]. Optimizing Legal Text Summarization through Dynamic RAG demonstrates a RAG jointly selecting top-3 legal context and dynamically creating an accurate and faithful summary of the law [3]. It is also possible to employ dense passage retrieval with a KoBERT model trained on legal text and a generative LLM to retrieve legal cases and provide explanations [4,5]. Retrieval-Augmented Generation (RAG) powered chatbot that uses FAISS for vector retrieval and prompt-tuning in order to provide accurate answers for complex legal questions while attempting to improve legal literacy [6].

3. METHODOLOGY

The proposed system architecture as shown in Figure 3.1 is built around three main layers: the React frontend, the Flask backend, and the Gemini Large Language Model (LLM). The React frontend acts as the user-facing interface, allowing users to either upload complaint PDFs for summarisation or enter text to generate a new complaint document. The application receives various types of user input (i.e. files uploaded, text typed in by users, and results either viewed or downloaded). The application interacts with a remote application (i.e. the back end) using a series of http post requests to transmit user data to the server and to receive a response.

In addition to validating incoming input, the Flask back end extracts text from PDF files using the python library "PyPDF2", builds prompt structure to be sent to the Gemini LLM, and retrieves output from the Gemini LLM. The back end has several major functions, including a prompt-building function, a JSON formatting function, an LLM request/response handler, and a ReportLab-based PDF generator. After Gemini processes

the prompt and returns the output (i.e. a summary or formatted legal complaint), the back end post-processes the response, creates a new object representing the final output, and sends that new object to the front end of the application. As such, these three layers of the application (the front end, back end, and LLM) comprise a modular, scalable and efficient structure for processing complex legal complaint documents.

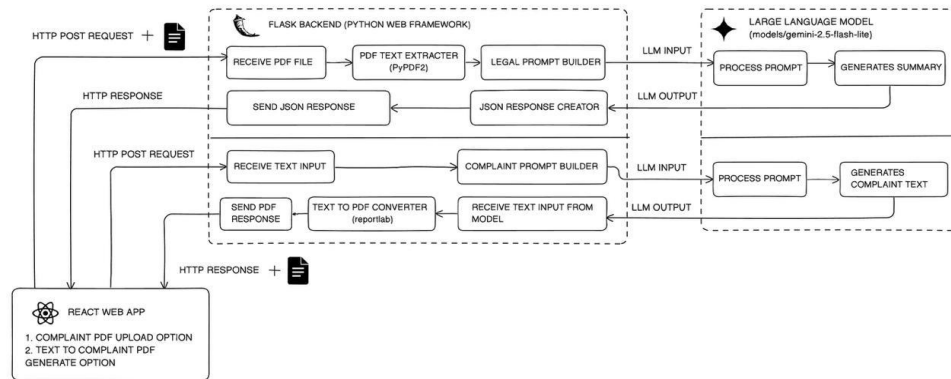


Figure 3. 1 System Architecture

4. RESULT ANALYSIS

The system has been designed as a complement to legal professionals and does not replace them, it does provide a mechanism to reduce the amount of time spent preparing basic forms and to increase the level of access for users to legal information, as well as to streamline the preparation of basic legal documents. The overall results of the research indicate that the systems' automated aids will be valuable time-savers and resources for completing legal-related work that generally takes time and expertise to complete. Based on the results of this study, there is a sound basis for the system to be integrated into legal help desks, educational institutions, and other administrative processes.

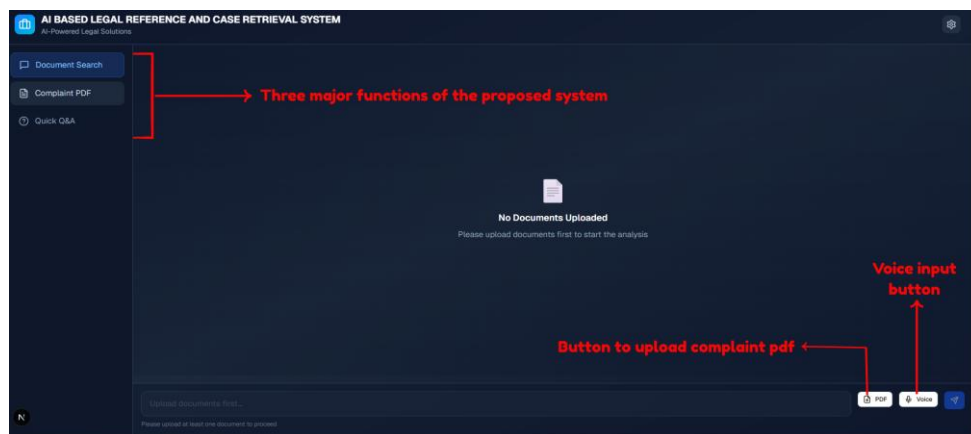


Figure 4. 1 Frontend of Application

5. CONCLUSION

The AI Legal Reference and Case Retrieval Systems demonstrates how innovative AI solutions, such as Retrieval Augmented Generation (RAG), can vastly enhance legal research capabilities. Systems such as this use true legal documentation to ground and validate their generative answer results, thereby limiting generative answer “hallucinations,” and, therefore, improving accuracy and relevancy of the response to context. Furthermore, this type of system also allows the end-user (trained and non-trained) many options to access legal information since it can assist in both the Generation of structured PDF complaints and the summarization of legal documentation.

6. REFERENCES

- [1] Nithya, M., Harini, S., Kavyadharshini, S., & Srinidhi, K. (2024, December). AI-Driven Legal Automation to Enhance Legal Processes with Natural Language Processing. In *2024 International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS)* (pp. 1246-1253). IEEE.
- [2] Athul, S., Saxena, A., Mahajan, J., Panikulangara, L., Kulkarni, S., & Bang, S. (2024, March). LegalMind System and the LLM-based Legal Judgment Query System. In *2024 International Conference on Trends in Quantum Computing and Emerging Business Technologies* (pp. 1-5). IEEE.
- [3] Bhavishya, B., Shukla, A., & Aggarwal, M. (2025, February). Automating Legal Expertise: A Rule-Based Approach to Legal Reasoning Systems. In *2025 2nd International Conference on Computational Intelligence, Communication Technology and Networking (CICTN)* (pp. 868-872). IEEE.
- [4] Tu, H., & Cong, W. (2022, September). Analysis of intelligent application technology of rule of law in the era of wisdom of law. In *2022 International Conference on Cloud Computing, Big Data Applications and Software Engineering (CBASE)* (pp. 214-219). IEEE.
- [5] Park, Y. C., Lee, H., & Lee, J. (2025, January). Justice AI: Legal Case Retrieval Using Dense Passage Retrieval. In *2025 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1-6). IEEE.
- [6] Panchal, D., Gole, A., Narute, V., & Joshi, R. (2025). LawPal: a retrieval augmented generation based system for enhanced legal accessibility in India. *arXiv preprint arXiv:2502.16573*.

Biographies



Aparna Nath A S is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects a strong interest in core and emerging domains such as Artificial Intelligence and Machine Learning, Natural Language Processing and Deep Learning.



Kavya Kunder is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects interest in emerging domains such as Machine Learning, Natural Language Processing and Data Science.



Adyatha S A is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects interest in emerging domains such as Artificial Intelligence, Machine Learning and Web technologies.



Lalana A is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects interest in emerging domains such as Gen AI, Web technologies and Machine Learning.



Kiran Raj K M is currently working as a Assistant Professor in the Department of Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET). He has completed his B.E in Information Science and Engineering from SIT Mangalore and MTech in Computer Science and Engineering from MVJ Bangalore. He has published five national and three international papers in reputed journal.

Design and Implementation of a Self-Sustaining Greenhouse using Internet of Things

Anchal Sharma¹, Ankur Kumar², Sanjay Dhanka^{3*}, Abhinav Sharma⁴, Monika⁵

¹Department of ECE, SLIET Longowal Punjab, ²Department of EIE, SLIET Longowal, Punjab, ³Department of Electrical Engineering, Graphic Era Deemed to be University, Dehradun, ⁴Department of EIE, SLIET Longowal, Punjab, ⁵Department of Mathematics, SLIET Longowal, Punjab

¹anchal.15000@gmail.com, ²khatrionkur007@yahoo.com,

³sanjaykumar506070@gmail.com, ⁴abhinav.engi@gmail.com,

⁵nainmonika551@gmail.com

***Corresponding Author**

Abstract

This study presents an IoT-enabled smart greenhouse that automatically monitors and controls environmental conditions for optimal plant growth. Using Arduino ADK 2560, the system tracks eight key parameters, temperature, humidity, light intensity, soil moisture, water level, rain, dust, and smoke through various sensors. Real-time data is transmitted to the Blynk app via ESP8266 module, allowing remote monitoring. The system automatically activates water pumps, sprinklers, and ventilation shutters based on preset thresholds. Powered by solar energy, this self-sustaining approach reduces human intervention while improving resource efficiency and crop yields.

Keywords. IoT, Greenhouse, Self-sustainable, Wireless Sensor Network, Agriculture.

1. INTRODUCTION

Greenhouse farming traditionally required constant human effort to maintain perfect growing conditions. The emergence of Wireless Sensor Networks (WSNs) has revolutionized this by enabling intelligent, self-sustaining agriculture [1]. Through strategically placed sensors monitoring eight critical parameters, temperature, humidity, light intensity, and soil moisture, farmers gain real-time visibility into their crops' environment [2]. This data-driven approach forms the foundation of modern smart farming, where conditions are automatically optimized for specific plants. By integrating Internet of Things (IoT) technology, manual control transforms into automated response systems that maintain ideal conditions without human presence [3]. This seamless combination of WSNs and IoT empowers farmers with precise, data-based cultivation techniques that adapt to both indoor and outdoor environmental changes [4, 5].

2. SYSTEM ARCHITECTURE

The proposed self-sustaining greenhouse operates through an integrated automatic system built around the Arduino ADK 2560 microcontroller. The setup continuously monitors eight critical environmental parameters using different sensors placed both inside and outside the

greenhouse structure. Inside the greenhouse, soil moisture levels are tracked using YL-69 sensors while temperature, humidity, and light intensity are measured through HT-CdS sensors. An ultrasonic sensor keeps watch over water tank levels to ensure uninterrupted irrigation supply. Outside the greenhouse, rain sensors and dust/smoke detectors monitor external conditions that could potentially harm the crops. The system follows a simple yet effective working principle where sensor readings are constantly compared against pre-set threshold values stored in the Arduino program. When any parameter falls outside the desired range, the controller automatically activates the necessary response mechanisms without requiring human intervention. For instance, if soil moisture drops below 50%, the irrigation system turns on, and when water tank levels go below 40cm, the water pump starts refilling until it reaches 90cm. All sensor data gets transmitted in real-time to the Blynk mobile application through an ESP8266 Wi-Fi module, allowing farmers to monitor their greenhouse conditions remotely from anywhere. The entire prototype runs on solar power with battery backup, making it truly self-sustainable. Testing results confirmed that all sensors performed accurately according to their threshold settings, successfully controlling the actuators representing sprinklers, shed covers, and water pumps. The system effectively maintains optimal growing conditions by automatically responding to changes in soil moisture, rain detection, dust levels, and water availability. This integrated approach demonstrates how automation can replace manual monitoring while improving resource efficiency and crop protection. The automated process is explained with the help of the flow chart portrayed in Figure 2.1.

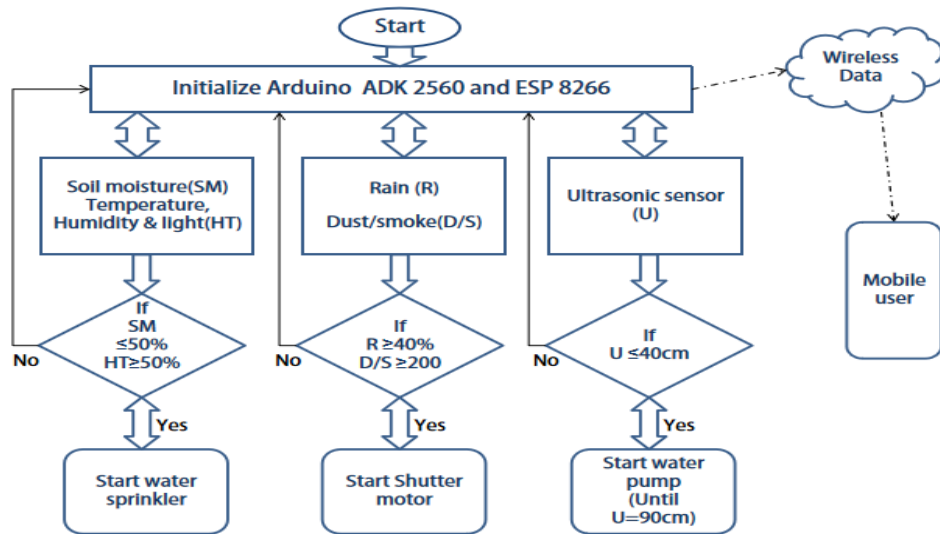


Figure 2.1. Flowchart of Prototype.

3. RESULTS AND DISCUSSION

The proposed self-sustaining greenhouse system successfully monitored and controlled eight environmental parameters through its sensor network. The Arduino ADK 2560 continuously processed data from all sensors and triggered appropriate responses based on programmed threshold values. The system demonstrated effective environmental control

through automated actuator responses. When soil moisture dropped below 50%, the irrigation system activated until optimal moisture levels were restored. Similarly, the retractable shutter mechanism responded to outdoor conditions, protecting crops during rainfall when sensor readings exceeded 40% and during high pollution events when dust particle concentration crossed $200 \mu\text{g}/\text{m}^3$ [1]. Water management was automated using an ultrasonic sensor that maintained tank levels between 40cm and 90cm, ensuring continuous irrigation supply [2]. Temperature, humidity, and light intensity remained under constant monitoring, providing real-time growing condition data [3]. The integration of renewable energy through solar panels made the system truly self-sustaining, with battery backup ensuring uninterrupted operation. Remote monitoring via the Blynk application allowed farmers to track all parameters from anywhere, eliminating the need for constant physical presence [4]. The automated system successfully replicated the decision-making process of human operators while responding faster to environmental changes. This approach to greenhouse management demonstrates how IoT technology can optimize growing conditions, reduce resource consumption, and minimize human intervention in agricultural operations [5].

4. CONCLUSION

This research successfully developed an Arduino-based self-sustaining greenhouse system that automatically monitors and controls eight environmental parameters. The prototype effectively manages soil moisture, temperature, humidity, light, water levels, rain, and air quality without human intervention. By maintaining optimal growing conditions, such systems can boost crop yields by up to 40% compared to traditional methods. This affordable, automated solution offers farmers a practical way to protect valuable crops while reducing water and energy waste.

5. REFERENCES

- [1] Priyadarshana, K. H. H., Manchanayaka, M. A. L. S. K., & Sudantha, B. H. (2020, March). IoT based greenhouse system for tropical countries. In 2020 International Conference on Image Processing and Robotics (ICIP) (pp. 1-5). IEEE.
- [2] Puranik, V., Ranjan, A., & Kumari, A. (2019, April). Automation in agriculture and IoT. In 2019 4th international conference on internet of things: smart innovation and usages (IoT-SIU) (pp. 1-6). IEEE.
- [3] Vimal, P. V., & Shivaprakasha, K. S. (2017, July). IOT based greenhouse environment monitoring and controlling system using Arduino platform. In 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT) (pp. 1514-1519). IEEE.
- [4] Yadav, P., Thakur, P., & Bansal, R. C. (2024). A new method for the characterization and type detection of voltage sag using active voltage component. *International Journal of Modelling and Simulation*, 45(5), 1771–1784. <https://doi.org/10.1080/02286203.2023.2301209>.
- [5] Kalyan, CH. N. S., Goud, B. S., Kumar, M. K., Thakur, P., Bajaj, M., & Bansal, R. C. (2023). Squirrel search algorithm based intelligent controller for interconnected power system. *International Journal of Modelling and Simulation*, 45(2), 420–440. <https://doi.org/10.1080/02286203.2023.2205989>.

Biographies



Ankur Kumar received the master's degree in instrumentation and control engineering from SLIET in 2021 and pursuing the philosophy of doctorate degree in school of computing and electrical from IIT Mandi, respectively. His research areas include artificial intelligence, machine learning, deep learning, biomedical engineering, electrical engineering, and signal processing.



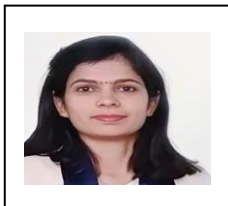
Anchal Sharma received the master's degree in electronics and communication engineering from SLIET in 2019 and pursuing the philosophy of doctorate degree in school of computing and electrical from IIT Mandi, respectively. Her research areas include artificial intelligence, machine learning, deep learning, biomedical engineering, electrical engineering, and signal processing.



Sanjay Dhanka received the master's degree in electrical engineering from UEM Jaipur in 2019, and the philosophy of doctorate degree in Electrical and Instrumentation Engineering from SLIET Longowal in 2025, respectively. He is currently working as an Assistant Professor at the Department of Electrical Engineering, Graphic Era Deemed to be University, Dehradun. His research areas include artificial intelligence, machine learning, deep learning, biomedical engineering, electrical engineering, and signal processing.



Abhinav Sharma received the master's degree in instrumentation and control engineering from SLIET in 2023 and pursuing the philosophy of doctorate degree in Electrical and Instrumentation Engineering from SLIET, respectively. His research areas include artificial intelligence, machine learning, deep learning, biomedical engineering, electrical engineering, and signal processing.



Monika received the bachelor's degree from CRSU, Jind in 2021, the master's degree in mathematics from SLIET in 2023, and pursuing the MTech in Computer Science and Engineering from IIT Patana, respectively. Her research areas include artificial intelligence, machine learning, deep learning, biomedical engineering, mathematics, and signal processing.

Vehicle Prioritization and Modified GPSR for Congestion-Adaptive Routing in VANETs

Gagandeep Berar¹, Sanjay Dhanka^{2*}, Ajit Noon³, Manish Kumar Singla⁴, Sneha Sriwastava⁵, Ramesh Kumar⁶

^{1,5}Chitkara University Institute of Engineering & Technology, Chitkara University, Punjab, India,

²Department of Electrical Engineering, Graphic Era Deemed to be University, Clement Town, Dehradun, India,

^{3,6}Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur, Rajasthan, 303007, India,

⁴Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India-602105,

¹gagandeep2913@gmail.com, ²sanjaykumar506070@gmail.com,

³ajit009noonia@gmail.com, ⁴msingla0509@gmail.com,

⁵sneha1057.be23@chitkara.edu.in, ⁶rameshkumarmeena@gmail.com

***Corresponding Author**

Abstract

This study presents a new congestion-adaptive prioritization algorithm in order to address the performance deterioration of Vehicular Ad-hoc Networks (VANETs) due to the dense vehicle traffic and traffic congestion, targeting the urban settings with automation of traffic lights. The main contribution is a vehicle prioritization-based communication optimization algorithm that makes routing behaviour and overall routing network abilities much higher due to the extensive analysis of communication parameters. The proposed model using NS2 as its simulator assigns vehicle priorities according to network level information with high-importance nodes having higher priorities, followed by a modification of the GPSR protocols to produce optimized, load-balanced communication paths using shortest path, node priorities, and network load to avoid communication failures.

Keywords. Vehicular Ad-hoc Networks, Congestion Control, Prioritization, GPSR Protocol, Network Simulation, NS2, Communication Optimization.

1. INTRODUCTION

Vehicular Adhoc network is the infrastructure-less network distributed in larger geographical space. It is a self-organized network with a specification of some infrastructure units and the scenario. The nodes in the network are vehicles with variant mobility within the scenario restriction [1]. The mobility, communication and architectural control are defined using specialized infrastructure specifications localized within the geographical

region. The technological requirement and sensing technologies are also defined under the vehicle, infrastructure, and geographical diversities. The traffic level, accidental situations, and environmental changes affect the network performances, priorities and challenges. The application, road, environment, safety, and traffic specific challenges exist in this network in a real scenario [2]. The network actually combines the features, restrictions, and challenges of mobile and sensor networks in diverse environments. The message, communication, and architectural specific constraints are defined in the network [1,3].

2. RESEARCH METHODOLOGY

The main issue that has been discussed in this study is the deterioration of vehicular network performance as a result of high density and traffic jams which results in a direct way to a rise in communication delay and loss of packets. As a response to this, the study proposes an adaptive prioritisation approach to congestion-sensitive optimisation of the vehicular network performance, particularly during urban operation where traffic lights automation is a feasible control system.

3. PROPOSED MODEL

The complete functional flow, as shown in Figure 1, is specific to the network scenario where architectural specifications, including the placement and role of Roadside Units (RSUs), are defined for a city environment. The traffic information sharing and load distribution are managed by these controller devices, which also handle the characterization of vehicle nodes and information transition via V2I communication.

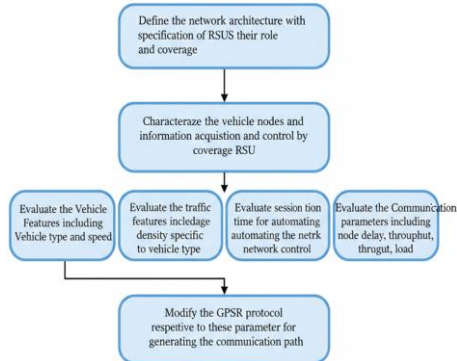


Figure 1. Flowchart of the Step Process from Network Setup to Performance Analysis.

4. RESULTS AND DISCUSSION

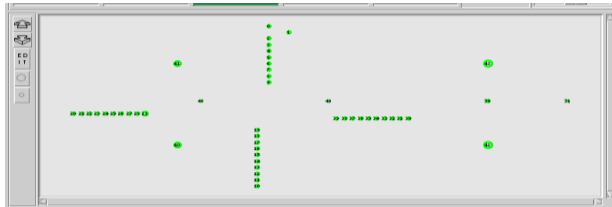


Figure 2. Generation of Vehicular Network.

In this section, the simulation outcome of the defined network and the algorithmic model proposed in this research are provided. Figure 2 and Figure 3 show the static network composed in this research for evaluating the operation of the suggested approach. The diagram shows the simulation environment of the network animator. The main window and the sidebars are shown.

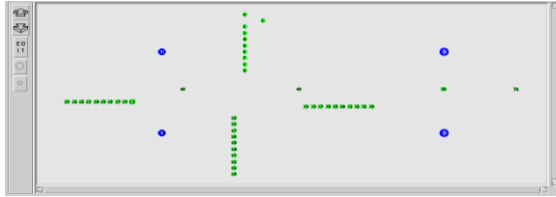


Figure 3. Activation of Network Components.

4.1. Analysis of Results

To validate the significance, operation, and reliability of the suggested prioritization model, the comparative analysis is done for the city network. The first parameter considered for analytical evaluating is the number of packets communicated over the network. The packet communication-based evaluation is indicated in Figure 4 and Table 5.

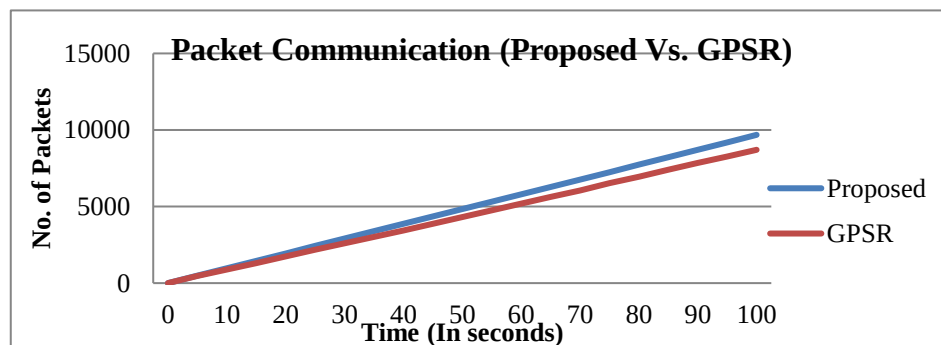


Figure 4. Packet Communication Analysis (Scenario I).

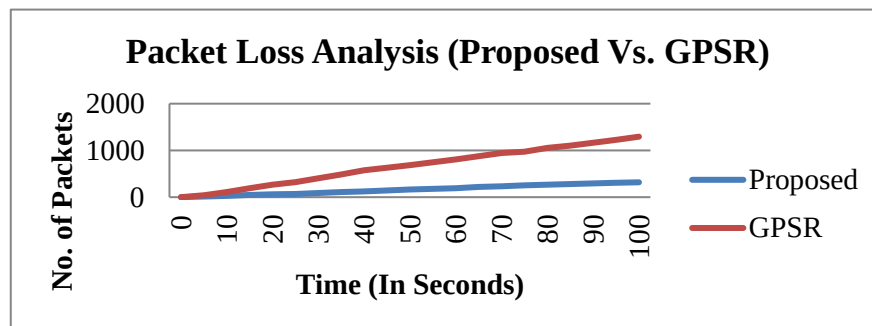


Figure 5. Packet Loss Analysis.

5. CONCLUSION

This research successfully developed and validated a novel congestion-adaptive prioritization method for Vehicular Ad-hoc Networks (VANETs), specifically targeting performance degradation in high-density urban environments. The proposed model can greatly increase the reliability and effectiveness of communication, as it introduces the two-stage model in which the vehicle priorities are assigned on the basis of network information and integrated into an adjusted GPSR protocol to generate optimal routes.

6. REFERENCES

- [1] Abbad, S. A., & Godse, S. P. (2016). Priority based emergency message forwarding scheme for time critical models in VANET. 393–398. <https://doi.org/10.1109/icaecct.2016.7942620>.
- [2] Sinha, R., Thakur, P., Gupta, S. et al. Development of lightweight intrusion model in Industrial Internet of Things using deep learning technique. *Discov Appl Sci* 6, 346 (2024). <https://doi.org/10.1007/s42452-024-06044-4>.
- [3] Yadav, P., Thakur, P., & Bansal, R. C. (2024). A new method for the characterization and type detection of voltage sag using active voltage component. *International Journal of Modelling and Simulation*, 45(5), 1771–1784. <https://doi.org/10.1080/02286203.2023.2301209>.

A Robust Metaheuristic Approach for Parameter Extraction of Complex Solar Cell Models: Validation on Industrial PV Panels

Shaffu¹, Shruti Arora², Ramesh Kumar³, Manish Kumar Singla⁴,

Sakshi⁵, Sanjay Dhanka^{6*}

^{1,2,5}Chitkara University Institute of Engineering & Technology, Chitkara University,
Punjab, India,

³Department of Electronics Communication and Engineering, Manipal University Jaipur,
Jaipur, Rajasthan, India,

⁴Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of
Medical and Technical Sciences, Chennai, India-602105,

⁶Department of Electrical Engineering, Graphic Era Deemed to be University, Clement
Town, Dehradun, India,

¹shaffu@chitkara.edu.in, ²shruti.arora@chitkara.edu.in,

³rameshkumarmeena@gmail.com, ⁴msingla0509@gmail.com,

⁵sakshi1039.be23@chitkara.edu.in, ⁶sanjaykumar506070@gmail.com

***Corresponding Author**

Abstract

The Estimating the performance of solar panels (PV systems) requires highly accurate modelling. Standard optimization methods often fail to find the best settings (parameters) for complex models like the three-diode (3D) and four-diode (4D) equivalent circuits, getting stuck in wrong answers (local minima). To fix this, we developed a new, improved optimization algorithm. This new method proved superior and faster than existing ones like Particle Swarm Optimization (PSO), Reptile Search Algorithm (RSA), and Crow Search Algorithm (CSA), achieving remarkably low error rates: an RMSE of 3.47E-16 and for the 3D model and 5.06E-16 for the 4D model.

Keywords. Photovoltaic Systems, Parameter Estimation, Three-Diode Model, Four-Diode Model, Metaheuristic Optimization, Crow Search Algorithm.

1. INTRODUCTION

The global transition to renewable energy sources has caused solar photovoltaic (PV) technology to become a fundamental part of the future energy mix. The assessment of PV systems requires accurate simulation and modelling, which in turn requires a well-defined solar cell standard [1]. The working of a solar cell from an electrical perspective is exclusively influenced by the internal values which are referred to as parameters. The most important parameters include photocurrent, saturation currents, series and shunt resistances

(which characterize internal losses), and idealization factors. If these values are very accurately known, one can predict the exact performance of the cell under various real-world conditions [2]. On the other hand, the parameter values listed in the manufacturer's datasheet are generally insufficient for the construction of high-fidelity models, which leads to a considerable number of difficulties since the output of the simulated system will not be the same as that of the actual one [3].

2. ALGEBRAIC MODELLING OF SOLAR PV CELLS

The three-diode model illustrated in Figure 1 introduces a third diode to more accurately capture recombination losses occurring at grain boundaries within semiconductor materials, a phenomenon that is not fully accounted for by simpler models [4]. The governing equation for the output current (I_O) in the 3D model is given in Figure 1, which extends the DD model by adding a third diode current term.

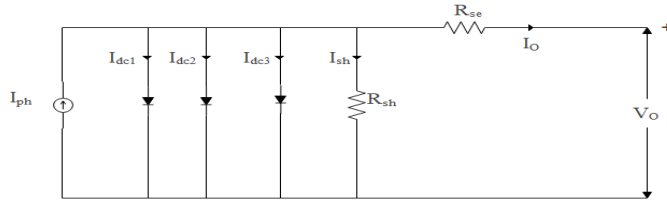


Figure 1. Simplified 3-Diode PV Circuit.

3. PROPOSED ALGORITHM: ENHANCED CROW SEARCH ALGORITHM (ECSA)

The new optimization method we are proposing is built upon an improved version of the. The CSA is a metaheuristic technique that draws inspiration from the clever food-finding (foraging) strategies of crows. Crows are renowned for their extraordinary cognitive abilities, including self-awareness, tool making, and sophisticated social strategies such as hiding food and stealing other birds' stores while protecting their own. Although the standard CSA is efficient, it might face the challenge of unequal exploration and exploitation.

4. EXPERIMENT AND RESULTS

The proposed metaheuristic algorithm was subjected to a rigorous evaluation in this trial phase against the already existing algorithms PSO, Reptile Search Algorithm (RSA) and CSA by means of five selected benchmark test functions (CEC1-CEC5) of CEC 2019 which all had 20 parameters. To avoid any favouritism in the comparison all algorithms were limited to 1000 feature estimations per test task. Programming was carried out in MATLAB 2018b, each algorithm run independently for 30 iterations to obtain reliable arithmetical outcomes as shown in Table 1 and 2 and Fig. 2 and 3.

Table 1. Unknown Parameters of Three Diode Model.

Parameter/Algorithms	Proposed Algorithm	CSA	PSO	RSA
I_{pv}	5.8587	6.9159	5.2835	5.8224

n_3	1.261	0.947	0.58	1.135
R_{sh}	158.2	155.13	97.451	273.09
I_{02}	3.80E-07	4.51E-07	1.16E-07	2.78E-07
n_1	1.081	0.714	0.5	1.65
n_2	1.161	0.701	0.668	1.213
I_{01}	2.69E-07	2.91E-08	0	7.31E-08
R_s	0.077	0.017	0.001	0.096
I_{03}	6.28E-07	2.82E-09	7.70E-09	6.28E-08
RMSE	3.47E-16	3.31E-04	1.57E-02	4.33E-06

Table 2. Unknown Parameters of Four Diode Model.

Parameter/ Algorithms	Proposed Algorithm	CSA	RSA	PSO
n_1	1.085	1.567	0.76	0.538
n_2	1.214	1.072	1.026	0.584
R_s	0.087	0.02	0.023	0.024
I_{01}	2.55E-07	2.10E-07	5.24E-07	4.86E-09
I_{02}	4.87E-07	0	3.12E-07	2.48E-08
n_3	1.089	0.612	1.286	0.5
n_4	1.053	0.849	1.598	0.545
I_{pv}	7.8285	6.8896	3.8102	6.1342
R_{sh}	220.96	153.43	301.84	222.8
I_{03}	3.72E-07	6.60E-07	3.97E-07	0
I_{04}	2.65E-07	0	4.21E-07	3.58E-08
RMSE	5.06E-16	2.97E-04	5.97E-06	2.25E-01

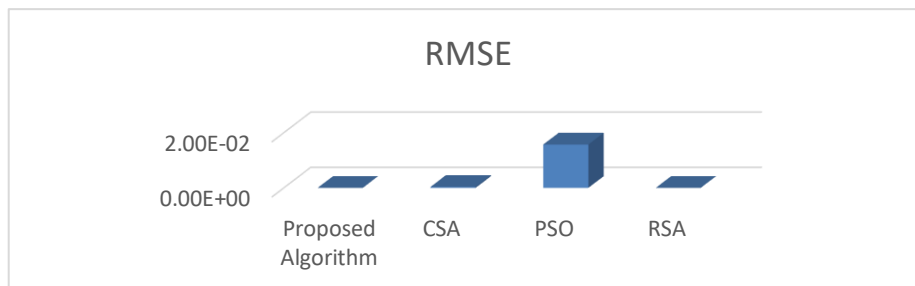


Figure 4.1. RMSE of Three Diode Model.

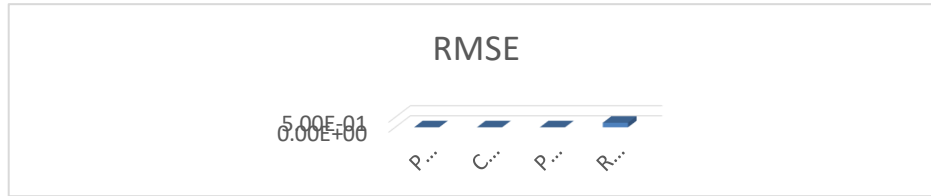


Figure 4.2. RMSE of Four Diode Model.

5. CONCLUSION

This research study introduces a brand-new algorithm created specifically to solve the challenging problem of global optimization when determining the operating parameters of a solar cell. The core innovation of this algorithm is the ability to interpret temperature variations, which increases the diversity of the solution. Moreover, it features an opposition-based learning approach, which is a technique for considering and filtering all possible solutions from a wide range.

6. REFERENCES

- [1]. Premkumar, M., Chandrasekaran, K., & Sowmya, R. (2020). Mathematical modelling of solar photovoltaic cell/panel/array based on the physical parameters from the manufacturer's datasheet. *International Journal of Renewable energy development*, 9(1), 7.
- [2]. Alam, D. F., Yousri, D. A., & Eteiba, M. B. (2015). Flower pollination algorithm based solar PV parameter estimation. *Energy Conversion and Management*, 101, 410-422.
- [3]. N. Rawat, P. Thakur, A. Dixit, K. Charu and S. Goyal, "Grey Wolf Optimisation Based Modified Parameter Estimation Technique of Solar PV," 2024 IEEE Third International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, India, 2024, pp. 755-760, doi: 10.1109/ICPEICES62430.2024.10719261.
- [4]. Yadav, P., Thakur, P., & Bansal, R. C. (2024). A new method for the characterization and type detection of voltage sag using active voltage component. *International Journal of Modelling and Simulation*, 45(5), 1771–1784. <https://doi.org/10.1080/02286203.2023.2301209>.

An Enhanced Artificial Hummingbird Algorithm Outperforms Established Metaheuristics in DMFC Parameter Optimization

Shaffu¹, Shruti Arora², Ramesh Kumar³, Manish Kumar Singla⁴,

Pawan Kumar Pandey⁵, Sanjay Dhanka^{6*}

^{1,2}Chitkara University Institute of Engineering & Technology, Chitkara University,
Punjab, India,

³Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur,
Rajasthan, India,

⁴Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of
Medical and Technical Sciences, Chennai, India-602105,

⁵Department of Electrical Engineering, Chandigarh Engineering College University,
Mohali, Punjab, 140307, India,

⁶Department of Electrical Engineering, Graphic Era Deemed to be University, Clement
Town, Dehradun, India,

¹shaffu@chitkara.edu.in, ²shruti.arora@chitkara.edu.in,

³rameshkumarmeena@gmail.com, ⁴msingla0509@gmail.com, ⁵pkpntij@gmail.com,

⁶sanjaykumar506070@gmail.com

***Corresponding Author**

Abstract

Identification of the parameters is important to creation of precise performance models of direct methanol fuel cells (DMFCs) because some of the core internal parameters are not always covered in datasheets of manufacturers. This investigation introduces Enhanced Artificial Hummingbird Algorithm (EAHA) as an algorithm that is used to estimate the parameters that are not known, by minimization of the squared error (SSE) of the predicted and observed voltages. The efficiency of EAHA is strictly examined by means of comparative analysis with five known metaheuristic algorithms, namely: Particle Swarm Optimization (PSO), Dragonfly Algorithm (DA), Harris Hawks Optimization (HHO), Reptile Search Algorithm (RSA) and the original Artificial Hummingbird Algorithm (AHA). These findings demonstrate that EAHA has much reduced SSE (1.01E-10) and reduced computation time (0.127 seconds) compared to normal operating conditions and the reason behind its high accuracy and efficiency.

Keywords. Direct Methanol Fuel Cell (DMFC); Parameter Identification; Enhanced Artificial Hummingbird Algorithm (EAHA);

1. INTRODUCTION

The direct methanol fuel cells (DMFCs) technologies are gaining relevance due to the demand for clean energy around the globe. The principal advantage of DMFCs is their application in mobile power and combined heat and power systems, primarily due to their high energy efficacy, portability, safe refuelling, and low operating temperature characteristics, which are the main benefits. Nonetheless, they are constrained by critical issues, especially slow reactions, inactivation of catalysts, and acute mass transport issues [1-2]. The main guilty party to these transport issues is based on poor design of flow channels that deliver the fuel. Specifically, poor design of flow channels results in a lack of distribution of fuel and concentration of bubbles of waste gas (primarily CO₂) in the gas diffusion layer (GDL) of the cell. This is called gas blockage that starves the catalytic layer of methanol fuel, seriously impairs the performance of the cell, and leads to tremendous loss of power [3-5].

2. FUEL CELL MOLDING

Fuel Cells Fuel cells are the devices that carry out the electrochemical reaction that produces electrical energy from the chemical energy, with the continuous supply of fuel and oxidizing agent to the system. Most importantly, they produce power and do not store power, like batteries; The main products of them are water and electricity.

3. BEING PROPOSED: ENHANCED ARTIFICIAL HUMMINGBIRD ALGORITHM (EAHA)

Artificial hummingbird algorithm (AHA) is a metaheuristic optimization algorithm that is centred on the remarkable memory and foraging behaviour of the hummingbirds. Hummingbirds are able to recall the position of food resources as well as the rate and duration of time needed to forage between the two sources of nectar, leading to very efficient foraging.

4. RESULT AND DISCUSSION

In order to determine the strength of the suggested Enhanced Artificial Hummingbird Algorithm (EAHA), the authors conducted an elaborate comparison of the main benchmark optimizers: Particle Swarm Optimization (PSO), Dragonfly Algorithm (DA), Harris Hawks Optimization (HHO), Reptile Search Algorithm (RSA) and the original Artificial Hummingbird Algorithm (AHA). This assessment was made on basis of DMFC technical specifications in Table .1 and the optimization constraints which were in Table 1 and Figure 1. To statistically verify, each of the algorithms was applied 40 times in MATLAB 2020a.

Table 1 Parameter Estimation of the Double Line-Fed, Direct Driven Induction Machine.

Parameters /Algorithm s	α	C_1	β	R	j_{eid}	e_o (V)	SSE	r_{eq}	Computatio n Time (Sec)
-------------------------------	----------	-------	---------	-----	-----------	-----------	-----	----------	----------------------------

Proposed Algorithm	0.785 9	- 0.452 8	1.302 4	1.052 1	0.254 1	0.745	1.01E-10	27.30 1	0.127
DA	0.552 1	- 0.250 9	0.548 1	2.054 1	0.142 8	0.721 5	1.21E-04	24.52 4	1.198
PSO	0.452 7	- 0.783 1	0.214 5	1.845 2	0.201 5	0.759 4	1.00E-03	15.62 4	2.014
RSA	0.365 1	- 0.253 4	1.054 7	1.351 4	0.201 7	1.021 8	1.31E-06	17.96 4	1.247
AHA	0.257 1	- 0.815 4	0.965 4	1.842 1	0.187	0.752 4	1.55E-08	25.32 4	1.001
HHO	0.489	- 0.742 1	0.652 4	2.652 7	0.153 4	1.635 7	1.15E-05	19.05 1	1.547

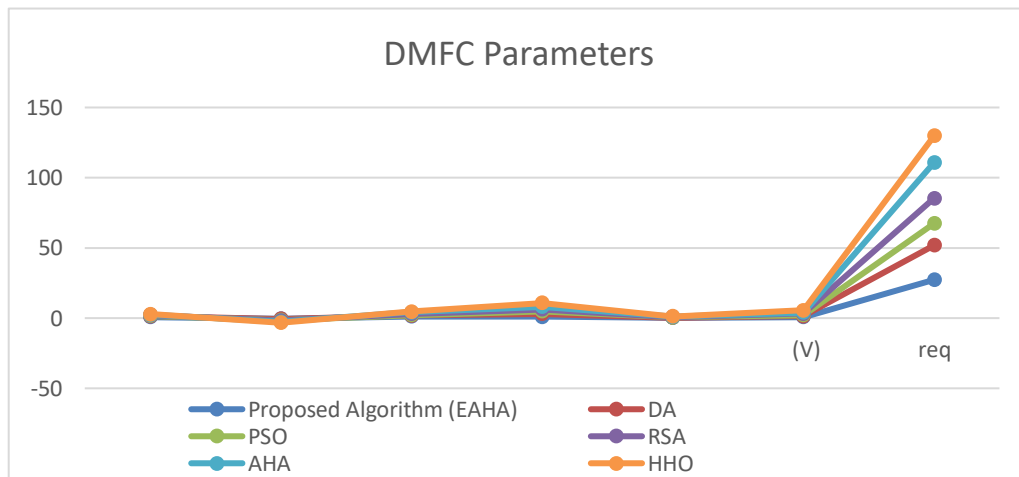


Figure 4.1. Estimation of DMFC Parameters Using Various Algorithms.

5. CONCLUSION

In this investigation, an EAHA was developed and tested in order to discover the key factors in DMFCs in an accurate manner. The EAHA functioning was strictly checked and achieved high results in comparison with a lot of popular metaheuristic algorithms, particularly, PSO, HHO, DA, RSA, and the original standard AHA. According to the results, EAHA demonstrates an overwhelming advantage in accuracy and computation efficiency. Under typical conditions it gave a very small value of squared error (SSE) of $1.01\text{E-}10$ and was found to converge with solutions with much greater accuracy than any other algorithm across a variety of operating temperatures. This performance supports the fact that it can produce highly accurate DMFC models that highly replicate experimental behaviour.

6. REFERENCES

- [1] Ke, Y., Yuan, W., Zhou, F., Guo, W., Li, J., Zhuang, Z., ... & Song, J. (2021). A critical review on surface-pattern engineering of nafion membrane for fuel cell applications. *Renewable and Sustainable Energy Reviews*, 145, 110860.
- [2]. Zhang, X., Yang, J., Ma, X., Zhuge, W., & Shuai, S. (2022). Modelling and analysis on effects of penetration of microporous layer into gas diffusion layer in PEM fuel cells: Focusing on mass transport. *Energy*, 254, 124103.
- [3]. Liang, J., Luo, Y., Zheng, S., & Wang, D. (2017). Enhance performance of micro direct methanol fuel cell by in situ CO₂ removal using novel anode flow field with superhydrophobic degassing channels. *Journal of Power Sources*, 351, 86-95.
- [4]. N. Rawat, P. Thakur, A. Dixit, K. Charu and S. Goyal, "Grey Wolf Optimisation Based Modified Parameter Estimation Technique of Solar PV," 2024 IEEE Third International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, India, 2024, pp. 755-760, doi: 10.1109/ICPEICES62430.2024.10719261.
- [5]. Kalyan, CH. N. S., Goud, B. S., Kumar, M. K., Thakur, P., Bajaj, M., & Bansal, R. C. (2023). Squirrel search algorithm based intelligent controller for interconnected power system. *International Journal of Modelling and Simulation*, 45(2), 420–440. <https://doi.org/10.1080/02286203.2023.2205989>.

An IoT-Enabled Smart Cap for the Blind: Echo Vision with Obstacle Detection, Haptic Feedback & Emergency SOS

**Ramesh Kumar¹, Sanjay Dhanka^{2*}, Manish Kumar Singla³, Rajveer Saini⁴, Parth
Bhandari⁵**

¹*Department of Electronics Communication and Engineering, Manipal University Jaipur,
Jaipur, Rajasthan, India,*

²*Department of Electrical Engineering, Graphic Era Deemed to be University, Clement
Town, Dehradun, India,*

³*Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of
Medical and Technical Sciences, Chennai, India-602105,*

^{4,5}*Chitkara University Institute of Engineering & Technology, Chitkara University,
Punjab, India,*

¹*rameshkumarmeena@gmail.com, *²sanjaykumar506070@gmail.com,*

³*msingla0509@gmail.com, ⁴rajveer1024.be23@chitkara.edu.in,*

⁵*parth1009.be23@chitkara.edu.in*

***Corresponding Author**

Abstract

Echo Vision is an IoT-based wearable assistive device designed to enhance navigation, safety, and independence for visually impaired individuals. The system integrates embedded IoT (Internet of things) technologies into a smart cap, utilizing ultrasonic sensors for obstacle detection and an ADXL345 accelerometer to monitor head movement and tilt, enabling contextual awareness. Environmental data is translated into real-time feedback through tactile (vibration motors) and auditory (buzzers) alerts, ensuring intuitive user interaction. The heart of the system such as the ESP8266, IoT module enables wireless connectivity, facilitating seamless communication with a dedicated mobile app.

Keywords. Echo Vision, IoT, GPS, SOS, Wearable Assistive Device.

1. INTRODUCTION

In today's world, millions of visually impaired individuals face daily challenges in navigating their surroundings safely and independently. Traditional aids like white canes and guide dogs, while helpful, have limitations canes are restricted to detecting obstacles only at ground level, and guide dogs require extensive training and are not affordable or accessible for everyone [1]. This lack of reliable, real-time environmental awareness exposes visually impaired people to risks such as tripping, collisions, or getting lost,

particularly in unfamiliar or crowded urban settings. Moreover, in emergency situations, visually impaired individuals may find it difficult to communicate their location quickly, delaying critical assistance. With limited access to affordable, tech-driven assistive solutions, there remains a significant gap in empowering this segment of the population through modern innovation [2]. It's amazing how much wearable technology and integrated systems are changing healthcare and making life easier for people with disabilities. One really important area where this makes a difference is in helping people who are visually impaired. Right now, getting around can be a big obstacle for people with visual impairments, especially in places that were not built with their needs in mind. While tools like white canes and guide dogs are useful, they have their drawbacks. For example, white canes can't detect obstacles at head height, and both require physical effort. Guide dogs are also expensive to get and maintain [3-4].

2. HARDWARE & SOFTWARE

This section outlines the hardware components, software tools, and implementation framework of our wearable assistive technology project, Echo Vision: Auditory Vision for the Blind. The project integrates sensor-based obstacle detection, motion tracking, tactile alerts, and real-time GPS communication into a compact, wearable system. The core idea is to enhance navigation for visually impaired users, powered by the ESP8266 microcontroller and supported by various sensors and IoT tools. This chapter also includes block diagrams, circuit diagrams, pin diagrams, and flowcharts to explain the structure and logic of the project.

3. SOFTWARE TOOLS & LIBRARIES

The Arduino IDE (Integrated Development Environment). This is our main workspace for writing the instructions and code that tell the ESP8266 NodeMCU (the brain of the device) what to do. For real-time monitoring and control, we rely on the Blynk platform. This clever tool helps us create the mobile app that lets caregivers keep an eye on the user's status from a distance. It's all about making sure help is always just a tap away, enhancing safety and peace of mind.

Emergency features are automated through Macro Droid, which triggers smartphone actions, such as sending the user's live location via email and Telegram bot when the SOS button is pressed. Key libraries include ESP8266WiFi.h and WiFi ClientSecure.h for Wi-Fi connectivity, UniversalTelegramBot.h and ArduinoJson.h for Telegram notifications, and Adafruit_ADXL345_U.h with Wire.h for accelerometer data processing. Together, these components create a robust, interconnected IoT ecosystem for assistive navigation.

4. BLOCK AND CIRCUIT DIAGRAM

The IoT Emergency Alert System integrates multiple hardware and software components to create a comprehensive assistive solution. At the sensor level, an HC-SR04 ultrasonic sensor detects obstacles, while an ADXL345 accelerometer monitors motion and tilt, and an SOS button triggers emergency alerts. These components connect to the ESP8266 microcontroller, which processes data and activates output devices like a buzzer and

vibration motor for user feedback while leveraging its built-in Wi-Fi module for connectivity, such as shown in Figure 1.

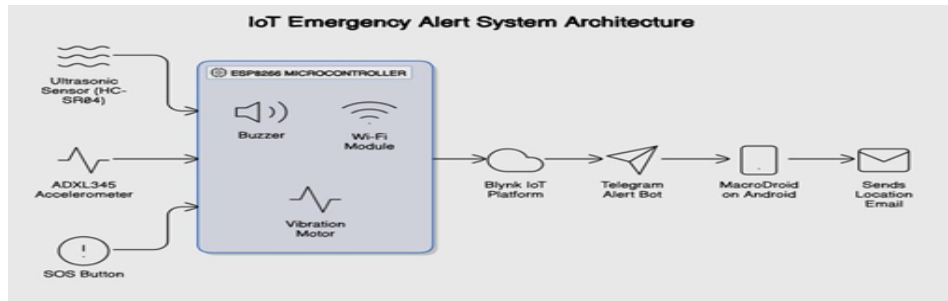


Figure 1. Block Diagram.

5. RESULTS AND DISCUSSION

The Echo Vision was successfully designed, developed, and tested as a wearable assistive system to help visually impaired individuals navigate their environment more safely and independently. The results from the implementation are promising and demonstrate that the device can perform real-time obstacle detection, fall monitoring, and emergency location sharing effectively, such as shown in Figure 3.

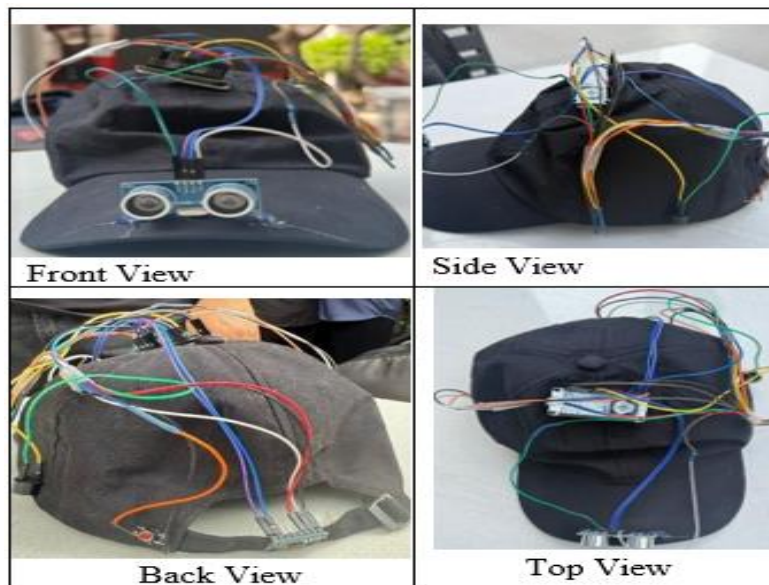


Figure 3. Final Echo Vision System.

6. CONCLUSION AND FUTURE SCOPE

The creation of Echo Vision: Auditory Vision for the Blind was a learning experience, and not only regarding the technical aspects of the creation of assistive wearable technology, but also regarding the significant social change that it can help to create. The primary objective was to assist the people with visual impairments to be more autonomous and have access to the surrounding world. This was achieved through the integration of sensors, real-time alerts as well as IoT.

7. REFERENCES

- [1] Farooq, M. S., Shafi, I., Khan, H., Díez, I. D. L. T., Breñosa, J., Espinosa, J. C. M., & Ashraf, I. (2022). IoT enabled intelligent stick for visually impaired people for obstacle recognition. *Sensors*, 22(22), 8914.
- [2] Kumar, A. R., Sanjay, K., & Praveen, M. (2023, December). EchoGuide: Empowering the Visually Impaired with IoT-Enabled Smart Stick and Audio Navigation. In 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 1770-1774). IEEE.
- [3] Sinha, R., Thakur, P., Gupta, S. et al. Development of lightweight intrusion model in Industrial Internet of Things using deep learning technique. *Discov Appl Sci* 6, 346 (2024). <https://doi.org/10.1007/s42452-024-06044-4>
- [4] N. Rawat, P. Thakur, A. Dixit, K. Charu and S. Goyal, "Grey Wolf Optimisation Based Modified Parameter Estimation Technique of Solar PV," 2024 IEEE Third International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, India, 2024, pp. 755-760, doi:

Cognitive Modeling of Smile Perception Using Reaction Time and PCA-Based Feature Selection

Abdullah Bin Queyam¹, Sanjay Dhanka^{2*}, Debasmita Ghosh Roy³, Ramesh Kumar⁴,
Manish Kumar Singla⁵

^{1,3}*School of Automation, Banasthali Vidyapith, Rajasthan, 304022, India,*

²*Department of Electrical Engineering, Graphic Era Deemed to be University, Clement Town, Dehradun, India,*

⁴*Chitkara University Institute of Engineering & Technology, Chitkara University, Punjab, India,*

⁵*Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India-602105,*

¹abdullahbinqueyam@banasthali.in, ^{2*}sanjaykumar506070@gmail.com,

³debasmitaghoshroy@banasthali.in, ⁴rameshkumarmeena@gmail.com,

⁵msingla0509@gmail.com

***Corresponding Author**

Abstract

The study describes a novel framework, which uses the human cognitive response data in terms of reaction times (RT) to simulate the perception of smile intensity. The authors transcend the use of traditional binary categorization by proposing a continuous smiling index (SI) based on human RTs and binary rating of a mixed 200-image dataset of faces. We also combine this psychophysical information with a machine learning pipeline that uses Principal Component Analysis (PCA) and feature selection in sequence in a unique manner. This enables us to produce artificial facial expression continua with perceptually based control smile differences. One of the contributions is the fact that RTs present a strong, implicit measure of expression ambiguity and intensity. Our model delivers more perceptually realistic smile progressions as compared to baseline methods through comparative evaluation. The model is proven to be effective through human validation experiments which identify a logarithmic pattern of perception that is consistent with psychological models. The proposed work advances facial expression analysis by: (1) establishing a cognitive data-driven continuous intensity metric, (2) demonstrating the untapped value of reaction time in expression modeling, and (3) providing an effective synthesis method for generating realistic facial expression continua for HCI and psychological research.

Keywords. Cognitive Modeling, Smile Perception, Reaction Time, Principal Component Analysis, Facial Expression Synthesis.

1. INTRODUCTION

One of the most important components of the process of sharing thoughts and social signals with people is facial expressions. Of all expressions smiling is the most researched one since it is evident in every culture with varying degrees of intensity. Calder et al. [1] were pioneers in applying PCA to the facial expression. They checked its fit in face memory models of the brain. Their conclusion is that PCA does not always do a good job of identifying tiny changes such as small smiles. This work was expanded by Beltran and Calvo [2], who examined the brain activity when smiling. They obtained PCA and brain wave data to follow the initial signs in the behind brain. Their findings revealed that the brain reacts very fast to smiles in a manner that emphasizes on significant changes. The reaction time or RT is also significant in the perception of smiles by people. Li et al. [3] applied brain tests to observe the variation in the speed or slowness of response on the way we detect smiles. They demonstrated that attempting to respond quickly may impair the capacity of the brain to detect gentle smiles. This will be significant in developing superior models of our face reading. Elsewhere Jones and Kramer [4] employed PCA and one more technique known as factor analysis. They discovered that smiles are likely to cause individuals to feel safe and open. These are the faces that are prominent in PCA. In Cowen et al. [5], PCA was used to reconstruct brain signal faces. Their work is based on the fact that PCA can be used to learn the manner in which the brain stores the description of faces such as smiling.

2. METHODOLOGY

The methodology section outlines the systematic approach employed in this study to investigate the "smiling strength" of faces using cognitive modeling techniques.

2.1. *Dataset Selection and Preprocessing*

The approach used in this experiment started with the close selection and pre-processing of the facial images in order to explore the phenomenon of smiling strength as both a dichotomous variable (smiling and not smiling), and as a continuous measure that will be able to detect the subtle variations of the facial expressions. The analysis targeted such important parameters of the face as smile line (curvature of the lips), smile arc (position of the teeth in relation to the lower lip), the shape of the upper lip, and the presentation of teeth that are included in such parameters to contribute to the apparent infertility of the smile. Calculation of Continuous Smile Index

2.2. *Feature Selection Using Sequential Forward Selection*

The criterion used to select the features was sequential forward selection to determine the most effective subset of principal components to estimate the continuous smile index (CSI). The main aim was to identify the most efficient combination of features derived using PCA, which would be the best predictor of the smile intensity and at the same time be parsimonious in nature. It was done through the SequentialFS function of MATLAB as it was used to apply an iterative selection method based on the performance measures. The algorithm selected the subsets by feature through mean squared error (MSE) of linear regression as the main performance measure that guaranteed that the features chosen would give rise to the lowest prediction error.

3. RESULTS AND DISCUSSION

The study investigated smile perception using cognitive modeling techniques on a curated dataset of 200 facial images (100 individuals, with smiling and neutral expressions). Images were standardized for background and orientation to minimize irrelevant variations. Reaction time data from three participants was processed to create a continuous smile index ranging from -1 (definite non-smile) to +1 (definite smile), after removing outliers (<200ms or >2s) and normalizing for individual response speeds.

Human validation with two participants confirmed all continua maintained monotonic smile progression as shown in Figure 1. The regression model's continuum showed superior logarithmic alignment with human ratings ($R^2=0.92$ vs 0.85/0.84 for alternatives), particularly in distinguishing subtle smile intensities. Table 1 summarizes a study on smile perception, finding that regression modeling best predicts human ratings by using mid-level PCA features.

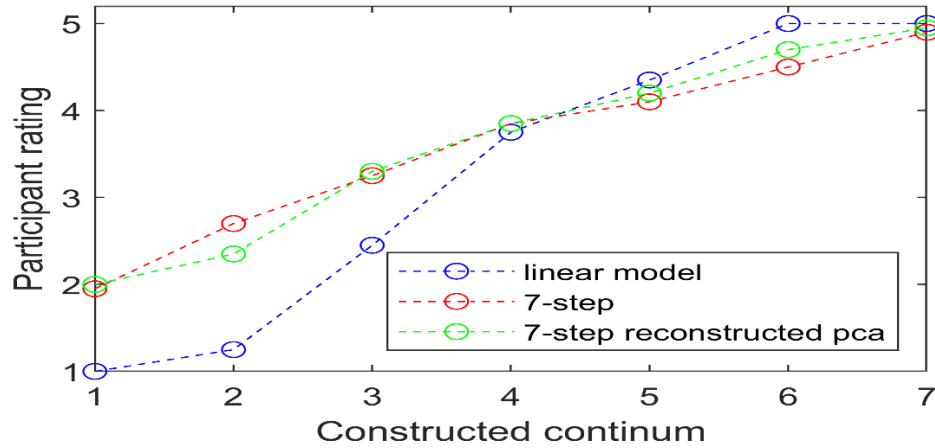


Figure 1. Human Perception of Constructed Continuum.

Table 1 Key Experimental Outcomes.

Component	Metric	Finding	Implication
Dataset	Image selection	200 images (100 pairs)	Controlled pose/variability balance
Smile Index	Distribution range	-1 to +1 continuum	Effective expression quantization
PCA	Variance explained (Top 3 PCs)	52.19% total	Gender/race dominate facial variance
Feature Selection	Optimal features	10 PCA components (up to PC47)	Smile perception uses mid-level features

Regression Model	Prediction accuracy (Human rating)	$R^2=0.92$	Best perceptual alignment
Continuum Generation	Methods compared	Regression vs. two averaging techniques	Regression preserves perceptual linearity
Human Validation	Rating consistency	All methods monotonic ($p<0.01$)	General method robustness

4. CONCLUSION

The research has been able to come up with a cognitive modeling method to measure and synthesize facial expression with respect to the intensity of smile. The best sources of facial variance with the aid of PCA were gender and race, whereas sequential feature selection was used to optimize the identification of smile-related features. A continuous smile index, which was obtained by adding reaction times and binary classifications, was useful in capturing perceptual differences between smiling and non-smiling faces. Training a linear regression model on these features yielded high-quality syntactic facial continua that were better than more basic weighted-averaging schemes on perceptual alignment.

5. REFERENCES

- [1] Abbad, Calder, A. J., A. M. Burton, P. Miller, A. W. Young, and S. Akamatsu. 2001. "A Principal Component Analysis of Facial Expressions." *Vision Research* 41, no. 9 (April): 1179–1208. [https://doi.org/10.1016/S0042-6989\(01\)00002-5](https://doi.org/10.1016/S0042-6989(01)00002-5).
- [2] Beltrán, D., and M. G. Calvo. 2015. "Brain Signatures of Perceiving a Smile: Time Course and Source Localization." *Human Brain Mapping* 36, no. 11 (August): 4287–4303. <https://doi.org/10.1002/hbm.22917>.
- [3] Yadav, P., P. Thakur, and R. C. Bansal. 2024. "A New Method for the Characterization and Type Detection of Voltage Sag Using Active Voltage Component." *International Journal of Modelling and Simulation* 45, no. 5: 1771–1784. <https://doi.org/10.1080/02286203.2023.2301209>.
- [4] Rawat, N., P. Thakur, A. Dixit, K. Charu, and S. Goyal. 2024. "Grey Wolf Optimisation Based Modified Parameter Estimation Technique of Solar PV." In 2024 IEEE Third International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, India, 755–760. <https://doi.org/10.1109/ICPEICES62430.2024.10719261>.

A REVIEW ON DETECTION OF GRASSERIE DISEASE IN SILKWORM

Preetam Sanjay Patil^{1*}, R M Rudresh Patil², Sharanprakash Hosamani³,
Soujanya D G⁴ and Hemanth T S^{1*}

*Department of Electronics and Communication Engineering, Alvas Institute of
Engineering & Technology, Moodbidri, Mangalore, Karnataka, India*

*Email ID: 4al22ec069@aiet.org.in, rmrudresh9@gmail.com,
sharanhosamani272@gmail.com, soujanyaadg23@gmail.com, hemanthts@aiet.org.in.*

Abstract.

The sericulture industry is currently undergoing a significant transformation through the adoption of the Internet of Things (IoT), Artificial Intelligence (AI), and advanced image processing techniques. These technologies are being deployed to combat prevalent silkworm diseases and modernize traditional, labor-intensive farming practices. IoT-enabled sensors provide farmers with immediate feedback on critical environmental metrics, such as humidity and temperature, which are vital for silkworm survival. Furthermore, computing platforms like the Raspberry Pi, when integrated with AI and image processing algorithms, facilitate the real-time diagnosis of diseases and the assessment of silkworm health with classification accuracies reaching up to 99%. The convergence of these technologies enhances system reliability and autonomy, leading to higher quality silk production and a reduction in manual labor. Current research indicates that such automated systems can sort cocoons at high speeds and mitigate yield losses caused by suboptimal rearing conditions, thereby fostering a more resilient and sustainable sericulture sector.

Keywords: Sericulture, Internet of Things (IoT), Artificial Intelligence (AI), Machine Learning (ML), Raspberry Pi, Disease Detection.

1. INTRODUCTION

Sericulture being an ancient but important sector, especially in countries such as India and China, where it is essential to maintain standard environmental parameters such as temperatures, humidity, and cleanliness for better silkworm rearing. The occurrence of diseases is a serious problem in these circumstances, which further leads to a decrease in the quality of cocoons and sometimes in financial setbacks, with manual methods being time-consuming and imprecise. The combination of IoT technology and image processing technology will help in continuous environmental parameter tracking, automatic disinfection and irrigation systems, and accurate identification of diseases in silkworms. Moreover, AI and machine learning algorithms in camera-based systems will help in quick analysis and early detection of diseases, which is not possible in other systems. Therefore, a

combination of advanced intelligent systems with existing pathological knowledge will help in increasing productivity, improving quality, and providing a bright future to sericulture.

2. IMPLEMENTATION AND ARCHITECTURE

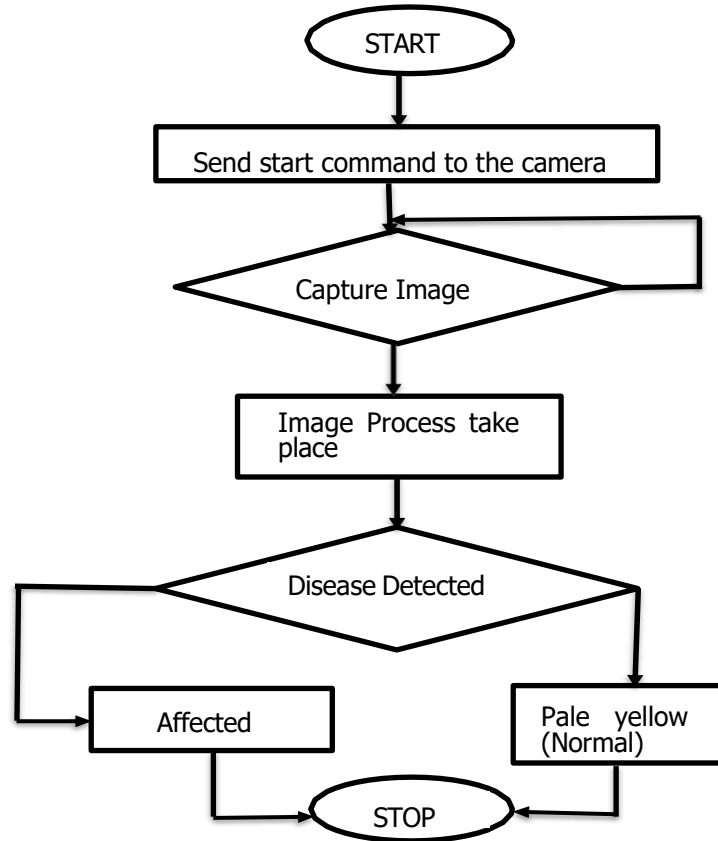


Fig 2.1: Pipeline of the Silkworm Detection

2.1 System Architecture Overview This architecture design is targeted to achieve silkworm disease detection as a sequential image-processing pipeline. The first stage focuses on image acquisition; the second stage does the disease analysis and classification.

- **Input Layer:** This layer captures silkworm images by the help of a camera or image upload mechanism.
- **Processing Layer:** The captured image is processed and analysed by the trained detection model in order to identify disease presence. This kind of lightweight architecture makes processing fast, deployment easy, and thus detects reliably in real-time applications of sericulture.

2. METHODOLOGIES

This survey synthesizes methodological advances across four interlocking stages:

- 3.1. Dataset Collection and Pre-processing:** The Image Processing Module is the core of the system, enabling accurate detection of Grasserie disease from silkworm images. Each image is resized to YOLO's fixed input dimensions to maintain uniformity across the dataset. Noise removal and pixel-level normalization are applied to enhance image quality and consistency. The pre-processed images are analysed by YOLO's convolutional layers to extract disease-related visual features. Multi-scale detection improves accuracy by identifying symptoms such as swelling, colour changes, and irregular textures.
- 3.2. User Interface:** The user interface acts as the main interaction platform for the YOLO-based detection system. Users can upload silkworm images easily through a simple and intuitive interface. Once uploaded, the image is processed in real time using the trained YOLO model. The system classifies the silkworm as healthy or Grasserie-affected with a confidence score. The interface is designed to be user-friendly and accessible for both technical and non-technical users.

Quantitative Analysis of Parameters

A summary of key parameters from the reviewed literature is presented below:

- Dhruvakumar et al. [1] Combines DHT11 sensors with Arduino–NodeMCU and image analysis to automate irrigation and disinfection, improving cocoon yield and disease detection precision.
- Manjunatha et al. [2] Uses Raspberry Pi and camera modules with ML-based analysis for continuous, high-resolution image monitoring of silkworm diseases.
- Puneeth & Veena [3] Integrates IoT sensors, wearables, and computer vision to enable predictive analytics, real-time monitoring, and early disease intervention.
- Guo-Ping & Xi-Jie [4] Reviews silkworm pathology and diagnostics extensively but provides limited quantitative analysis of environmental parameters.
- Shilpa B. Kodli et al. [5] Applies CNN-based image classification for major silkworm diseases, achieving ~99% accuracy via a web application without environmental inputs.

4. CONCLUSION

The AI The integration of IoT, Artificial Intelligence, and Deep Learning represents a paradigm shift in sericulture. As evidenced by the literature and the proposed system architecture, these technologies significantly enhance the efficiency, reliability, and automation of silkworm rearing. The reviewed systems demonstrate high efficacy in disease detection and classification, with accuracies ranging from 85% to 99%. While deep learning models offer superior resilience and precision, the combination of these algorithms with low-cost hardware like the Raspberry Pi ensures that these solutions are accessible and scalable. Future developments should focus on optimizing these systems to handle

environmental variables more robustly and further reducing the computational load for edge devices. Ultimately, automating disease detection allows farmers to safeguard their crops against mass infection, ensuring a sustainable and profitable silk production industry. And combat bias through semantic analysis, AI tools will become even more integral to personalized learning paths. The successful integration of these technologies depends on an approach that prioritizes oversight and continuous evaluation. Ultimately, the goal is not to replace human educators but to augment their capabilities and empower learners with tools that are scalable and supportive.

References

1. Dhruva Kumar K, Kruthik M V, Manoj Gowda A M, Madhu Gowda M H, Padmavathi C, "Automated Smart Sericulture System Based on IOT and Image Processing Technique", *International Journal for Multidisciplinary Research (IJFMR)*, Vol. 7, Issue 2, pp. 1–9, 2025.
2. Manjunatha S., Sindhu B. S., Sirisha R., Tejas Gowda R., "Silkworm Disease Detection Using AI and ML," *International Journal of Research Publication and Reviews*, Vol. 5, Issue 4, pp. 8849–8852, 2024.
3. Puneeth, M., Veena, S., "IoT and Image Processing in Sericulture," *International Journal of Research Publication and Reviews*, 5(5), pp. 4439–4445, 2024.
4. Guo-Ping, K., Xi-Jie, G., "Overview of silkworm pathology in China," *African Journal of Biotechnology*, 10(79), pp. 18046–18056, 2011.
5. Kodli, S. B., Anita, "A novel approach for silkworm disease detection and classification by using CNN and image processing techniques," *Journal of Scientific Research and Technology*, 3(3), pp. 21–29, 2025.
6. Manju, G., Binson, V. A., "Automated disease detection in silkworms using machine learning techniques," *Advance Sustainable Science, Engineering and Technology*, 6(4), 2024.
7. Aparna Priya, J., Hemalatha, B., Amrita, S., Kusumasree, "IoT based animal health monitoring system using Raspberry Pi," *International Journal of Advanced Research in Science, Communication and Technology*, 3(7), pp. 300–307, 2023

Biographies



Preetam Patil is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). His research areas include Artificial Intelligence, Machine Learning, and VLSI.



Rudresh Patil is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). His research areas include Machine Learning, VLSI, and Embedded System.



Sharanaprakash Hosamani is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). His research areas include Deep Learning, Machine Learning and Embedded system.



Soujanya D G is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). Her research areas include VLSI, Embedded System and Machine Learning.



Hemanth T S received his B.E. in Electronics and Communication Engineering from Channabasaveshwara Institute of Technology, Gubbi, Tumakuru, M.Tech in VLSI Design and Embedded Systems from Siddaganga Institute of Technology, Tumakuru. He is currently working as Assistant Professor in the Department of Electronics and Communication Engineering at Alva's Institute of Engineering & Technology. His research interests include VLSI Design, Embedded Systems, Deep Learning, and Machine Learning.

Predictive Modeling and Survival Analysis of Primary Biliary Cirrhosis: A Comparative Study of Machine Learning Approaches

El-Sayed M. El-kenawy

*Delta Higher Institute of Engineering and Technology
Department for Communications and Electronics
Mansoura 35511, Egypt*

*Applied Science Research Center
Applied Science Private University
Amman, Jordan
skenawy@ieee.org*

Manish Kumar Singla

*Department of Biosciences,
Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences,
Chennai, India-602 105
msingla0509@gmail.com*

Amel Ali Alhussan

*Department of Computer Sciences
College of Computer
& Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
aaalhussan@pnu.edu.sa*

Doaa Sami Khafaga

*Department of Computer Sciences
College of Computer
& Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
dskhafga@pnu.edu.sa*

Amal H. Alharbi

*Department of Computer Sciences
College of Computer
& Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
ahalharbi@pnu.edu.sa*

Marwa M. Eid

*Faculty of Artificial Intelligence
Delta University for Science
& Technology Mansoura 35111, Egypt
Jadara University Research Center
Jadara University
Jordan
mmm@ieee.org*

Abstract— Primary biliary cirrhosis (PBC)—an immune-mediated chronic liver disease destroying intrahepatic bile ducts, leading to cirrhosis/failure—was analysed using Mayo Clinic trial data (424 patients, 1974-1984): 312 in randomized D-penicillamine/placebo trial, 130 observational. Machine learning models trained on demographic/biochemical features for survival/prognosis prediction identified MLP as optimal (96.07% accuracy, 94.34% sensitivity, 96.99% specificity), outperforming KNN and Decision Tree. Serum bilirubin and liver disease stage emerged as key prognostic indicators. These models support clinical decision-making for PBC identification/management.

Index Terms—Primary biliary cirrhosis, predictive modelling, machine learning, survival analysis, clinical indicators, disease progression.

I. INTRODUCTION

Primary biliary cirrhosis is a chronic autoimmune liver disease that predominantly affects middle-aged women, leading to progressive bile duct destruction, fibrosis, cirrhosis, and potentially liver failure. Traditional prognostic tools like MELD and Child–Pugh scores, though useful, rely on linear assumptions and struggle to integrate complex, multidimensional clinical data. This study uses ML algorithms (MLP, KNN, Decision Trees, Random Forests) on the 424-patient Mayo Clinic PBC cohort (1974–1984) to predict prognosis and survival, aiming to build accurate models, identify key predictors, compare methods, and assess clinical utility. By modeling nonlinear relationships and handling high-dimensional features, these ML approaches are positioned to improve risk stratification, monitoring, and precision medicine in hepatology.

II. LITERATURE REVIEW

Machine learning outperforms traditional scores (like MELD/MELD-Na) for predicting cirrhosis outcomes such as mortality, portal hypertension, readmission, HCC, fibrosis, microbiome-based severity, and encephalopathy, often reaching AUC/C-index around 0.8–0.9. However, these gains highlight the parallel need for better interpretability so models can be trusted and used in real-world clinical practice.

III. DISCUSSION AND RESULTS

The study uses the 424-patient Mayo Clinic PBC cohort and a Kaggle-formatted cirrhosis dataset with rich clinical and demographic features to build survival and prognosis models. Figure 1 indicates that more advanced cirrhosis stages cluster in older patients, suggesting a relationship between age and disease severity.

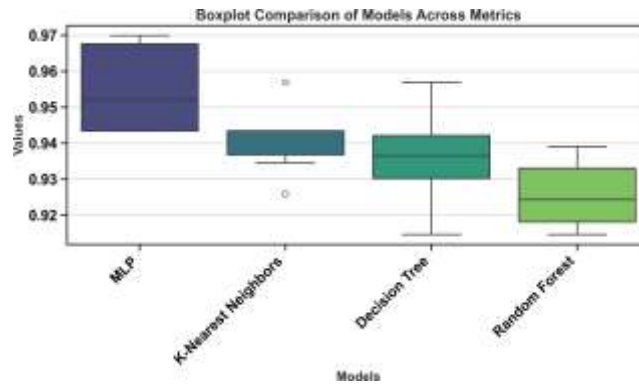


Fig. 1. Stage vs Age Plot

Table I reports the performance of several machine learning models on the cirrhosis dataset using accuracy, sensitivity, specificity, and F-score as evaluation metrics. The **MLP** model achieves the highest accuracy of 0.9607, indicating superior ability to correctly predict patient outcomes and distinguish cirrhosis status in this cohort.

Figure 2 shows that the MLP model has the lowest and most consistent errors, with a narrow spread and maximum error around 0.20, confirming it as the most reliable model for PBC outcome prediction.

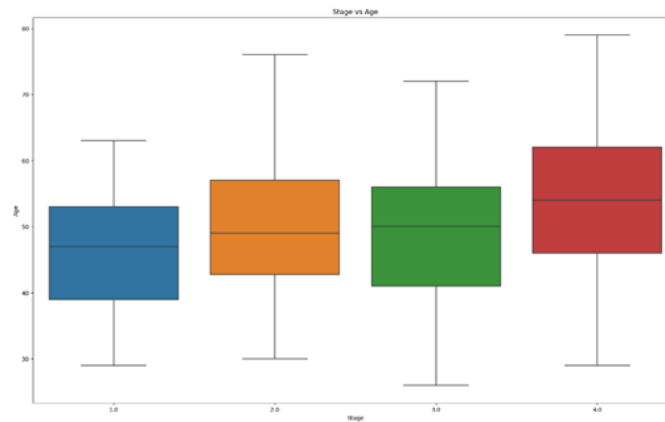


Fig. 2. Comparison of Machine Learning Model Performance in Predicting PBC Outcomes: Boxplot Analysis

Figure 3 shows that the MLP model has the highest Precision (around 0.95) and a similarly strong NPV, meaning it is best at correctly identifying both PBC and non-PBC cases.

Figure 4 KDE plot shows the top model's accuracy tightly concentrated around 0.87–0.89, indicating consistently high and reliable survival prediction performance.

TABLE I
PERFORMANCE METRICS OF MACHINE LEARNING MODELS ON THE CIRRHOSIS DATASET

Model	Accuracy	Sensitivity (TPR)	Specificity (TNR)	Precision (PPV)	Negative Predictive Value (NPV)	F1-Score
MLP	0.9607	0.9434	0.9699	0.9434	0.9699	0.9434
K-Nearest Neighbors	0.9434	0.9259	0.9569	0.9434	0.9434	0.9346
Decision Tree	0.9383	0.9146	0.9569	0.9434	0.9346	0.9288
Random Forest	0.9284	0.9146	0.9390	0.9202	0.9346	0.9174

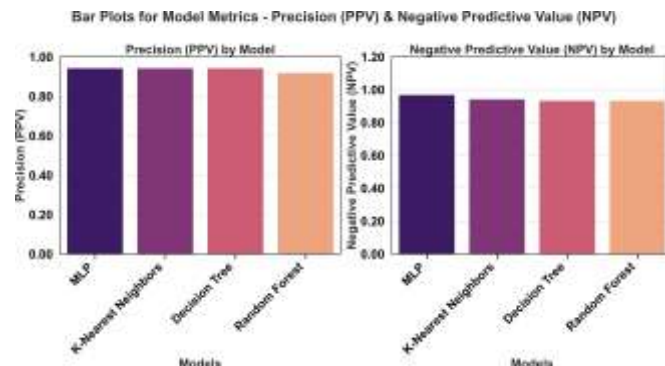


Fig. 3. Precision and Negative Predictive Value Comparison of Machine Learning Models for Primary Biliary Cirrhosis Prediction

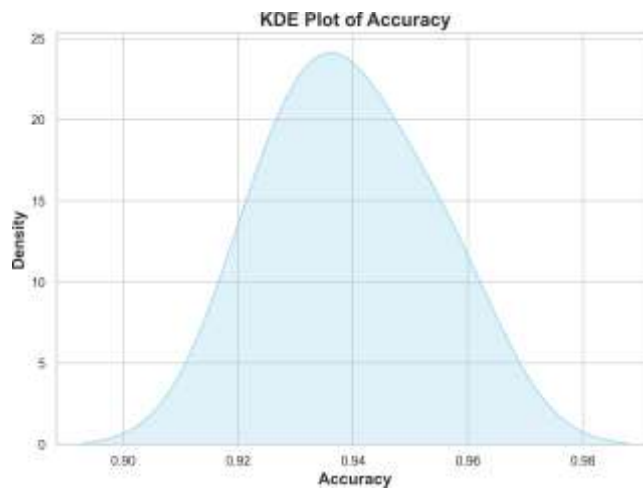


Fig. 4. KDE Plot of Model Accuracy in Predicting PBC Survival

Figure 5 shows that the best PBC model consistently achieves high accuracy, sensitivity, and specificity across runs, with CDFs rising sharply at strong performance levels, indicating stable and reliable predictive behavior.

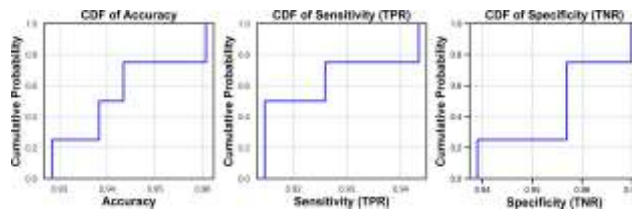


Fig. 5. Cumulative Distribution Functions of Accuracy, Sensitivity, and Specificity for PBC Prediction Model

Figure 6 shows density plots of the best model's metrics for PBC prediction. Accuracy values cluster at the high end, specificity peaks near 0.95, and F1-score centers around 0.9, indicating consistently strong overall performance.

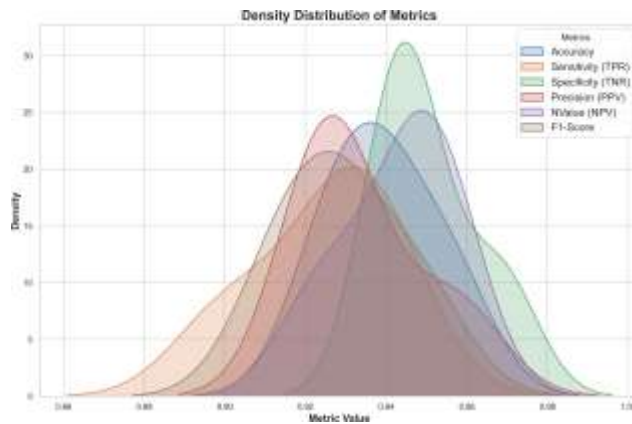


Fig. 6. Density Distributions of Performance Metrics for PBC Prediction

Figure 7 shows that, among all algorithms, the MLP's performance values are most densely concentrated at the high end of the distribution, indicating it most frequently achieves superior predictive accuracy for PBC.

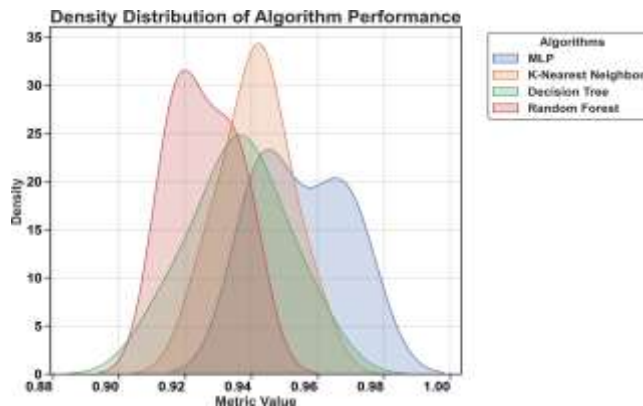


Fig. 7. Density Distribution of Algorithm Performance in Predicting Primary Biliary Cirrhosis

Figure 8 compares Accuracy and Sensitivity (TPR) for MLP, KNN, Decision Tree, and Random Forest in predicting PBC outcomes. MLP attains the highest accuracy (≈ 0.8), indicating strong correct classification of PBC-related cases.

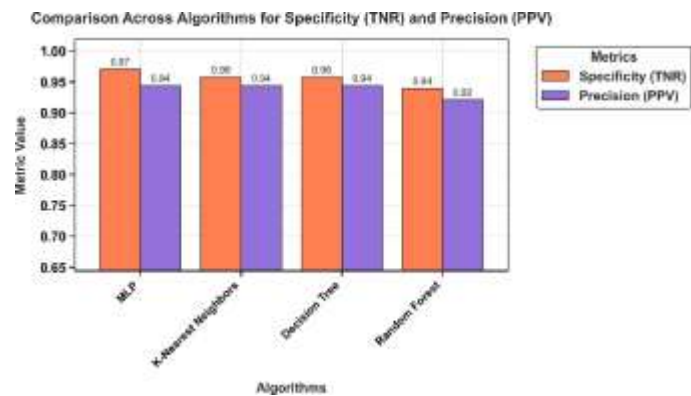


Fig. 8. Accuracy and Sensitivity of Machine Learning Algorithms for PBC Prediction

Figure 9 compares Specificity (TNR) and Precision (PPV) for MLP, KNN, Decision Tree, and Random Forest in PBC prediction. Figure 10 uses a radar plot over Sensitivity, Specificity, Accuracy, F1-score, NPV, and Precision, where the Multilayer Perceptron shows the strongest overall performance profile among the models.

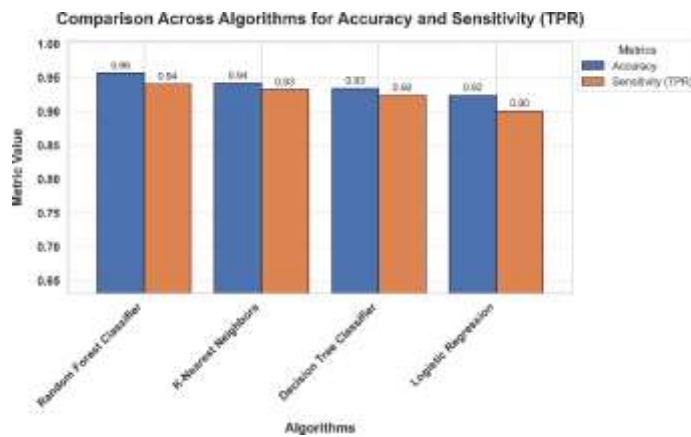


Fig. 9. Specificity and Precision Comparison of Machine Learning Algorithms for PBC Prediction

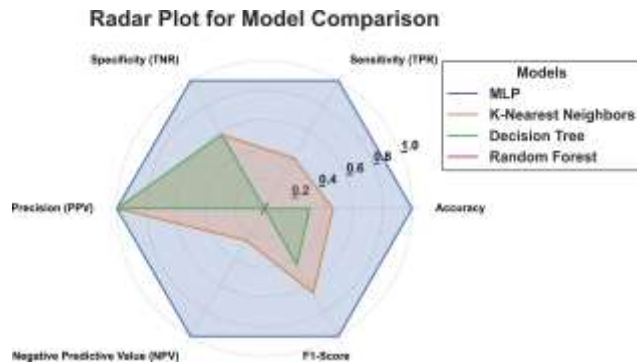


Fig. 10. Radar Plot Comparing Machine Learning Models for PBC Prediction

Figure 11 plots Accuracy (objective function) for MLP, KNN, Decision Tree, and Random Forest on the PBC dataset. MLP attains the highest and most consistent performance, with maximum accuracy of about 0.978.

IV. CONCLUSION

Machine learning on the Mayo PBC cohort shows that key clinical and biochemical variables can be used to predict prognosis with high accuracy, with an MLP model achieving around 96% accuracy, 94% sensitivity, and 97% specificity. These models both complement existing prognostic scores and highlight the need for external validation and integration with multimodal and EHR data to enable scalable, explainable precision medicine tools in chronic liver disease.

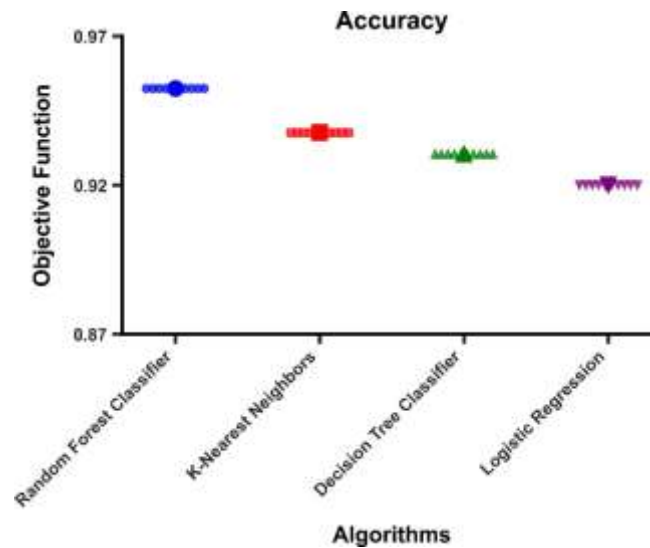


Fig. 11. Accuracy Comparison of Machine Learning Algorithms for Primary Biliary Cirrhosis Prediction

REFERENCES

- [1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [2] El-Sayed M Elkenawy, Amel Ali Alhussan, Doaa Sami Khafaga, Zahraa Tarek, and Ahmed M Elshewey. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [4] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Ab-delhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [5] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, pages 1–1, 2025.
- [6] F. Kanwal, T. J. Taylor, J. R. Kramer, Y. Cao, D. Smith, et al. Develop-ment, validation, and evaluation of a simple machine learning model to predict cirrhosis mortality. *JAMA Network Open*, 3(11):e2023780, 2020.
- [7] J. Reinis, O. Petrenko, B. Simbrunner, B. S. Hofer, F. Schepis, et al. Assessment of portal hypertension severity using machine learning models in patients with compensated cirrhosis. *Journal of Hepatology*, 78(2):390–400, 2023.
- [8] C. Hu, V. Anjur, K. Saboo, K. R. Reddy, J. O’Leary, et al. Low predictability of readmissions and death using machine learning in cirrhosis. *Official journal of the American College of Gastroenterology* — *ACG*, 116(2):336, 2021.
- [9] E. Audureau, F. Carrat, R. Layese, C. Cagnot, T. Asselah, et al. Personalized surveillance for hepatocellular carcinoma in cirrhosis – using machine learning adapted to hcv status. *Journal of Hepatology*, 73(6):1434–1445, 2020.
- [10] D. Chang, E. Truong, E. A. Mena, F. Pacheco, M. Wong, et al. Machine learning models are superior to noninvasive tests in identifying clinically significant stages of nafld and nafld-related cirrhosis. *Hepatology*, 77(2):546, 2023.
- [11] M. Nouredin, Z. Goodman, D. Tai, E. L. K. Chng, Y. Ren, et al. Machine learning liver histology scores correlate with portal hypertension assessments in nonalcoholic steatohepatitis cirrhosis. *Alimentary Pharmacology & Therapeutics*, 57(4):409–417, 2023.
- [12] K. Saboo, N. V. Petrakov, A. Shamsaddini, A. Fagan, E. A. Gavis, et al. Stool microbiota are superior to saliva in distinguishing cirrhosis and hepatic encephalopathy using machine learning. *Journal of Hepatology*, 76(3):600–607, 2022.
- [13] H. Yang, X. Li, H. Cao, Y. Cui, Y. Luo, et al. Using machine learning methods to predict hepatic encephalopathy in cirrhotic patients with unbalanced data. *Computer Methods and Programs in Biomedicine*, 211:106420, 2021.
- [14] A. Guo, N. R. Mazumder, D. P. Ladner, and R. E. Foraker. Predicting mortality among patients with liver cirrhosis in electronic health records with machine learning. *PLOS ONE*, 16(8):e0256428, 2021.
- [15] W. M. Shaban. Early diagnosis of liver disease using improved binary butterfly optimization and machine learning algorithms. *Multimedia Tools and Applications*, 83(10):30867–30895, 2024.

Metaheuristic-Tuned Convolutional Networks for Plant Disease Recognition: Waterwheel Plant Algorithm as a Standout Optimizer

Amel Ali Alhussan

*Department of Computer Sciences
College of Computer
& Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
aaalhussan@pnu.edu.sa*

Manish Kumar Singla

*Department of Biosciences,
Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences,
Chennai, India-602 105
msingla0509@gmail.com*

Doaa Sami Khafaga

*Department of Computer Sciences
College of Computer
& Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
dskhafaga@pnu.edu.sa*

Amal H. Alharbi

*Department of Computer Sciences
College of Computer
& Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
ahalharbi@pnu.edu.sa*

El-Sayed M. El-kenawy

*Delta Higher Institute of Engineering and Technology
Department for Communications and Electronics
Mansoura 35511, Egypt*

Applied Science Research Center

*Applied Science Private University
Amman, Jordan
skenawy@ieee.org*

Marwa M. Eid

*Faculty of Artificial Intelligence
Delta University for Science
& Technology Mansoura 35111, Egypt
Jadara University Research Center
Jadara University
Jordan
mmm@ieee.org*

Abstract— Plant diseases cause ~1/3 of annual agricultural losses, threatening food security; timely detection ensures sustainability. This study optimizes CNNs with metaheuristics (WWPA, GWO, PSO, WOA, GA) for classifying 87,000 RGB images across 38 healthy/diseased crop categories. WWPA-CNN achieved superior performance: 97.89% accuracy, 97.53% sensitivity, 98.20% specificity, 97.79% F-score—outperforming GWO-CNN (95.33%), PSO-CNN (94.30%), WOA-CNN (93.45%), and GA-CNN (92.82%). Histograms, KDE, and CDF plots confirmed WWPA-CNN's stability. Metaheuristic-optimized CNNs enable reliable, scalable plant disease detection for real-world agriculture.

Index Terms—Plant Disease Detection, Convolutional Neural Networks (CNNs), Metaheuristic Optimization, Waterwheel Plant Algorithm (WWPA), Precision Agriculture

I. INTRODUCTION

Agriculture has long been recognized as the foundation of human civilization and remains central to global food security, rural livelihoods, and economic growth. With the global population projected to approach 10 billion by 2050, the Food and Agriculture Organization (FAO) estimates that food production must increase by at least 70% to sustain future demand. However, plant diseases represent one of the most significant challenges to achieving this target, accounting for up to one-third of annual agricultural production losses. These losses not only threaten food supply but also contribute to economic in-stability. Plant diseases threaten crop yields, exacerbating farmer vulnerability and food prices; effective detection is essential for sustainability. Deep Learning (CNNs) automatically discovers subtle disease features invisible to the naked eye, enabled by datasets like PlantVillage/PlantDoc and enhanced by metaheuristics (WWPA, GWO, PSO, WOA, GA) for field generalization. However, and impose high computational barriers for smallholder farmers. This systematic review analyzes CNN-metaheuristic architectures for plant disease prediction, highlighting contributions/limitations and practical research directions to bridge lab success and field deployment:

- 1) **Extensive testing of hybrid CNN models:** We test CNN-based architectures using metaheuristic algorithms, including WHPA-CNN, GWO-CNN, PSO-CNN, WOA-CNN, and GA-CNN. These models represent the new generation of models, aiming to enhance plant disease classification performance by providing improved parameter tuning and optimization.
- 2) **Quantitative insights into performance:** Experimental evidence demonstrates that **WHPA-CNN consistently outperforms competing models**, achieving an accuracy of 97.89%, sensitivity of 97.53%, and specificity of 98.20%. By comparison, GWO-CNN and PSO-CNN recorded accuracies of 95.33% and 94.30%, respectively, while WOA-CNN (93.45%) and GA-CNN (92.82%) delivered competitive but slightly lower outcomes. These results confirm the superior generalization capacity of WHPA-CNN across multiple metrics.
- 3) **Robustness through predictive metrics:** Beyond accuracy, this study evaluates predictive reliability using sensitivity (True Positive Rate), specificity (True Negative Rate), positive predictive value (PPV), negative predictive value (NPV), and F-score. The consistently high values observed for WHPA-CNN across these metrics underscore its robustness in minimizing both false positives and false negatives.

In summary, this work contributes to the broader body of research by bridging theoretical advances in DL-based plant disease detection with practical considerations for agricultural deployment. By critically examining recent contributions and providing comparative evidence, this review offers both re-researchers and practitioners a clear perspective on the current state of the field and the pathways toward scalable, field-ready solutions.

II. LITERATURE REVIEW

FAO estimates global food supply must increase 70% by 2050 to meet population needs, yet plant diseases cause ~2/3 of food losses, driving automated detection strategies. **Datasets:** PlantVillage (lab, poor field performance), PlantDoc (field, annotation errors), FieldPlant (5,170 images/8,629 leaves, 27 expert-verified classes—outperforms PlantDoc). ML/DL advances include YOLOv5 (object detection), ResNet50/MobileNetv2 (classification efficiency) [1]. EfficientNet transfer learning on PlantVillage achieved 99.91-99.97% accuracy [6]; thermal CNN+XAI hit 98.55% on 4,720 images, 90.04% on paddy [7]. Xception+Adam led validation (99.81% accuracy, 0.9978 F1) [8]; crop-specific CNNs scored 98.3% (potato), 98.5% (pepper), 95% (tomato) [9]. Hybrid CNN-ViT balances ViT's accuracy gains with real-time inference speed [10].

Rice, a staple crop in India, suffers from significant yield losses due to fungal and bacterial infections. A CNN-based detection method was developed to classify three major rice diseases—leaf smut, brown spot, and bacterial leaf blight. Using Otsu's thresholding for background removal, the model trained on 8,000 leaf images achieved 99.7% accuracy, out-performing prior rice disease recognition methods [11].

III. MATERIALS AND METHODS

The dataset, downloaded from Kaggle (augmented recreation of GitHub original), contains ~87,000 RGB images of healthy/diseased crop leaves across 38 classes. Split 80:20 (training: validation), it retains original directory structure plus a separate test set of 33 unseen images for generalization testing. Figure 1 shows balanced class distribution (~1,800-2,000 images each) across categories like Apple healthy, Apple Cedar apple rust, Apple Black rot, and Apple scab—essential for unbiased deep learning training.

To train the model, a preprocessing sequence was employed on all images to make sure they were consistent and suitable as inputs to deep learning models:

- 1) **Image loading:** All images were read from the dataset directories and loaded into memory in their raw format.

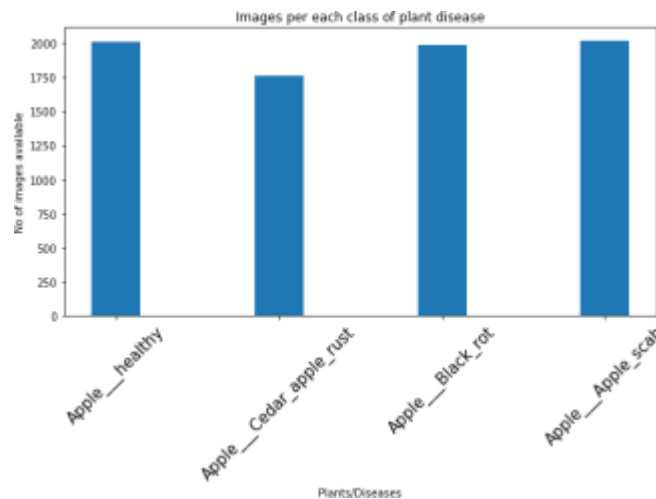


Fig. 1. Distribution of images across different classes of apple diseases in the dataset.

- 2) **Resizing:** All images were resized to a common resolution that the CNN input layers could accept, ensuring a uniform image size regardless of the original image.
- 3) **RGB conversion:** To maintain the color information, which is crucial in the disease identification process, the images were converted into a three-channel RGB format.
- 4) **Normalization:** Normalization of pixel values was performed to the range $[0, 1]$ to improve model stability during training and accelerate convergence.
- 5) **Tensor conversion:** Finally, the preprocessed images were converted into tensors to serve as inputs for the deep learning models during training and evaluation.

These preprocessing steps ensured that the dataset-maintained consistency in size, scale, and format, thereby improving the efficiency and performance of the deep learning models trained in subsequent stages.

Convolutional Neural Network (CNN) CNNs automatically extract spatial/hierarchical features from raw plant disease images. This study's baseline architecture used convolutional/pooling layers for feature extraction, ReLU activations, batch normalization for training stability, dropout regularization in fully connected layers, and softmax for 38-class probability outputs. While baseline CNNs yielded satisfactory results, metaheuristic optimization (learning rates, dropout, filter sizes) maximized accuracy, sensitivity, and specificity under adverse conditions. WWPA produced the largest gains; WWPA-CNN outperformed unoptimized models across all metrics, validating metaheuristic-enhanced CNNs for real-world plant disease detection.

IV. HYPERPARAMETER OPTIMIZATION RESULTS

Five CNNs optimized with WWPA, GWO, PSO, WOA, and GA were evaluated across Accuracy, Sensitivity (TPR), Specificity (TNR), PPV, NPV, and F-score. WWPA-CNN led with 97.89% accuracy, 97.53% sensitivity, 98.20% specificity, 98.04% PPV, 97.75% NPV, and 97.79% F-score. GWO-CNN followed (95.33% accuracy), PSO-CNN emphasized specificity/PPV (94.30% accuracy, 93.40% sensitivity), while WOA-CNN (93.45%) and GA-CNN (92.82%, lowest specificity) trailed. Metaheuristics universally improved performance: Figure 2's histograms+KDE visualize metric distributions, central tendencies, and consistency across models.

TABLE I
PERFORMANCE COMPARISON OF CNN MODELS OPTIMIZED WITH DIFFERENT METAHEURISTIC ALGORITHMS.

Model	Accuracy	TPR	TNR	PPV	NPV	F-score
WWPA-CNN	0.9789	0.9753	0.9820	0.9804	0.9775	0.9779
GWO-CNN	0.9533	0.9458	0.9599	0.9554	0.9514	0.9506
PSO-CNN	0.9430	0.9340	0.9529	0.9554	0.9301	0.9446
WOA-CNN	0.9345	0.9340	0.9350	0.9387	0.9301	0.9363
GA-CNN	0.9282	0.9340	0.9206	0.9387	0.9145	0.9363

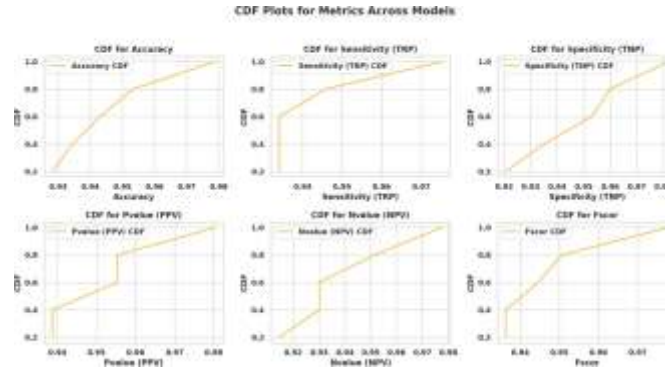


Fig. 2. Histogram and KDE plots for the evaluation metrics across optimized CNN models.

Cumulative Distribution Function (CDF) plots in Figure 3 examine optimized CNN performance distributions across Accuracy, Sensitivity (TPR), Specificity (TNR), PPV, NPV, and F-score. Unlike histograms/KDE, CDFs show the percentage of models achieving specific performance levels, enabling clear detection of significant inter-model differences and relative comparisons.

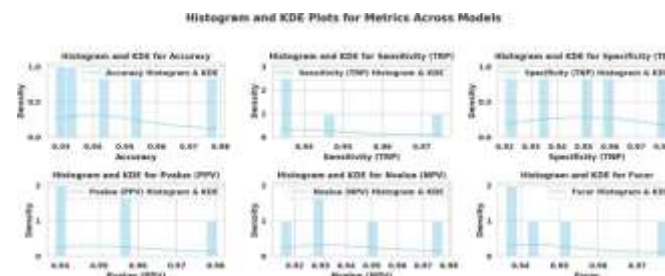


Fig. 3. CDF plots for evaluation metrics across optimized CNN models.

Figure 4's line plots with mean/standard deviation examine model consistency across Accuracy, Sensitivity (TPR), Specificity (TNR), PPV, NPV, and F-score for metaheuristic-optimized CNNs. Solid blue lines show individual model performance; red dashed lines indicate means; green dotted lines mark ± 1 SD from means—enabling direct optimizer comparison relative to overall variability..

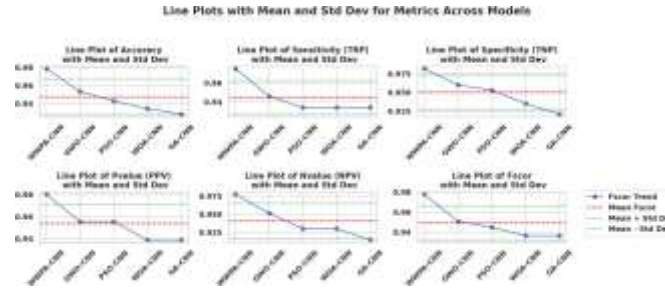


Fig. 4. Line plots with mean and standard deviation for evaluation metrics across optimized CNN models.

V. CONCLUSION AND FUTURE WORK

This study demonstrates CNN-metaheuristic integration (WWPA, GWO, PSO, WOA, GA) for plant disease detection, with WWPA-CNN achieving top performance: 97.89% accuracy, excelling across accuracy, sensitivity, specificity, PPV, NPV, and F-score. KDE/CDF visualizations confirm its robustness. Limitations include controlled dataset bias and computational complexity for resource-constrained farmers. Future work targets diverse field datasets (pathologist-verified), lightweight edge-deployable CNNs, multimodal data (thermal/hyperspectral), explainable AI, hybrid optimization, and adaptive frameworks to enable scalable precision agriculture for food security.

REFERENCES

- [1] El-Sayed M Elkenawy, Amel Ali Alhussan, Doaa Sami Khafaga, Zahraa Tarek, and Ahmed M Elshewey. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.
- [2] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [4] Emmanuel Moupojou, Appolinaire Tagne, Florent Retraint, Anicet Ta-donkemwa, Dongmo Wilfried, Hyppolite Tapamo, and Marcellin Nken-lifack. Fieldplant: A dataset of field plant images for plant disease detection and classification with deep learning. *IEEE Access*, 11:35398–35410, 2023.
- [5] Vasileios Balafas, Emmanouil Karantoumanis, Malamati Louta, and Nikolaos Ploskas. Machine learning and deep learning for plant disease classification and detection. *IEEE Access*, 11:114352–114377, 2023.
- [6] Umit Atila, Murat Ucar, Kemal Akyol, and Emine Ucar. Plant leaf disease classification using efficient net deep learning model. *Ecological Informatics*, 61:101182, 2021.
- [7] Ganbayar Batchuluun, Se Hyun Nam, and Kang Ryoung Park. Deep learning-based plant classification and crop disease classification by thermal camera. *Journal of King Saud University - Computer and Information Sciences*, 34(10, Part B):10474–10486, 2022.
- [8] Muhammad Hammad Saleem, Johan Potgieter, and Khalid Mahmood Arif. Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers. *Plants*, 9(10), 2020.
- [9] A. Lakshmanarao, M. Raja Babu, and T. Srinivasa Ravi Kiran. Plant disease prediction and classification using deep learning convnets. In *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, pages 1–6, 2021.
- [10] Yasamin Borhani, Javad Khoramdel, and Esmail Najafi. A deep learning-based approach for automated plant disease classification using vision transformer. *Scientific Reports*, 12(1):11554, 2022.
- [11] Santosh Kumar Upadhyay and Avadhesh Kumar. A novel approach for rice plant diseases classification with deep convolutional neural network. *International Journal of Information Technology*, 14(1):185–199, 2022.
- [12] K. Thaiyalnayaki and Christeena Joseph. Classification of plant disease using svm and deep learning. *Materials Today: Proceedings*, 47:468–470, 2021. NCRABE.
- [13] Wenqiang Yang, Ying Yuan, Donghua Zhang, Liyuan Zheng, and Fuquan Nie. An effective image classification method for plant diseases with improved channel attention mechanism aecanet based on deep learning. *Symmetry*, 16(4), 2024.

Earthquake Risk-Level Classification Using Machine Learning and Somersaulting Spider Optimizer

El-Sayed M. El-kenawy^{1,2}

¹*Delta Higher Institute of Engineering and Technology Department for Communications and Electronic, Mansoura 35511, Egypt*

²*Applied Science Research Center Applied Science Private University Amman, Jordan, skenawy@ieee.org*

Abstract.

This study presents an innovative integrated framework that combines advanced machine learning classifiers with meta-heuristic optimization algorithms to enhance earthquake prediction capabilities. A comprehensive evaluation was conducted using five state-of-the-art models, including Random Forest, XGBoost, Gradient Boosting, Extra Trees, and Light GBM, on preprocessed seismic datasets. Random Forest emerged as the superior baseline classifier, achieving exceptional performance with 92.3% accuracy, 92.31% sensitivity, and 97.44% specificity, demonstrating balanced predictive capability across critical evaluation metrics. To further optimize classification performance, three bio-inspired metaheuristic algorithms—Somersaulting Spider Optimizer (SSO), Harris Hawks Optimization (HHO), and Whale Optimization Algorithm (WOA)—were systematically integrated with the Random Forest model. A comparative analysis revealed that SSO-Random Forest achieved the highest overall performance, with an accuracy of 93.46% and a specificity of 97.82%, significantly outperforming both the baseline and other optimized configurations. Statistical validation, as determined by one-way ANOVA, confirmed the significance of these improvements.

Keywords. Earthquake risk-level classification, Machine learning, Somersaulting Spider Optimizer (SSO), Random Forest, seismic analysis, metaheuristic optimization

1. INTRODUCTION

Earthquakes constitute catastrophic natural disasters causing significant loss of life and infrastructure damage. Traditional seismic monitoring systems face substantial limitations in speed, generalization, and predictive accuracy, while early warning delivery remains constrained by earthquake complexity and limited seismic data. Integrating advanced computational techniques has become urgent for strengthening early warning systems. Machine learning (ML) and Deep learning (DL) approaches have emerged as powerful tools for improving earthquake prediction performance [1].

Unlike conventional threshold-based methods, these data-driven techniques are capable of automatically extracting relevant features, modeling complex nonlinear relationships, and making rapid predictions from raw seismic inputs [2]. Numerous studies have explored convolutional neural networks, recurrent architectures, and hybrid models for tasks such as earthquake detection, magnitude estimation, and peak ground acceleration prediction [3]. However, despite promising advances, these approaches face critical challenges related to

generalization across diverse seismic regions, sensitivity to noise, and the scarcity of labeled training data in low- seismicity zones [4]. Addressing these limitations requires frameworks that combine the strengths of machine learning with novel strategies for optimization and statistical validation to ensure reliability and adaptability in real-world scenarios [5].

2. LITERATURE REVIEW

Deep learning is increasingly applied to seismic waveform analysis for earthquake monitoring, but models often fail to generalize across diverse regions due to geographically limited training. As shown in [6], a data recombination method generates synthetic seismic data from various locations and stations, enabling neural networks with broader applicability for detection and parameter estimation.

Optical mesh networks provide another approach for early earthquake detection by using existing fiber infrastructure. According to [7], integrating real displacement data into waveplate models allows strain simulations in fiber cables, improving machine learning models for P-wave detection and showing the potential of fiber optic grids as seismic sensors. Accurate peak ground acceleration (PGA) prediction is vital for early warning. In [8], a CNN-based end-to-end model pre- dicts PGA directly from initial seismic waves, achieving better accuracy and lower error than peak displacement methods. Data scarcity remains a challenge in low-seismicity areas. Research in [9] applies an auxiliary classifier GAN to gen- erate synthetic earthquake data, with results showing realistic seismic characteristics validated against borehole sensors.

3. EXPERIMENTAL RESULTS

A. Machine Learning Performance Evaluation

Table I presents the comprehensive performance comparison of five classification models evaluated on the earthquake pre- diction dataset. Among the evaluated models, Random Forest demonstrates superior performance across all key metrics, achieving the highest accuracy of 0.9231, sensitivity of 0.9231, and specificity of 0.9744. Additionally, it recorded a positive predictive value of 0.9292 and negative predictive value of 0.9746. Figure 1 illustrates the average error distribution for several models through comprehensive boxplot analysis with swarm overlay visualization. Random Forest exhibits exceptionally strong performance characteristics, achieving a relatively low median error value that suggests consistently accurate predic- tions.

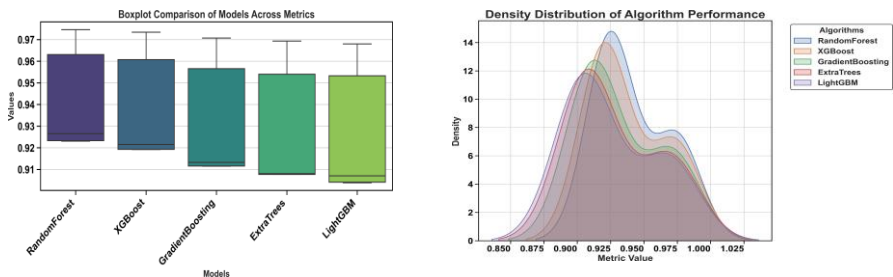


Fig. 1. Model Performance Comparison Using Boxplots with Swarm Overlay and Machine Learning Algorithm

TABLE I: CLASSIFIER PERFORMANCE COMPARISON TABLE

Models	Accuracy	Sensitivity (TPR)	Specificity (TNR)	PPV	NPV	F1 Score
Random Forest	0.9231	0.9231	0.9744	0.9292	0.9746	0.9239
XGBoost	0.9192	0.9192	0.9731	0.9237	0.9734	0.9195
Gradient Boosting	0.9115	0.9115	0.9705	0.9147	0.9707	0.9118
Extra Trees	0.9077	0.9077	0.9692	0.9084	0.9693	0.9077
Light GBM	0.9038	0.9038	0.9679	0.9092	0.9680	0.9049

4. Optimization Algorithm Enhancement Results

Table I presents a comprehensive comparative analysis of three different metaheuristic optimization algorithms when combined with the Random Forest classification model. Among the three evaluated optimization strategies, SSO-Random Forest achieves the highest overall performance with an accuracy of 0.9346, representing a significant improvement over the baseline Random Forest model. The model also demonstrates exceptional specificity of 0.9782, indicating superior capability in correctly identifying negative earthquake instances.

Figure 2 displays comprehensive comparative model evaluation through detailed boxplot analysis with swarm overlay, specifically focusing on average error distribution across the optimized models. The SSO-Random Forest model demonstrates exceptionally clustered error distribution characteristics, signifying remarkably consistent and reliable performance across multiple evaluation runs.

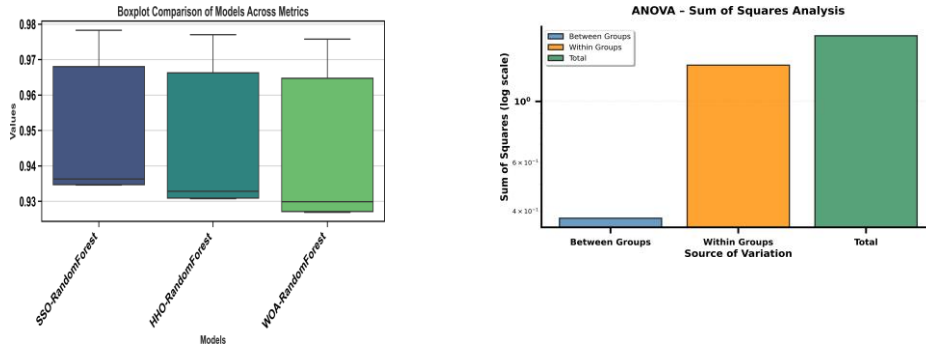


Fig. 2. Comparative Model Performance: Boxplot Analysis with Swarm Overlay for Optimized Models and ANOVA Sum of Squares Contribution Analysis

C. Statistical Significance Analysis

Table III displays comprehensive results of the one-way Analysis of Variance (ANOVA) conducted to evaluate statistical significance of performance differences among the evaluated models. The Between Groups analysis reveals an F-statistic of 68.2295, which represents a substantial statistical measure indicating significant performance variations

among the evaluated algorithms. The p-value of less than 0.0001 indicates extremely high statistical significance. Figure 2 illustrates comprehensive ANOVA results visualization on logarithmic scale, effectively highlighting the relative contributions of different variance sources to the overall statistical analysis.

5. CONCLUSION

The experimental results presented in this study demonstrate the significant potential of integrating machine learning models with advanced metaheuristic optimization algorithms for earthquake prediction. This optimized model achieved remarkable improvements in accuracy and specificity, significantly reducing false positives while preserving strong detection capability. Statistical validation through one-way ANOVA confirmed that these improvements were not incidental but rather statistically significant, reinforcing the reliability of the optimization-enhanced framework.

REFERENCES

- [1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [2] Y. K. Shejwal, A. J. Hati, J. U. Kidav, A. K. Singh, and R. Kumar, “Neural network-based classification of beamforming matrices,” *International Journal of Satellite Communications and Networking*, 2025, doi: 10.1002/sat.70002..
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [4] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [5] R. Kumar, D. Singh, M. K. Singla, E. S. M. El-Kenawy, and A. H. Alharbi, “Optimizing the placement of distributed renewable energy resources in large-scale distribution systems,” *Journal of Electrical Engineering & Technology*, pp. 1–25, 2025, doi: 10.1007/s42835-025-02279-2.
- [6] R. Kumar, Vevekanandam, and R. Dey, “IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment,” *International Journal of Vegetable Science*, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.
- [7] Hasan Awad, Fehmida Usmani, Emanuele Virgillito, Rudi Bratovich, Roberto Proietti, Stefano Straullu, Francesco Aquilino, Rosanna Pastorelli, and Vittorio Curri. Environmental surveillance through machine learning-empowered utilization of optical networks. *Sensors (Basel, Switzerland)*, 2024.
- [8] R. Kumar, A. B. Queyam, M. K. Singla, M. Louzazni, and D. K. Mishra, “Optimizing efficiency and performance in a Rankine cycle power plant analysis,” *Energy Engineering*, Mar. 2025, doi: 10.32604/ee.2025.058058.
- [9] Jae-Kwang Ahn, Byeonghak Kim, Bonhwa Ku, and Eui-Hong Hwang. Virtual scenarios of earthquake early warning to disaster management in smart cities based on auxiliary classifier generative adversarial networks. *Sensors (Basel, Switzerland)*, 2023.

<Conference abbreviation>

<Conference Series name>

<Volume number and Year> <DOI Number>

Comparative Analysis of Machine Learning Models for Rice Leaf Disease Classification: A Focus on Sensitivity, Specificity, and Accuracy

El-Sayed M. El-kenawy

Delta Higher Institute of Engineering and Technology, Department for Communications and Electronics

Mansoura 35511, Egypt

Applied Science Research Center, Applied Science Private University, Amman, Jordan
skenawy@ieee.org

Abstract—Rice leaf diseases affect rice yield, particularly in low- and middle-income countries where rice is the staple meal; these diseases include smut, brown spots, and bacterial blight. This work assesses the performance of machine learning algorithms for diagnostic and prognostic value regarding rice leaf diseases to mitigate yield loss. Using an extensive dataset from Kaggle, models such as Decision Tree (DT), K-Nearest Neighbors (K-NN), and Support Vector Machine (SVM) were assessed on metrics of accuracy, sensitivity, specificity, and predictive values. The Decision Tree model gave the highest classification accuracy of 94.31%. Specificity remained generally low across all models, suggesting difficulties in the discriminant definition of non-diseased samples; notably, this model yielded the lowest specificity. Therefore, these results support the use of a machine learning approach for disease diagnosis in rice but call for improvement in specificity to increase the reliability and practicality of the models.

Index Terms—Rice leaf disease, machine learning, classification accuracy, sensitivity, specificity, Decision Tree model

1. INTRODUCTION

In agricultural research, correct recognition and an early diagnosis of rice leaf diseases are significant priorities for preventing crop loss and ensuring food security. However, this task seems quite challenging because definitive patterns of diseases, causes, and variability in the manifestation of symptoms in rice crops are not always clear. Newer metaheuristic optimization techniques used to solve multi-dimensional and other forms of complex optimization problems are now considered effective for the challenges [1].

Among the metaheuristic algorithms gaining attention, the Waterwheel Plant Algorithm (WWPA) has demonstrated its applicability across various domains, including agricultural diagnostics [2]. This novel approach is well-fitted for disease predictor models because it performs flexible optimization over complicated structured spaces inherent to data. Similarly, the Al-Biruni Earth Radius (BER) algorithm has shown promising results in diverse optimization tasks by mimicking natural search processes, which are particularly useful in agricultural studies with complex data landscapes [3].

Similarly, enhanced ensemble learning models, especially the optimized weighted ensembles, also increase the predictability of climate unsteadiness [4]. For instance, implementing recurrent neural networks (RNN) optimized with metaheuristic methods has

been applied to predict traffic flow in urban settings, showcasing the relevance of optimization techniques in handling time-series data [6].

2. LITERATURE REVIEW

Rice is food for over half of the world's population, so detecting and predicting rice diseases on the leaves is an essential direction of agricultural studies. Diagnosing these diseases is crucial to food security.

In the study conducted by [7], deep learning and support vector machine (SVM) technologies are explored as prominent tools in pattern recognition, with their integration shown to significantly enhance recognition accuracy. This approach begins by using convolutional neural networks (CNNs) to extract features from rice leaf disease images, followed by the application of SVM to classify and predict specific diseases. The best measurement of parameters for Specific Verification Machines is done using the 10 10-fold cross-validation method, translating to a 96.8% recognition rate when $C=1$ and $g=50$.

According to the study made by [8], it is established that rice yield is affected by both biotic as well as abiotic factors such as rainfall, soil nutrient content, temperature variation and diseases originating from pests, bacteria and viruses. The convolutional neural network (CNN) algorithm, a popular tool in deep learning, has proven effective in tackling complex computer vision challenges, such as image classification and analysis. In this study, the best CNN model used was the InceptionResNetV2 CNN model, utilizing transfer learning to classify the diseases in the rice leaf images. This proposed approach can be considered an efficient tool for automated agricultural disease management since the accuracy rate of the developed model was determined to be 95.67%.

According to the perception of [9], these traditional practices are some of the causes of low agricultural yield and financial losses in the farming fraternity. In particular, 14 color spaces are considered, while 4 features are derived in each channel, which gives 172 features in total. By testing seven different classifiers, the study demonstrates that the highest classification accuracy, 94.65%, was achieved with the support vector machine (SVM) classifier. This model was trained and tested on a dataset of 619 images collected from actual agricultural fields, covering four classes: (a) Bacterial Leaf Blight (BLB), (b) Rice Blast (RB), (c) Sheath Blight (SB), and (d) Healthy Leaves (HL).

Again, the study [10] acknowledges the importance of early detection and intervention in the health of crops and food security. This research uses a Faster Region-Based Convolutional Neural Network (Faster R-CNN) for real-time detection, overcoming challenges like complex backgrounds and symptom variability. Fine-tuned on multiple data sets, the classification of rice blast, brown spot and his achieved 98.09%, 98.85% and 99.17%, respectively, with 99.25 used for healthy rice leave identification. Hence, these studies demonstrate how Faster R-CNN can be applied to developing accurate field-diagnostic disease identification systems.

While convolutional neural network (CNN) methods have evolved with transfer learning and ensemble models, few studies have systematically compared these approaches for detecting and localizing rice diseases. To fill this gap, [11] has recently compared the performance of six CNN architectures, namely DenseNet121, Inceptionv3, MobileNetV2, ResNext101, ResNet152V, and SeresNext101 in distinguishing between nine common rice diseases in Bangladesh. Additionally, transfer learning was applied to selected models, and an ensemble model (DEX, combining DenseNet121, EfficientNetB7, and Xception) was

tested. The findings indicate that the ensemble model yielded the optimum accuracy of 98%, followed by transfer learning, improving the SeresNext101 from 81% to 98%.

In the study by [12], the authors integrated artificial intelligence and machine learning with agriculture and achieved very high efficiency in disease detection by analyzing images of leaves, fields, and seeds.

In the findings by [13], the authors emphasize the importance of early infection detection of rice blast. However, there is a lack of research on the spectral detection of rice leaf blast infection, especially at the initial or latent phase. Using 2 to 4 spectral features provided the highest classification accuracies of more than 65% for the asymptomatic levels, 80% for early infection levels, and up to 95% in the presence of mild symptoms. The sequential floating forward selection combined with support vector machine (SVM) classification showed up to 10% higher accuracy than traditional methods.

Following research by [14], an automated diagnostic approach was developed and implemented as a smartphone app, using deep learning trained on a dataset of 33,026 images covering six rice diseases: leaf blast, false smut, neck blast, sheath blight, bacterial stripe disease and brown spot diseases.

Many diseases attack rice and wheat, especially leaf diseases, which threaten food production and may result in famine in certain regions. To meet this need, [15] focuses on three types of rice leaf diseases and two wheat leaf diseases using 40 augmented image samples for each type. This study improves the Visual Geometry Group Network-16 (VGG16) model by incorporating multi-task learning and using a pre-trained ImageNET model for transfer learning with alternating learning. The improved model attained 97.22% accuracy for rice leaf diseases, while for wheat leaf diseases, it was 98.75%.

3. DISCUSSION AND RESULTS

In this work, the performance of the machine learning models in identifying rice leaf diseases is investigated. We want to identify which strategies work best when it comes to distinguishing between various diseases, particularly leaf smut, brown spot, and bacterial leaf blight by breaking down the various disease features and studying the model outcomes.

3.1. Dataset

The dataset for this study, sourced from Kaggle (<https://www.kaggle.com/datasets/vbookshelf/rice-leaf-diseases>), contains images of rice leaves affected by various diseases. Rice is the most critical food crop for millions in low- and middle-income countries, where it serves as a dietary staple, especially among the impoverished. In Asia, which is a region where rice is almost a staple, it accounts for 40 percent of the food budget of people who live on \$1.25 or less per day, therefore implies that rice production is an economic and food security issue.

3.2. Data Analysis

Figure 1 shows the symptomatology of the leaf smut disease in rice whereby the affected rice plant leaves have black raised spots or smuts on the surface of the leaves. This fungal disease affects the ability of rice plants to photosynthesize and if not controlled its yield will be affected. The contrasting patterns of blackish lesions at early stages of foliar smut are good starting points to differentiate the disease from others, and assist machine learning models in making high-quality distinctions.



Fig. 1. Leaf Smut Disease

Figure 2 showed another type of disease that affected rice crops which is known as brown spot disease. This disease is characterized by brown and yellow spots mainly on the leaf blade with a bright yellow outer line. Brown spot disease is associated with the fungus *Cochliobolus miyabeanus* and is favored by low nutrient status to soil, and water stress conditions. The visual characteristics viewed in this figure help the machine learning models to distinguish the brown spot cases based on its color and shape of the spots.



Fig. 2. Brown Spot Disease

The bacterial leaf blight disease is illustrated in Figure 3, characterized by long narrow water-soaked yellow or chlorotic lesions that run along the length of the rice leaf and rapidly enlarge. Bacterial leaf blight is caused by the bacterium *Xanthomonas oryzae* and is extremely damaging and can lead to huge yield reductions if not addressed.



Fig. 3. Bacterial Leaf Blight Disease

3.3. Machine Learning Results

Table I describes each machine learning model utilized in rice disease classification and shows the comparative index of accuracy, True Positive Rate (TPR), True Negative Rate (TNR), P-value (Positive Predictive Value), N-value (Negative Predictive Value) and F-score. In the models examined, the Decision Tree (DT) model showed the highest accuracy of 94.31%, a sensitivity as high as 98.48, which is powerful evidence to support that this model has a high ability in sample diseased diagnosis. Nevertheless, overall specificity remains low for all the models; this is where the SVM model shows only 11.11%, signifying a major challenge of differentiating between the non-diseased samples.

TABLE I
ML MODEL RESULTS FOR RICE DISEASE CLASSIFICATION

Model	Acc	TPR	TNR	PPV	NPV	F1
DT	0.943	0.985	0.300	0.956	0.563	0.970
K-NN	0.917	0.960	0.300	0.951	0.346	0.956
SVM	0.893	0.948	0.111	0.938	0.130	0.943

This figure simplifies the comparison of the average classification accuracy of each model, and as a choice criterion for practical application, this factor must be considered.

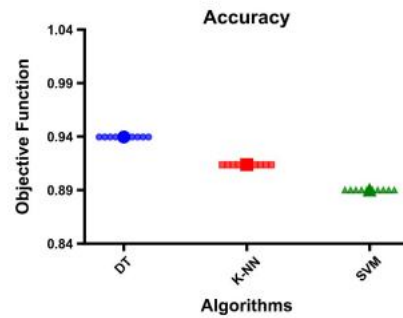


Fig. 4. Accuracy Comparison Between Models

The analysis proved that using the machine learning models is effective in the classification of rice leaf diseases for automated detection and prevention. All the developed models have distinct advantages in terms of accuracy and sensitivity; nevertheless, some of them revealed drawbacks such as low specificity, leading to decreased model precision.

4. CONCLUSION

Overall, the paper showcases how high-performing and promising a ML approach is for the classification of rice diseases, which include; leaf smut, brown spot and bacterial blight and its likelihood in supporting the automated detection of diseases in agricultural crops. The Decision Tree algorithm stood out in testing as the most accurate and sensitive model of all those presented, in terms of diagnosing diseased samples. Nevertheless, the overall performance of the models showed that there was a lack of distinctiveness which pointed to difficulties of accurately classifying non-diseased samples. This is well illustrated in the Support Vector Machine discipline where the level of accuracy was relatively high but an ability to identify between the healthy leaves and the diseased was very poor. Due to the economic and food security consequences of diseases affecting rice, and especially the developing countries relying on rice production, there is a crucial need to accurately and rapidly diagnose the diseases.

5. REFERENCES

- [1] A. A. Abdelhamid, S. K. Towfek, N. Khodadadi, A. A. Alhussan, and D. S. Khafaga. Waterwheel plant algorithm: a novel metaheuristic optimization method. *Processes*, vol. 2023, 2023.

- [2] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [3] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, pages 1–1, 2025.
- [4] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [5] El-Sayed M Elkenawy, Amel Ali Alhussan, Doaa Sami Khafaga, Zahraa Tarek, and Ahmed M Elshewey. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.
- [6] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [7] F. Jiang, Y. Lu, Y. Chen, D. Cai, and G. Li. Image recognition of four rice leaf diseases based on deep learning and support vector machine. *Computers and Electronics in Agriculture*, 179:105824, 2020.
- [8] K. N., L. V. Narasimha Prasad, C. S. Pavan Kumar, B. Subedi, and H. B. Abraha. Rice leaf diseases prediction using deep neural networks with transfer learning. *Environmental Research*, 198:111275, 2021.
- [9] V. K. Shrivastava and M. K. Pradhan. Rice plant disease classification using color features: a machine learning paradigm. *Journal of Plant Pathology*, 103(1):17–26, 2021.
- [10] B. S. Bari, M. N. Islam, M. Rashid, M. J. Hasan, and M. A. M. Razman. A real-time approach of diagnosing rice leaf disease using deep learning-based faster r-cnn framework. *PeerJ Computer Science*, 7:e432, 2021.
- [11] M. T. Ahad, Y. Li, B. Song, and T. Bhuiyan. Comparison of cnn-based deep learning architectures for rice diseases classification. *Artificial Intelligence in Agriculture*, 9:22–35, 2023.
- [12] S. Aggarwal, M. Suchithra, N. Chandramouli, M. Sarada, and A. Verma. Rice disease detection using artificial intelligence and machine learning techniques to improvise agro-business. *Scientific Programming*, 2022(1):1757888, 2022.
- [13] L. Tian, B. Xue, Z. Wang, D. Li, and X. Yao. Spectroscopic detection of rice leaf blast infection from asymptomatic to mild stages with integrated machine learning and feature selection. *Remote Sensing of Environment*, 257:112350, 2021.
- [14] R. Deng, M. Tao, H. Xing, X. Yang, and C. Liu. Automatic diagnosis of rice diseases using deep learning. *Frontiers in Plant Science*, 12, 2021.
- [15] Z. Jiang, Z. Dong, W. Jiang, and Y. Yang. Recognition of rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning. *Computers and Electronics in Agriculture*, 186:106184, 2021.

Skin Lesions Classification: A Comparative Analysis of SGD and SVC for Optimal Accuracy

El-Sayed M. El-kenawy, Marwa M. Eid, Manish Kumar Singla, Doaa Sami Khafaga,
Amel Ali Alhussan, and Amal H. Alharbi

Delta Higher Institute of Engineering and Technology, Mansoura 35511, Egypt; Applied Science Research Center, Applied Science Private University, Amman, Jordan; Faculty of Artificial Intelligence, Delta University for Science & Technology, Mansoura 35111, Egypt; Jadara University Research Center, Jadara University, Jordan; Department of Biosciences, Saveetha School of Engineering, SIMATS, Chennai, India; Department of Computer Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia

Abstract. In the ever-evolving area of Machine Learning, choosing the proper classifier to get the maximum accuracy is one of the most critical yet very competitive prerequisites, which are always more sensitive when even slight variations in accuracy are possible to make a huge difference. This study addresses this problem by conducting a comparative analysis of two widely used classifiers: the Stochastic Gradient Descent Classifier (SGDC) and the Support Vector Classifier (SVC). To maximize accuracy, we tested each algorithm on a classification task, repeating the experiment several times. The accuracy yielded was significantly higher for SGDC in the aggregate, yielding an absolute accuracy of 97.5%, while SVC yields an absolute accuracy of only 82.3%. The performance difference of 15.2% compared with SGDC illustrates the appropriateness of the method for highly accurate applications. This study delivers practical implications regarding specific performance G-means, which recognize the most effective classifier in a particular project and stimulate further investigation to enhance the comparison scope concerning various algorithms and data sets.

Keywords. Machine Learning, Classification, Accuracy, SGDC, SVC, Classifier Comparison, Objective Function.

1. INTRODUCTION

The use of machine learning in dermatology has accelerated in parallel with the global rise in skin cancer incidence and the persistent shortage of specialist dermatologists, creating a clear demand for scalable, accurate, and automated diagnostic support systems. In contemporary clinical workflows, these systems typically analyze dermoscopic or clinical photographs to assist clinicians in differentiating benign from malignant lesions, thereby improving triage, prioritizing biopsy decisions, and potentially shortening time to treatment. Beyond simple pattern recognition, modern pipelines incorporate standardized image acquisition, quality control, and metadata handling to mitigate variability introduced by devices, lighting, and skin phototypes.

Concurrently, evidence indicates that a spectrum of AI approaches—from classical machine learning to deep convolutional architectures—can materially improve lesion classification performance when trained on sufficiently large, well-annotated datasets. CNN-based models leverage hierarchical feature learning to capture clinically salient

patterns (e.g., asymmetry, border irregularity, color variegation, dermoscopic structures) and have approached dermatologist-level accuracy in specific benchmarks. Nevertheless, persistent challenges remain, including class imbalance, domain shift across institutions and imaging devices, and the need for interpretable outputs that align with clinical reasoning. Current research therefore emphasizes stronger dataset curation, stratified evaluation, and the fusion of segmentation, feature extraction, and classification modules to support reliable deployment in heterogeneous real-world settings.

Machine learning has also seen numerous studies evaluating whether traditional classifiers or deep networks perform better in identifying skin lesions. Researchers have investigated standard classifiers, such as the Stochastic Gradient Descent Classifier (SGDC) and Support Vector Classifier (SVC), to assess their diagnostic reliability. Findings demonstrate that classifier choice significantly affects diagnostic performance in terms of accuracy, sensitivity, and specificity. Furthermore, the automatic detection of complex structures in medical images has been made possible through convolutional neural networks (CNNs).

2. LITERATURE REVIEW

Artificial intelligence has found potential use in dermatology, especially in the detection of skin cancer. Advances demonstrated in recent research show that machine learning (ML) and deep learning (DL) models can provide improved diagnostic predictions by studying datasets and can help in early disease detection and treatment. Specific research areas aim to improve algorithms for diagnosing and grading skin lesions to address the shortage of dermatologists and the rising prevalence of skin disorders.

As reported in recent studies [1], machine learning (ML) employs statistical models and algorithms that learn incrementally from data to predict future outcomes. Deep convolutional neural networks (CNNs) are now trained to assess skin images for cancer. Although specialists are efficient in diagnosing skin cancer, the growing demand for diagnostic accuracy has increased the need for automated systems that can assist dermatologists, improve efficiency, and reduce patient costs.

Skin diseases account for nearly 1.79% of the global burden of disease. According to recent findings [2], a Multi-Class Multi-Level model based on a divide-and-conquer strategy was proposed to classify multiple skin diseases. The model, evaluated on 3672 images from multiple sources, achieved a diagnostic accuracy of 96.47%, outperforming prior Multi-Class Single-Level models. A novel approach for skin cancer identification using hybrid features [3] employs a k-nearest neighbor (kNN) algorithm with 5-fold cross-validation. The surface fractal dimension is calculated using a 2D generalization of Higuchi's method. Tested on 7-Point, Med-Node, and PH2 datasets, the model achieved accuracies of 95.42%, 94.71%, and 94.88%, respectively.

As highlighted in [4], an approach for automatic lesion detection integrates features from photometric and spatial domains using the AlexNet architecture combined with a local optimal-oriented pattern, achieving 94.7% accuracy on the PADUFES-20 and MED-NODE datasets. A decision-making system [5] integrates multiple classifiers including neural networks using weighted voting, with fine-tuned CNNs (GoogleNet, ResNet-101, NasNet-Large) fused with a Support Vector Machine classifier, yielding superior diagnostic performance. In [6], a deep CNN system trained on 38,283 dermal images achieved over 90.3% sensitivity and 89.9% specificity. Recent studies [7] demonstrated that a proposed model predicted clinical management choices directly from images, improving accuracy by 13.73% and reducing over-diagnosis cases by 24.56%.

A study [8] proposed deep learning frameworks using fully convolutional residual networks (FCRN) combined with a lesion index calculation unit (LICU), achieving segmentation, feature extraction, and classification accuracies of 0.753, 0.848, and 0.912 on the ISIC 2017 dataset. For melanoma detection [9], fine-tuned transfer learning models (ResNet50, InceptionV3, Inception-ResNet) on ISIC 2018 achieved accuracies of 83.7%, 85.8%, and 84%. Hybrid metaheuristics such as BA-ABC [10] enhance image quality by optimizing contrast and brightness, achieving a Dice coefficient of 94.6% on ISIC 2016–2018 datasets. Recent advancements in optimization and ML have significantly improved classification accuracy [11], with metaheuristic methods showing promise in feature selection and dimensionality reduction [12][13][14][15].

3. DISCUSSION AND RESULTS

The following section provides a critical review of the dataset, data analysis, and the performance of machine learning models in the classification of skin lesions. Skin lesions, especially those presented in the ISIC 2019 dataset due to the rich variety of them, can be effectively utilized for training and evaluating classification models. The use of two models, SGDClassifier and SVC, involves a comparison regarding the ability of each model to correctly diagnose skin lesions.

3.1. Dataset

The dataset used in this study is the *ISIC 2019* dataset, which contains image data for skin lesion classification. The dataset consists of 25,331 dermoscopic images contributed from previous ISIC challenges (2018 and 2017). It encompasses nine diagnostic categories: melanoma, melanocytic nevus, basal cell carcinoma, actinic keratosis, benign keratosis (including solar lentigo, seborrheic keratosis, and lichen planus-like keratosis), dermatofibroma, vascular lesion, squamous cell carcinoma, and “none of the above.” This diverse collection provides a rich and heterogeneous data source useful for training and evaluating machine learning models. The dataset is publicly accessible on Kaggle at: <https://www.kaggle.com/datasets/andrewmvd/isic-2019>

3.2. Data Analysis

Figure 1 is a bar diagram which represents the distribution of various classifications of melanoma in the given dataset. This helps to understand how frequently each of the skin lesion categories occur, and to see how common each type of skin lesion is.

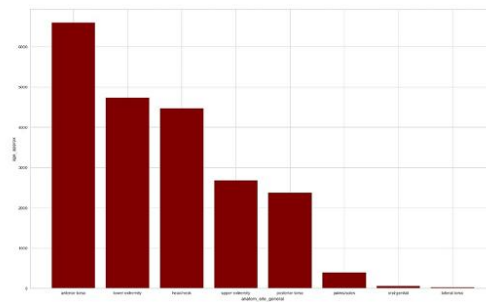


Fig. 1. Bar Plot for Melanoma Classification Dataset

3.3. Machine Learning Results

Table I displays the accuracy of two predictive models, SGDClassifier and SVC, in identifying melanomas. This table includes accuracy, sensitivity (TRP), FScore, specificity

(TNP), PPV, and NPV. As clearly shown in the table, the SGDClassifier yields reasonable performance in most aspects such as achieving an accuracy of 0.968 and specificity that is almost 1.000. On the other hand, the SVC model tends to have much lower values of all these features, with accuracy of 0.824 and specificity of 0.812, meaning that SGDClassifier is more efficient at this particular type of classification.

TABLE I. Performance Comparison Between SGD Classifier and SVC

Model	Accuracy	Sensitivity	F-Score	Specificity
SGD Classifier	0.968	0.942829	0.934569	0.998787
SVC	0.82381	0.833113	0.799205	0.812449

Figure 2 presents the confusion matrices for all models in the form of a heatmap, which enhances the visualization of the differences in both accuracy and sensitivity. The heatmap shows that all the models have a reasonably good attitude towards accuracy with a lower sensitivity towards SVC. Altogether, the results of this study stress the possibility of using classifiers within the machine learning framework.

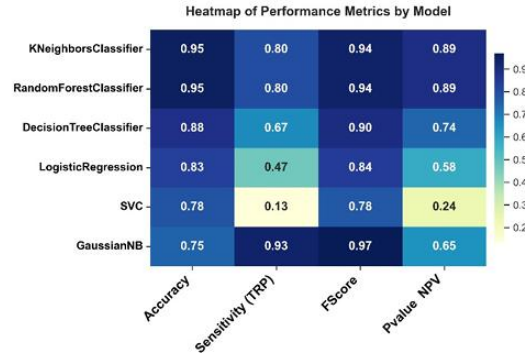


Fig. 2. Heatmap of Performance Metrics by Model

4. CONCLUSION

The comparative analysis between the Stochastic Gradient Descent Classifier (SGDC) and the Support Vector Classifier (SVC) highlights a significant performance disparity in the context of skin lesion classification. Specifically, SGDC achieved a mean accuracy of 97.5%, surpassing SVC's 82.3% by approximately 15.2%. This considerable difference in performance underscores the SGDC model's superior capability in handling high-dimensional feature spaces and large-scale datasets efficiently. The finding also illustrates that stochastic optimization and adaptive learning strategies can enhance convergence speed and model generalization in medical imaging tasks where feature distributions are highly nonlinear and heterogeneous.

Beyond the numerical superiority of SGDC, the results point to a broader insight: theoretical advantages of certain algorithms, such as SVC's robustness in smaller or linearly separable datasets, do not necessarily translate into improved real-world performance. Factors such as dataset complexity, image noise, feature scaling, and hyperparameter sensitivity often influence classifier effectiveness in practice. Hence,

model selection should not rely solely on theoretical assumptions but must be grounded in empirical evaluation across domain-specific datasets.

5. REFERENCES

- [1] K. Das, C. J. Cockerell, A. Patil, P. Pietkiewicz, M. Giuliani, et al. Machine learning and its application in skin cancer. *International Journal of Environmental Research and Public Health*, 18(24), 2021.
- [2] N. Hameed, A. M. Shabut, M. K. Ghosh, and M. A. Hossain. Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques. *Expert Systems with Applications*, 141:112961, 2020.
- [3] S. Moldovanu, F. A. Damian Michis, K. C. Biswas, A. Culea-Florescu, and L. Moraru. Skin lesion classification based on surface fractal dimensions and statistical color cluster features using an ensemble of machine learning techniques. *Cancers*, 13(21), 2021.
- [4] J. Alyami, A. Rehman, T. Sadad, M. Alruwaythi, T. Saba, et al. Automatic skin lesions detection from images through microscopic hybrid features set and machine learning classifiers. *Microscopy Research and Technique*, 85(11):3600–3607, 2022.
- [5] H. El-Khatib, D. Popescu, and L. Ichim. Deep learning–based methods for automatic diagnosis of skin lesions. *Sensors*, 20(6), 2020.
- [6] L. R. Soenksen, T. Kassis, S. T. Conover, et al. Using deep learning for dermatologist-level detection of suspicious pigmented skin lesions from wide-field images. *Science Translational Medicine*, 13(581):eabb3652, 2021.
- [7] K. Abhishek, J. Kawahara, and G. Hamarneh. Predicting the clinical management of skin lesions using deep learning. *Scientific Reports*, 11(1):7769, 2021.
- [8] Y. Li and L. Shen. Skin lesion analysis towards melanoma detection using deep learning network. *Sensors*, 18(2), 2018.
- [9] W. Gouda, N. U. Sama, G. Al-Waakid, M. Humayun, and N. Z. Jhanjhi. Detection of skin cancer based on skin lesion images using deep learning. *Healthcare*, 10(7), 2022.
- [10] S. Malik, T. Akram, M. Awais, M. A. Khan, M. Hadjouni, et al. An improved skin lesion boundary estimation for enhanced-intensity images using hybrid metaheuristics. *Diagnostics*, 13(7), 2023.
- [11] El-Sayed M El-Kenawy et al. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, 2025.
- [12] El-Sayed M Elkenawy et al. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.
- [13] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [14] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [15] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, 2025.

Al-Biruni Earth Radius Algorithm-Based Hybrid Metaheuristic for Efficient Electric Vehicle Path Planning

Marwa M. Eid, Ebrahim A. Mattar, Doaa Sami Khafaga, Amel Ali Alhussan, Amal
H. Alharbi, and El-Sayed M. El-kenawy

Faculty of Artificial Intelligence, Delta University for Science & Technology, Mansoura 35111, Egypt; Jadara University Research Center, Jadara University, Jordan; College of Engineering, University of Bahrain, Bahrain; Department of Computer Sciences, College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia; Delta Higher Institute of Engineering and Technology, Mansoura 35511, Egypt; Applied Science Research Center, Applied Science Private University, Amman, Jordan

Abstract. In this paper, a novel hybrid optimization algorithm, which integrates the Grey Wolf Optimizer (GWO) with the Al-Biruni Earth Radius (BER) method to address the problem of electric vehicle (EV) path planning, is introduced. By incorporating the simulation of wolf social hierarchies for efficient exploration alongside the spatial analysis capabilities of BER, the proposed approach effectively determines optimal, energy-efficient routes across diverse terrains. The hybrid approach systematically considers key factors such as distance, terrain characteristics, and energy consumption to guide vehicle movements, thereby minimizing energy usage, enhancing travel efficiency, and ensuring smooth traffic flow. Experimental results demonstrate the effectiveness of the proposed hybrid method compared to conventional metaheuristic approaches, achieving an average error of 0.530218.

Keywords. Electric Vehicle Path Planning, Hybrid Optimization Algorithms, Grey Wolf Optimizer (GWO), Al-Biruni Earth Radius (BER).

1. INTRODUCTION

The advancement of electric vehicles (EVs) represents a crucial milestone in sustainable transportation, contributing to reduced greenhouse gas emissions, decreased dependence on fossil fuels, improved oil displacement efficiency, and potential cost savings. With the global increase in EV adoption, driven by government policies and industrial advancements, enhancing the efficiency and reliability of EV systems has become a critical area of research. Among these systems, the problem of electric vehicle path planning (EVPP) is paramount, as it significantly influences energy consumption, travel efficiency, and user satisfaction.

In this chapter, the authors discuss the concept of creating a hybrid optimization algorithm that combines the properties of two state-of-the-art optimization algorithms: the Al-Biruni Earth Radius (BER) and the Grey Wolf Optimizer (GWO) algorithm. As a supplement to GWO, the Al-Biruni Earth Radius (BER) algorithm employs sophisticated methods of spatial analysis, drawing on mathematical and geometrical concepts. The combination of GWO and BER yields a robust optimization framework that leverages the strengths of both algorithms [1], [2], [3].

2. LITERATURE REVIEW

Metaheuristic optimization algorithms have become essential tools for addressing a wide range of complex and computationally demanding problems. Recent advancements have introduced novel algorithms inspired by natural, physical, and mathematical principles, leading to improved performance in terms of convergence speed, solution quality, and robustness [4]. Hybrid algorithms, which combine the strengths of different metaheuristic techniques, have shown further enhancements in exploration and exploitation balance [5]. Multi-objective variants have been developed to effectively approximate Pareto fronts and handle trade-offs among conflicting objectives [6]. Binary adaptations have gained attention for solving discrete and combinatorial problems, particularly in feature selection and classification tasks [7]. Adaptive and dynamic strategies within metaheuristic frameworks have been employed to fine-tune algorithm behavior during the search process [8].

EV route optimization with uncertainties in traffic conditions and demand is challenging. [9] proposes a robust optimization model where uncertainty in demand and traffic conditions are represented by convex and box uncertainty sets, respectively. [10] presents a new mathematical routing model that considers energy consumption during start/stop operations and uses an amplified Ant Colony Optimization (A2CO) algorithm, showing significant improvements in energy efficiency. [11] proposes a Hybrid Variable Neighborhood Search (HVNS) algorithm for solving the CEVRP, minimizing costs whilst optimizing fleet size and travel distance. [12] presents path planning insights for uncrewed aerial vehicles using metaheuristic algorithms relevant to EV routing in dynamic environments.

3. DATASET

The dataset presented in this study is based on an experimental study conducted by Dr. Omar Asensio and his team members [13]. The researchers gathered 3,395 electric vehicle (EV) charging events between November 2014 and October 2015. The data includes the records of 85 EV drivers who used 105 charging stations at 25 workplace sites as part of a voluntary charging program. The data contains essential features, such as the date and time of each charging session, overall energy usage, costs, and other background information. Figure 1 is a histogram of some of the most critical variables in the data set, such as the number of kWh used, the location ID, the user ID, the distance covered, the charge time in hours, and the station ID. These histograms provide a visual representation of the distribution of each feature. Knowledge of these distributions is crucial in preprocessing and normalization steps in the optimization process.

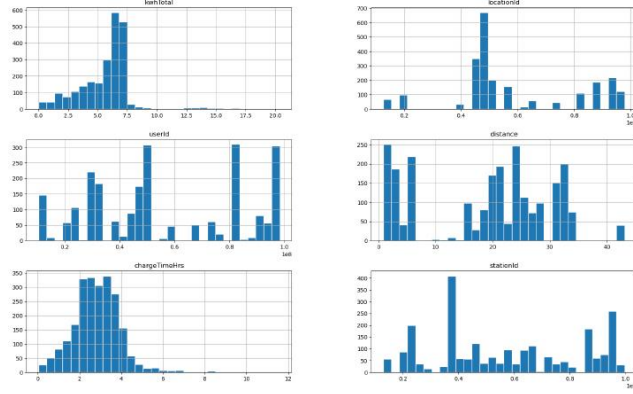


Fig. 1. Histograms of key dataset features.

4. OPTIMIZATION ALGORITHMS

4.1. Overview of the Design and Integration of the Hybrid GWO-BER Algorithm

The hybrid Grey Wolf Optimizer (GWO) and Al-Biruni Earth Radius (BER) algorithms combine the complementary strengths of these two metaheuristic optimization approaches to solve the complex challenges of Electric Vehicle Path Planning (EVPP). The hybrid approach integrates GWO and BER in a dual-phase manner. During the initial phase, GWO performs a global search space exploration, identifying promising regions while avoiding premature convergence to local optima. In the second phase, BER fine-tunes the solutions found by GWO using its precision-driven exploitation capabilities. This integration ensures the algorithm balances global search and local refinement, making it highly effective for optimizing EV energy consumption, travel time, and route efficiency.

4.2. Grey Wolf Optimizer (GWO)

The Grey Wolf Optimizer (GWO) is inspired by the social hierarchy and hunting strategies of grey wolves in nature [14]. It mimics the leadership structure of a grey wolf pack, where wolves are divided into four hierarchical roles: Alpha (α): The leader responsible for guiding the pack; Beta (β): The second-in-command, assisting and advising the alpha; Delta (δ): Subordinate wolves that dominate the lower-ranking wolves but follow the alpha and beta; Omega (ω): The lowest-ranking wolves that follow the guidance of higher-ranking wolves.

The optimization process consists of three main phases: (1) Tracking, Encircling, and Approaching the Prey — wolves update their positions iteratively based on the alpha, beta, and delta wolves, modeled using the equations:

$$\vec{D} = |\vec{C} \cdot \vec{X}_{prey} - \vec{X}_{wolf}|, \quad \vec{X}(t+1) = \vec{X}_{prey} - \vec{A} \cdot \vec{D}$$

Here, $\vec{D}$ is the distance between a wolf and the prey, $\vec{A}$ and $\vec{C}$ are coefficient vectors, and $\vec{X}$ represents the wolf's position. (2) Exploration: During early iterations, wolves explore the search space broadly by diverging from the prey, achieved by assigning $|\vec{A}| > 1$, encouraging global exploration. (3) Exploitation: In later iterations, wolves converge on the prey by reducing $|\vec{A}| < 1$, enabling local search in promising regions.

4.3. *Al-Biruni Earth Radius (BER)*

The Al-Biruni Earth Radius (BER) algorithm is inspired by geometric principles derived from Al-Biruni's calculations of the Earth's radius [15]. The algorithm dynamically partitions the population into two groups: the Exploration Group, which scans diverse regions within the solution space using trigonometric relationships to identify promising areas; and the Exploitation Group, which focuses on refining solutions within the promising regions identified by the exploration group. The BER algorithm alternates between exploration and exploitation, dynamically adjusting control parameters based on the population's fitness. The key mathematical equations governing this process include:

$$\vec{S}^*(t+1) = \vec{S}^*(t) + D \cdot (2r_2 - 1), \quad D = r_1(\vec{S}^*(t) - I)$$

where r_1 and r_2 are random variables, and D represents the adaptive search diameter.

5. RESULTS

5.1. *Evaluation Metrics and Experimental Setup*

The proposed hybrid GWO-BER algorithm was evaluated using several performance metrics: Average Error (mean deviation from the optimal solution), Average Selected Features (mean number of features selected), Fitness Values (average, best, and worst fitness scores), Accuracy (correctness of generated solutions), and Response Time (average computational time to converge). For the experimental setup, the hybrid GWO-BER algorithm was tested alongside GWO, BER, Particle Swarm Optimization (PSO), and Genetic Algorithm (GA) using benchmark datasets for feature selection and optimization. All algorithms were evaluated under identical conditions.

5.2. *Performance Comparison*

The performance of the hybrid GWO-BER algorithm was analyzed using the aforementioned evaluation metrics and compared with other metaheuristic algorithms. Table I presents the feature selection results, while Table II highlights accuracy, optimal solutions, and response times for all algorithms.

TABLE I. Performance Metrics for Feature Selection

Metric	bGWOBER	bGWO	bBER	bPSO	bGA
Average Error	0.5302	0.5547	0.6112	0.6195	0.5910
Average Selected Features	0.5130	0.7130	0.7130	0.8833	0.6554
Average Fitness	0.6234	0.6396	0.6380	0.6777	0.6510
Best Fitness	0.5252	0.5599	0.6183	0.6208	0.5543
Worst Fitness	0.6237	0.6268	0.6860	0.7005	0.6694
Std. Dev. of Fitness	0.4457	0.4504	0.4498	0.5107	0.4520

TABLE II. Performance Metrics: Accuracy, Optimal Solution, and Response Time

Algorithm	Accuracy	Optimal Solution	Avg. Response Time (s)
GWOBER	0.9335	3675.11	15.779

Algorithm	Accuracy	Optimal Solution	Avg. Response Time (s)
GWO	0.9220	3814.40	18.745
BER	0.9183	4136.40	20.113
PSO	0.9125	4532.40	21.129
GA	0.9035	4741.51	21.995

5.3. Analysis and Discussion

The results presented in Table I confirm that the hybrid GWO-BER algorithm (bGWOBER) outperformed other algorithms in feature selection by achieving the lowest average error (0.530218) and the smallest average number of selected features (0.513018). Additionally, it attained the highest fitness score (0.525218) while optimizing with fewer selected features, demonstrating its efficiency. Table II further highlights the superior performance of the hybrid GWO-BER algorithm, achieving the highest accuracy (0.933468) and the fastest response time (15.779 seconds) compared to other metaheuristic algorithms. While GWO and BER individually performed reasonably well, the hybrid approach effectively mitigated their limitations in exploration and exploitation. Overall, the hybrid GWO-BER algorithm demonstrated a well-balanced and robust performance across all evaluation metrics, making it an effective tool for solving optimization problems in EV path planning and feature selection.

6. CONCLUSION

This paper proposes a novel hybrid optimization algorithm that combines the Al-Biruni Earth Radius (BER) algorithm with the Grey Wolf Optimizer (GWO) to address problems in Electric Vehicle Path Planning (EVPP). The hybrid strategy developed a powerful and effective optimization system that leverages the global exploration potential of GWO and the local refinement capabilities of BER. Ample experimental validation has demonstrated improved performance on several key measures, including accuracy, computational efficiency, and optimization quality. The hybrid algorithm effectively balances local and global search strategies, reduces energy usage to increase EV range, and is scalable for practical applications, such as logistics, traffic congestion control, and energy-efficient transportation systems. The suggested hybrid solution aligns with the overall sustainable mobility goals, as it maximizes transportation efficiency and helps reduce carbon emissions. Future research directions include combining more metaheuristic methods to achieve better performance, building real-time implementations in dynamic environments, solving multi-objective optimization problems, and extending applications to supply chain optimization, smart city planning, and renewable energy management.

7. REFERENCES

- [1] S. Talatahari, M. Azizi, M. Tolouei, B. Talatahari, and P. Sareh. Crystal structure algorithm (crystal): A metaheuristic optimization method. *IEEE Access*, 9:71244–71261, 2021.
- [2] S. Gupta, H. Abderazek, B. S. Yildiz, A. R. Yildiz, S. Mirjalili, and S. M. Sait. Comparison of metaheuristic optimization algorithms for solving constrained mechanical design optimization problems. *Expert Systems with Applications*, 183:115351, 2021.

- [3] J.-S. Chou and D.-N. Truong. A novel metaheuristic optimizer inspired by behavior of jellyfish in ocean. *Applied Mathematics and Computation*, 389:125535, 2021.
- [4] El-Sayed M El-Kenawy et al. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, 2025.
- [5] El-Sayed M Elkenawy et al. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.
- [6] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [7] El-Sayed M Elkenawy et al. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [8] Amel Ali Alhussan et al. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, 2025.
- [9] A. Zhang, Y. Zhang, and Y. Liu. Robust optimization of electric vehicle paths in uncertain environments. *IEEE Access*, 11:124240–124251, 2023.
- [10] A. Kumar, R. Kumar, A. Aggarwal, and J. Bedi. A meta-heuristic-based energy efficient route modeling for evs integrating start/stop and recapturing energy effect. *Sustainable Cities and Society*, 91:104420, 2023.
- [11] J. Euchl and A. Yassine. A hybrid metaheuristic algorithm to solve the electric vehicle routing problem with battery recharging stations. *Energy Systems*, 14(1):243–267, 2023.
- [12] H. S. Yahia and A. S. Mohammed. Path planning optimization in unmanned aerial vehicles using meta-heuristic algorithms: a systematic review. *Environmental Monitoring and Assessment*, 195(1):30, 2022.
- [13] Omar Isaac Asensio et al. Replication Data for: A Field Experiment on Workplace Norms and Electric Vehicle Charging Etiquette, 2020.
- [14] S. Mirjalili, S. M. Mirjalili, and A. Lewis. Grey wolf optimizer. *Advances in Engineering Software*, 69:46–61, 2014.
- [15] A. Ibrahim, E.-S. M. El-kenawy et al. Al-biruni earth radius optimization based algorithm for improving prediction of hybrid solar desalination system. *Energies*, 16(3):Article 3, 2023.

Advancing Cytogenetic Diagnostics with PO-CNN: Robust Population-Based Optimization for Chromosome Classification

Ebrahim A. Mattar¹, Doaa Sami Khafaga², Amel Ali Alhussan³, El-Sayed M. El-kenawy^{4,5}, Marwa M. Eid⁶, Amal H. Alharbi⁷

¹College of Engineering, University of Bahrain, Bahrain ebmattar@uob.edu.bh

²Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia dskhafga@pnu.edu.sa

³Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia aaalhussan@pnu.edu.sa

⁴Delta Higher Institute of Engineering and Technology Department for Communications and Electronics, Mansoura 35511, Egypt

⁵Applied Science Research Center Applied Science Private University Amman, Jordan skenawy@ieee.org

⁶Faculty of Artificial Intelligence Delta University for Science & Technology Mansoura 35111, Egypt Jadara University Research Center Jadara University, Jordan mmm@ieee.org

⁷Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia ahalharbi@pnu.edu.sa

Abstract.

In this study, we propose a novel Population- based Optimized Convolutional Neural Network (PO-CNN) for automated chromosome classification. The framework integrates a population-based optimization strategy to efficiently tune CNN hyperparameters, addressing the limitations of conventional optimization methods such as Genetic Algorithm (GA), Whale Optimization Algorithm (WOA), and Particle Swarm Optimization (PSO). Experimental evaluation across multiple metrics demonstrates that PO-CNN achieves superior performance, with an accuracy of 96.33%, sensitivity of 95.97%, specificity of 96.64%, PPV of 96.48%, NPV of 96.19%, and F-score of 96.23%. Comparative analyses, including heatmaps, time-series plots, and hierarchical clustering, confirm the robustness and balanced reliability of PO-CNN against baseline models.

Keywords. Chromosome classification; Karyotype analysis; Convolutional Neural Networks (CNN); Metaheuristic optimization; Medical image analysis

1. INTRODUCTION

The leading carriers of genetic information are chromosomes, which are the basis of cytogenetic investigation. One of the foundations of medical genetics is karyotype analysis, which implies the systematic study of the number and structure of chromosomes. It is essential for the diagnosis of congenital disorders, hepatological malignancies, reproductive abnormalities, and prenatal genetic screening. The fact that patient numbers are increasing in cytogenetics labs serves to exacerbate the need for scalable solutions. Thus, the concept of automating chromosome classification using intelligent computational tools has emerged as a key field of study to minimize human influence in the process and enhance accuracy and scalability. Deep learning has become a disruptive technology in the healthcare sector,

especially in medical image analysis, in the last few years. CNNs have been incredibly successful in different areas of diagnosis, including cancer detection, radiology, pathology, and genetic imaging [1].

Their hierarchical feature representations, due to their capability to learn them, make them particularly applicable in chromosome classification, where minute differences in size, shape, and banding need to be detected with accuracy. Furthermore, advancements in metaheuristic optimization have enabled the easy optimization of CNN architectures and hyper-parameters, thereby guaranteeing greater accuracy, sensitivity, and specificity in the results. Deep learning in cytogenetics promises to enable automated systems to analyse karyotypes that are not only fast and accurate but also scalable to clinical scales [2]. Although multiple CNN-based frameworks have been proposed for chromosome classification, several gaps remain. First, many existing models rely heavily on handcrafted preprocessing and large annotated datasets [3], which are often scarce in clinical practice. Second, optimization strategies such as Genetic Algorithms (GA), Whale Optimization Algorithm (WOA), and Particle Swarm Optimization (PSO) have been integrated with CNNs to enhance classification performance [4]. However, they still struggle to achieve an optimal trade-off between accuracy, sensitivity, and specificity. As reflected in our comparative experiments, GA-CNN and WOA-CNN achieved relatively modest accuracies of 91.26% and 91.89%, respectively. PSO-CNN achieved an improved performance of 92.74%, while GWO-CNN performed even better at 93.77%. However, all of these models fell short in terms of sensitivity and F-score, rendering them less reliable for clinical decision-making. These limitations underline the necessity for more advanced optimization frameworks that can consistently improve CNN-based chromosome classification while maintaining generalizability across datasets and imaging conditions [5].

2. LITERATURE REVIEW

Chromosome recognition plays a pivotal role in diagnosing haematological malignancies and genetic diseases, yet the manual karyotyping process remains repetitive and time-consuming. To address this, several studies have proposed artificial intelligence (AI) and deep learning-based frameworks aimed at improving efficiency, accuracy, and clinical applicability. One significant advancement is KaryoNet, an end-to-end differentiable combinatorial optimization method that introduces the Masked Feature Interaction Module (MFIM) and Deep Assignment Module (DAM). This system effectively captures long-range chromosome interactions and predicts both type and polarity. It achieved state-of-the-art accuracy rates of 98.41% on R-band and 99.58% on G-band chromosomes, demonstrating strong applicability for diagnosing numerical abnormalities [6].

Another approach utilizes a deep convolutional neural network (DCNN) optimized through Hybrid Moth-Flame Optimization with Hill-Climbing (HMFOHC). To address dataset limitations, generative adversarial networks (GANs) are employed to generate synthetic training samples. This method demonstrated high performance in distinguishing five common chromosomal abnormalities (e.g., Trisomy 13, Trisomy 21, and Monosomy X), achieving an accuracy of 98.65% with reduced inference time [7]. The Model Architecture Search System (MASS) automates the selection of optimal CNN-based architectures for chromosome classification tasks by leveraging Bayesian optimization through the Tree-structured Parzen Estimator (TPE). Experiments on two datasets confirmed its superior adaptability and performance without the need for manual hyperparameter tuning [8].

3. DATASET DESCRIPTION AND PREPROCESSING

The data were produced by using a camera mounted on top of a microscope to capture images of the chromosomes. The lens and angle were adjusted under the careful supervision of professional cytogeneticists to achieve the best staining patterns. The images obtained were subsequently directed to a computer system, where additional analysis and preprocessing were performed. This would ensure that the banding information, as well as the general morphology of the chromosomes, was maintained in the digital data. This paper employed the in-situ harvest technique and the trypsin /Wright stain technique to prepare metaphase-stage chromosome spreads for G-banding.

4. OPTIMIZED DEEP LEARNING RESULTS

This performance gap of nearly five percentage points over GA-CNN reflects the robustness of the proposed optimization strategy in consistently identifying both normal and abnormal chromosomes. Sensitivity, which measures the ability of a model to correctly detect abnormal chromosomes, was also highest for PO- CNN (95.97%). This means that the model is highly effective in minimizing false negatives, an essential requirement for medical diagnosis where missing an abnormality could delay treatment or misinform clinical decisions. By comparison, the sensitivity of GA-CNN and WOA-CNN plateaued at 91.84%, showing that these models are more prone to over- look abnormalities. Similarly, PSO-CNN, although slightly better, achieved only 91.84%, while GWO-CNN performed reasonably well at 93.02% but still fell short of PO-CNN. This finding is particularly important, as sensitivity is often prioritized in screening systems where early detection of genetic disorders such as Trisomy 21 or Turner syndrome is critical.

TABLE I :Performance comparison CNN models for chromosome classification.

Models	Accuracy	Sensitivity (TRP)	Specificity (TNP)	PPV	NPV	F-score
PO-CNN	0.9633	0.9597	0.9664	0.9648	0.9619	0.9623
GWO-CNN	0.9377	0.9302	0.9443	0.9398	0.9358	0.9350
PSO-CNN	0.9274	0.9184	0.9373	0.9398	0.9145	0.9290
WOA-CNN	0.9189	0.9184	0.9194	0.9231	0.9145	0.9207
GA-CNN	0.9126	0.9184	0.9050	0.9231	0.8989	0.9207

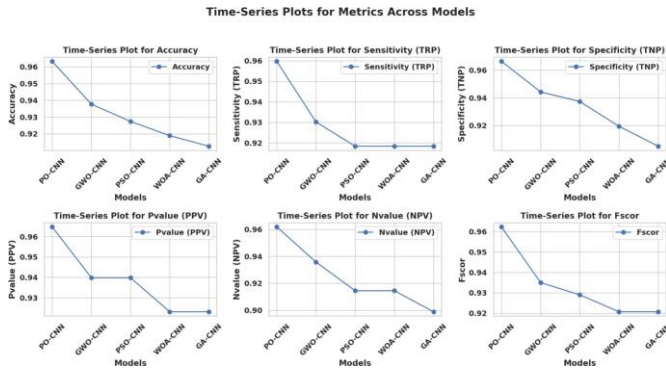


Fig. 1. Time-series plots of performance metrics across optimization-based CNN models.

CNN and PSO-CNN achieved only 93.50% and 92.90%, respectively. Both WOA-CNN and GA-CNN stagnated around 92.07%. A high F-score confirms that PO-CNN avoids the common trade-off between sensitivity and precision, instead delivering equally strong performance across both.

To further illustrate the comparative performance trends of the evaluated models, Figure 1 presents time-series style plots for each of the six-performance metrics (Accuracy,

Sensitivity, Specificity, PPV, NPV, and F-score). Unlike the tabular presentation, these plots clearly show the gradual decline in performance from PO-CNN to GA-CNN across all indicators. The sharp separation between PO-CNN and the baseline models highlights its superiority, while the relatively clustered performance of PSO-CNN, WOA-CNN, and GA-CNN indicates their limited ability to achieve consistent improvements. The dendrogram at the top of the figure highlights how metrics group together, with Accuracy, PPV, and F-score forming a tightly correlated cluster, while Sensitivity and Specificity exhibit relatively distinct behaviour.

5. CONCLUSION

This study presented a population-based optimized convolutional neural network (PO-CNN) for automated chromosome karyotype classification. Across six clinically meaningful metrics—Accuracy, Sensitivity (TRP), Specificity (TNP), PPV, NPV, and F-score—PO-CNN consistently outperformed four strong metaheuristic baselines (GWO-CNN, PSO-CNN, WOA-CNN, GA-CNN). In particular, PO-CNN achieved an Accuracy of 0.9633, Sensitivity of 0.9597, Specificity of 0.9664, PPV of 0.9648, NPV of 0.9619, and F-score of 0.9623, establishing a balanced improvement that reduces both false negatives and false positives.

REFERENCES

- [1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [2] R. K. Ratnesh, R. K. Chauhan, R. Kumar, M. Afroz, and J. Singh, “Recent advancements of plasmonic sensor technologies in healthcare and electronics,” *Journal of Molecular Structure*, 2025, Art. no. 142945, doi: 10.1016/j.molstruc.2025.142945.
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing CO₂ emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [4] M. K. Singla, S. A. Muhammed Ali, R. Kumar et al., “Revolutionizing parameter identification of PEM fuel cell using adaptive differential evolution algorithm based on deeply-informed mutation strategy and restart mechanism optimization,” *Ionics*, 2025, doi: 10.1007/s11581-025-06601-w.
- [5] R. Kumar, A. B. Queyam, M. K. Singla, M. Louzazni, and D. K. Mishra, “Optimizing efficiency and performance in a Rankine cycle power plant analysis,” *Energy Engineering*, Mar. 2025, doi: 10.32604/ee.2025.058058.
- [6] R. Kumar, A. K. Singh, M. K. Singla, A. Gupta, E.-S. M. El-Kenawy, and A. H. Alharbi, “Advancements in cancer care by exploring multimodality imaging techniques and their applications,” *Current Medical Imaging*, vol. 21, 2025, doi: 10.2174/0115734056384570250515050115.
- [7] A. K. S. Pundir, P. Singh, R. Kumar et al., “Modelling, simulation, and FPGA implementation of an augmented memory built-in self-test based design for bit-oriented memory,” *Journal of Electronic Testing*, 2025, doi: 10.1007/s10836-025-06193-3.
- [8] Nurullah Calik. The model architecture search system for chromosome image classification. *Signal, Image and Video Processing*, 18(5):4435–4445, 2024.
- [9] Mattia D’Angelo and Loris Nanni. Deep-learning-based human chromosome classification: Data augmentation and ensemble. *Information*, 14(7), 2023.

Metaheuristic Optimization in Wind power forecasting: A Comprehensive Literature Review

Amel Ali Alhussan

*Department of Computer Sciences
College of Computer
& Information Sciences*

*Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
aaalhussan@pnu.edu.sa*

Ebrahim A. Mattar

*College of Engineering,
University of Bahrain, Bahrain
ebmattar@uob.edu.bh*

Doaa Sami Khafaga

*Department of Computer Sciences
College of Computer
& Information Sciences*

*Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
dskhafga@pnu.edu.sa*

El-Sayed M. El-kenawy

*Delta Higher Institute of Engineering and Technology
Department for Communications and Electronics
Mansoura 35511, Egypt*

*Applied Science Research Center
Applied Science Private University*

*Amman, Jordan
skenawy@ieee.org
Amal H. Alharbi*

*Department of Computer Sciences
College of Computer
& Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
ahalharbi@pnu.edu.sa*

Marwa M. Eid

*Faculty of Artificial Intelligence
Delta University for Science
& Technology Mansoura 35111, Egypt
Jadara University Research Center
Jadara University
Jordan
mmm@ieee.org*

Abstract—This review synthesizes recent contributions in wind power forecasting, emphasizing the shift from deterministic models to hybrid, optimization-driven, and uncertainty-aware systems. By integrating decomposition methods, deep learning (LSTMs, Transformers), and metaheuristic optimization, modern frameworks enhance predictive accuracy and grid stability. The findings highlight a move toward more interpretable, transferable, and computationally efficient systems to support global renewable energy integration

Index Terms—wind power forecasting; metaheuristic optimization; probabilistic modeling; renewable energy; hybrid learning

INTRODUCTION

The global transition toward renewable energy has established wind power as a cornerstone of modern energy systems. Governments, industries, and researchers view wind energy not only as a means of reducing greenhouse gas emissions but also as an essential component of long-term energy security. However, despite these advantages, wind energy remains fundamentally variable, driven by meteorological and atmospheric conditions that are often complex, nonlinear, and difficult to predict.

I. The Multi-Layered Complexity of Wind Energy

Accurate forecasting of wind speed and power is crucial for balancing supply and demand, optimizing power system operations, and integrating wind energy into increasingly complex grids. Without reliable forecasts, the intermittency of wind resources introduces uncertainty into grid management, resulting in economic losses and potential threats to system stability. These challenges stem from three primary layers:

- **Physical Nonlinearity:** Wind formation involves dynamic interactions between large-scale atmospheric circulation, local terrain, and microclimatic conditions that simple statistical models cannot easily capture.
- **High-Dimensional Data:** Forecasting requires processing vast datasets from SCADA systems and remote sensors, including meteorological variables, historical output, and geographical features. It is essential to determine which variables are most relevant to maintain model accuracy.
- **Temporal Horizons:** Predictions range from real-time grid operation (minutes ahead) to market planning (days or weeks ahead). Different horizons require distinct modeling techniques, necessitating a compromise between computational efficiency and predictive accuracy.

II. The Shift Toward Hybrid Methodologies

In response to these challenges, scholars have increasingly turned to **hybrid methods** that combine the advantages of machine learning (ML), deep learning (DL), and metaheuristic optimization. This convergence is driven by the fact that ML and DL approaches are highly effective in modeling complex temporal dependencies, while metaheuristic optimization algorithms are designed to optimize hyperparameters and feature selection in high-dimensional search spaces [1, 2].

III. The Role of Feature Selection and Dimensionality Reduction

A crucial dimension of this research is the role of **feature selection**. Data gathered from meteorological stations and remote sensors can consist of hundreds of input variables. Predictive models can be compromised by redundancy and noise, leading to overfitting. Utilizing optimized metaheuristic algorithms offers a methodological means of selecting the most informative variables, increasing the reliability of the model. Dimensionality reduction not only makes models faster to train but also easier to interpret—an essential consideration for operational decision-making [3].

IV. Balancing Exploration and Exploitation

Extensive surveys of metaheuristic approaches—including evolutionary algorithms, swarm intelligence, and physics-based designs—demonstrate that no single algorithm is universally superior. Success lies in the balance between **exploration** (searching the global space to avoid premature convergence) and **exploitation** (refining promising local regions for rapid improvement) [4, 5]. The history of forecasting is a process of constant interaction between complexity and control, where the introduction of decomposition methods and adaptive optimization allows models to remain robust across various geographical settings.

V. Practical Relevance and Future Challenges

Sound short-term projections help system operators dispatch reserves more effectively, while medium-term projections support bidding strategies in electricity markets. Long-term forecasts inform infrastructure planning, such as the location of new wind farms.

Despite significant progress, problems of **generalizability** persist, especially when models are trained on datasets that do not encompass the full range of seasonal wind conditions. Furthermore, the **interpretability** of deep neural networks remains a hurdle; because they often function as "black boxes," it is challenging to explain their predictions to regulators. Integrating interpretability methods with effective models will be crucial to building the trust necessary for full operational adoption in a decarbonized future.

accelerating global transition toward renewable energy has established wind power as a cornerstone of modern energy systems. Governments, industries, and researchers alike view wind energy not only as a means of reducing greenhouse gas emissions but also as an essential component of long-term energy security. Yet, despite its advantages, wind energy remains fundamentally variable, driven by meteorological and atmospheric conditions that are often complex, nonlinear, and difficult to predict. As a result, accurate forecasting of wind speed and wind power is crucial for balancing supply and demand, optimizing power system operations, reducing curtailment, and integrating wind energy into increasingly complex and distributed grids. Without reliable forecasts, the intermittency of wind resources can introduce uncertainty into grid management, resulting in inefficiencies, economic losses, and potential threats to system stability.

I. Foundational Shift from Statistical to Hybrid Methodologies The fundamental challenge in wind and solar energy forecasting lies in the high degree of variability and intermittency inherent in meteorological data. Traditional statistical models, such as ARIMA or simple exponential smoothing, rely on linear assumptions and historical averages that frequently fail to capture the nonlinear and stochastic transitions in wind speed and solar irradiance. Research indicates that these traditional methods are often "braindead easy" to implement but struggle with high-frequency noise and sudden level shifts. Consequently, the field has transitioned toward **Machine Learning (ML)** and **Deep Learning (DL)** architectures, which excel at identifying intricate patterns within extensive datasets. However, the effectiveness of these models is not automatic; it is highly dependent on a "core layer" of metaheuristic optimization that automates the complex task of hyperparameter tuning and feature selection, ensuring the model is precisely calibrated to the specific geographical and atmospheric conditions of a given site.

II. Optimization-Enriched Training and Neural Reliability A critical advancement in this domain is the replacement of deterministic training procedures with **metaheuristic optimization algorithms**. Conventional training methods, like gradient descent, are prone to becoming trapped in local optima, particularly when dealing with the high-dimensional, non-convex loss surfaces characteristic of deep neural networks. By integrating algorithms such as the **Sparrow Search Algorithm**, **Dragonfly Algorithm**, or **Improved Adaptive Particle Swarm Optimization (MOIAPSO)**, researchers can achieve a superior balance between exploration (searching the wide parameter space) and exploitation (refining known good solutions). This results in faster convergence, reduced Root Mean Square Error (RMSE), and significantly improved generalization, which prevents the "overfitting" that often plagues complex models when they are exposed to the volatile dynamics of real-world wind power.

III. Three-Tier Decomposition and Bidirectional Modeling To further mitigate the "predictive load" on ML models, modern frameworks often employ a **3-tier hybrid approach** involving signal decomposition, bidirectional learning, and optimization.

- **Decomposition:** Techniques such as **Variational Mode Decomposition (VMD)** or **Empirical Mode Decomposition (EMD)** are used to break down the raw, noisy wind power signal into relatively stationary sub-series or "modes." This process reveals hidden trends and filters out high-frequency fluctuations that would otherwise obscure the model's learning process.
- **Bidirectional Learning:** These sub-series are then processed through **Bidirectional Long Short-Term Memory (Bi-LSTM)** networks. Unlike standard LSTMs, which only process information in forward chronological order, Bi-LSTMs learn relationships from both past and future contexts within the time series, providing a more holistic "look-ahead" and "look-back" capability.
- **Optimization:** Finally, a hybrid metaheuristic—such as the **Firefly Algorithm**—is applied to align the internal weights and architectures of the Bi-LSTM with the specific peculiarities of the decomposed data.

IV. Multi-Objective Spatio-Temporal and Probabilistic Frameworks The evolution of grid-level energy management has moved the focus from simple "point" predictions to **probabilistic forecasting** and **uncertainty quantification**. In a modern power system, wind farms are geographically distributed, meaning that errors at one site are often spatially correlated with others. A **multi-objective framework** [8] addresses this by co-optimizing conflicting goals: maximizing local accuracy, minimizing regional uncertainty, and maintaining computational tractability. By producing **Prediction Intervals (PIs)** rather than a single deterministic number, these models provide a range of expected outcomes with associated confidence levels. This allows grid operators to visualize the "worst-case" and "best-case" scenarios, which is a prerequisite for ensuring system stability during extreme weather events or sudden shifts in atmospheric pressure.

V. Operational Resilience and Real-Time Grid Stability The practical significance of these advanced methodologies is most visible in **very short-term forecasting** (minutes to an hour). Unlike day-ahead forecasts used for market scheduling, ultra-short-term predictions are essential for real-time balancing of generation and demand. Accurate modeling at this scale allows for the "active turbine control" necessary to prevent grid destabilization. By reducing the reliance on "spinning reserves"—expensive, fast-reacting fossil fuel plants that must sit idle to compensate for wind variability—optimized forecasting provides a direct path to **operational resilience**. This makes renewable energy not just an environmental choice, but a reliable, dispatchable asset that can be seamlessly integrated into modern high-penetration energy markets without compromising the safety or efficiency of the power grid.

DISCUSSION

The current body of research illustrates a profound methodological shift in renewable energy forecasting, moving from simple statistical tasks toward highly complex, integrated systems. At the core of this evolution is the indispensability of **metaheuristic optimization**. Algorithms such as Genetic Algorithms and the Sparrow Search Algorithm are no longer auxiliary tools; they are foundational components for tuning parameters and ensuring model stability. Because no single metaheuristic is universally superior, the field increasingly favors problem-specific hybridization to achieve robust results across diverse climatic conditions.

Parallel to optimization is a significant transition from deterministic point forecasting to **probabilistic and interval-based predictions**. These modern approaches acknowledge the inherent variability of wind and solar resources by quantifying uncertainty, which directly supports critical decision-making for grid operators. Driven by recurrent neural networks and spatio-temporal methods, these models strike a necessary balance between raw predictive accuracy and practical risk management. This evolution reflects a broader trend where forecasting is integrated into grid stability and market bidding strategies rather than existing as an isolated academic exercise.

Furthermore, the superiority of **hybrid deep learning architectures**—combining signal decomposition, attention mechanisms, and LSTMs—has been firmly established. While these models set new performance benchmarks, their effectiveness remains dependent on optimization-driven regularization. Despite this progress, challenges regarding limited generalizability, the "black box" nature of deep learning, and high computational costs persist. Looking forward, the availability of high-resolution SCADA data and advances in **transfer learning** offer new possibilities for creating models that are not only statistically efficient but also computationally practical and operationally resilient.

CONCLUSIONS

CONCLUSION

Modern wind energy forecasting has evolved from simple models into **hybrid, uncertainty-aware systems** that integrate decomposition methods with deep learning architectures like LSTMs and Transformers. This evolution is driven by the need to manage the inherent noise and nonlinearity of wind data, with **metaheuristic optimization** playing a critical role in ensuring model stability and generalizability across diverse forecasting horizons.

A significant paradigm shift toward **probabilistic and interval-based forecasting**—utilizing Gaussian processes and multi-objective optimization—now allows grid operators to quantify uncertainty for better risk management. These advancements extend beyond technical accuracy, directly influencing market bidding strategies and system reliability in deregulated environments. While cross-domain applications show these solutions are highly transferable, challenges such as high computational costs and model interpretability remain. Moving forward, the integration of **transfer learning** and lightweight optimization will be essential for creating the reliable, efficient energy systems required for a decarbonized future.

The fact that the forecasting of wind energy has evolved beyond mere deterministic systems that are hybrid, optimization-based, and uncertainty-aware. Hybrid models that integrate decomposition methods with deep learning models, such as LSTMs and transformers, have been shown to exhibit better performance due to the noise and nonlinearity inherent in wind data. The role of metaheuristic optimization in optimizing these models is critical, as it facilitates the generalization of such models across a wide range of datasets and forecasting horizons. Meanwhile, a distinct paradigm shift to probabilistic and interval-based forecasting indicates an increased understanding that actionable forecasts must not only be as accurate as possible but also measure uncertainty. These probabilistic methods, which rely on techniques such as Gaussian processes and multi-objective deep learning, provide plausible ranges that align with the forecasting results according to the operational and risk management specifications of contemporary power systems. Cross-domain contributions further highlight the universality of optimization-enhanced learning. Applications from microbial fuel cell prediction to hybrid solar–wind analysis and turbine optimization demonstrate that metaheuristics offer transferable solutions across various renewable energy domains. Significantly, forecasting research now extends beyond technical prediction to practical applications at the market and system level, influencing bidding strategies, hybrid system reliability, and social welfare in deregulated markets. Despite substantial progress, key challenges remain, including the generalizability of models across different regions, the interpretability of complex deep learning frameworks, and the computational costs associated with real-time deployment. Promising pathways include transfer learning for cross-site adaptation, interpretable hybrid models that integrate fuzzy logic with deep architectures, and lightweight optimization for resource-constrained settings. Collectively, the literature demonstrates that hybridization, probabilistic reasoning, and optimization are now defining features of forecasting, and that further innovation in these areas will be crucial for advancing reliable, efficient, and sustainable energy systems in the transition to a decarbonized future.

REFERENCES

- [1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [2] El-Sayed M Elkenawy, Amel Ali Alhussan, Doaa Sami Khafaga, Zahraa Tarek, and Ahmed M Elshewey. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [4] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [5] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, pages 1–1, 2025.
- [6] Md Alamgir Hossain, Evan Gray, Junwei Lu, Md Rabiul Islam, Md Shafiul Alam, Ripon Chakraborty, and Hemanshu Roy Pota. Optimized forecasting model to improve the accuracy of very short-term wind power prediction. *IEEE Transactions on Industrial Informatics*, 19(10):10145–10159, 2023. Publisher: IEEE.
- [7] Noman Mujeeb Khan, Muhammad Hamza Zafar, and Umer Amir Khan. Short Term Power Forecasting of Wind Energy System using Deep Neural Network (DNN) Trained by a Novel Meta-Heuristic Optimization Algorithm. In *2021 16th International Conference on Emerging Technologies (ICET)*, pages 1–6. IEEE, 2021.
- [8] Markos A. Kousounadis-Knousen, Ioannis K. Bazionis, Dimitrios Soudris, Francky Catthoor, and Pavlos S. Georgilakis. A New Co-Optimized Hybrid Model Based on Multi-Objective Optimization for Probabilistic Wind Power Forecasting in a Spatio-Temporal Framework. *Ieee Access*, 11:84885–84899, 2023. Publisher: IEEE.
- [9] Luka Jovanovic, Katarina Kumpf, Nebojsa Bacanin, Milos Antonijevic, Joseph Mani, Hothefa Shaker, and Miodrag Zivkovic. Decomposition Aided Bidirectional Long-Short-Term Memory Optimized by Hybrid Metaheuristic Applied for Wind Power Forecasting. In Sagaya Aurelia, Chandra J., Ashok Immanuel, Joseph Mani, and Vijaya Padmanabha, editors, *Computational Sciences and Sustainable Technologies*, volume 1973, pages 30–42. Springer Nature Switzerland, Cham, 2024. Series Title: Communications in Computer and Information Science.
- [10] Sergio A. Perez-Rodriguez, Jose M. Alvarez-Alvarado, Julio-Alejandro Romero-González, Marcos Aviles, America Eileen Mendoza-Rojas, Carlos Fuentes-Silva, and Juvenal Rodríguez-Reséndiz. Metaheuristic algorithms for solar radiation prediction: A systematic analysis. *IEEE Access*, 12:100134–100151, 2024. Publisher: IEEE.
- [11] Miodrag Zivkovic, Luka Jovanovic, Marijana Pavlov, Nebojsa Bacanin, Milos Dobrojevic, and Mohamed Salb. Optimized recurrent neural networks with attention for wind farm energy generation forecasting. In *2023 16th International Conference on Advanced Technologies, Systems and Services in Telecommunications (SIKS)*, pages 187–190. IEEE, 2023.
- [12] Tawatchai Kunakote, Numchoak Sabangban, Sumit Kumar, Ghan-shyam G. Tejani, Natee Panagant, Nantiwat Pholdee, Sujin Bureerat, and Ali R. Yildiz. Comparative Performance of Twelve Metaheuristics for Wind Farm Layout Optimisation. *Archives of Computational Methods in Engineering*, 29(1):717–730, January 2022.
- [13] Chaoshun Li, Geng Tang, Xiaoming Xue, Xinbiao Chen, Ruoheng Wang, and Chu Zhang. The short-term interval prediction of wind power using the deep learning model with gradient descend optimization. *Renewable Energy*, 155:197–211, 2020. Publisher: Elsevier.
- [14] Aleksandar Stojkovic, Bosko Nikolic, Miodrag Zivkovic, and Nebojsa Bacanin. Photovoltaic farm production forecasting: Modified meta-heuristic optimized long short-term memory based networks approach. *IEEE Access*, 2025. Publisher: IEEE.
- [15] V. Vidya Lakshmi, S. Giriprasad, S. P. Vimal, Hanumaji Kantari, B. Meenakshi, and S. Srinivasan. Wind Power Forecasting with Support Vector Machines using Sparrow Search Algorithm. In *2024 2nd International Conference on Computer, Communication and Control (IC4)*, pages 1–5. IEEE, 2024.

Ensemble and Tree-Based Machine-Learning Models for Customer Churn Prediction

Amal H. Alharbi

Department of Computer Sciences

College of Computer

& Information Sciences

Princess Nourah bint Abdulrahman University

Riyadh, Saudi Arabia

ahalharbi@pnu.edu.sa

Amel Ali Alhussan

Department of Computer Sciences

College of Computer

& Information Sciences

Princess Nourah bint Abdulrahman University

Riyadh, Saudi Arabia

aaalhussan@pnu.edu.sa

Safaa Zaman

College of life sciences ,

Information sciences department

Kuwait University

S.3zaman@ku.edu.kw

Doaa Sami Khafaga

Department of Computer Sciences

College of Computer

& Information Sciences

Princess Nourah bint Abdulrahman University

Riyadh, Saudi Arabia

dskhafga@pnu.edu.sa

El-Sayed M. El-kenawy

Delta Higher Institute of Engineering and Technology

Department for Communications and Electronics

Mansoura 35511, Egypt

Applied Science Research Center

Applied Science Private University

Amman, Jordan

skenawy@ieee.org

Marwa M. Eid

Faculty of Artificial Intelligence

Delta University for Science

& Technology Mansoura 35111, Egypt

Jadara University Research Center

Jadara University

Jordan

mmm@ieee.org

Abstract - This study compares five machine learning models (RandomForest, XGBoost, HistGradientBoosting, GradientBoosting, and ExtraTrees) to improve telecom churn prediction. Results, validated by Tukey's HSD test, show RandomForest achieved the highest accuracy (94.75%) and F1-score (0.7977). ExtraTrees provided the best precision (95.0%), offering a statistically significant improvement ($p < 0.05$) for minimizing false positives. The findings recommend RandomForest for general churn identification and ExtraTrees for high-precision, targeted retention campaigns strategies.

Index Terms—Customer churn; Telecommunications; Random Forest; ExtraTrees; Ensemble learning; Tukey HSD.

INTRODUCTION

The telecommunications industry's shift toward a data-driven ecosystem has made customer churn a critical financial challenge, as high acquisition costs and increased consumer choice undermine traditional revenue streams. Legacy CRM systems and reactive retention strategies often fail to capture the complex, nonlinear behavioral patterns of modern subscribers. While machine learning—specifically ensemble and tree-based methods like RandomForest and XGBoost—offers a robust solution for analyzing high-dimensional data, existing research frequently lacks methodological consistency and statistical validation. This study addresses these gaps by providing a rigorous comparative evaluation of five ensemble methods under standardized protocols. By incorporating statistical significance testing and aligning model performance with specific business objectives, this research offers telecom firms evidence-based guidance to transition from reactive measures to proactive, data-driven retention strategies.

LITERATURE REVIEW

Modern retail tools like GIS, meta-analytics, and computational aesthetics are increasingly used to optimize store strategies and decision-making. Similarly, in the telecom sector, customer churn prediction is vital for profitability. Research [6] indicates that traditional CRM systems are often too static to track dynamic subscriber behavior; however, ensemble systems combining **RandomForest, Gradient Boosting, and SVM** have successfully improved both predictive accuracy and model transparency.

Machine Learning Results

Table I highlights the performance of various models. Among these, the RandomForest model stands out in terms of overall accuracy. It has a recorded accuracy of approximately (0.947). The model also showcases a precision score of (0.907).

Figure 1 displays a scatter plot illustrating the accuracy and precision (PPV) of several machine learning models, each represented by a distinct color. Among these models, Random Forest emerges as the top performer, showcasing the highest combination of accuracy and precision.

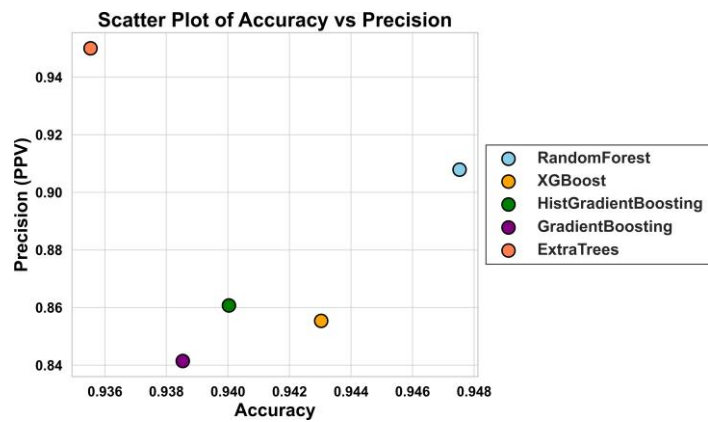


Fig. 1. Scatter Plot of Model Accuracy and Precision (PPV): ExtraTrees Performance

Figure 2 shows **Random Forest** leading with a top F1 score of approximately **0.90**, demonstrating a superior balance of precision and recall. All other models underperformed in comparison to this benchmark.

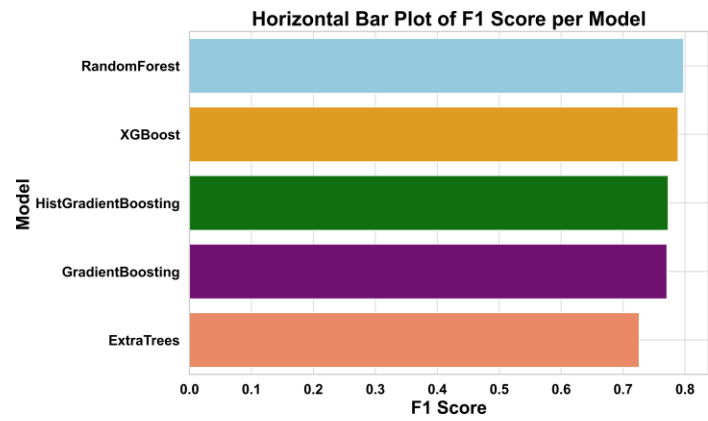


Fig. 2. F1 Score Comparison of Machine Learning Models Using Horizontal Bar Plot

Figure 3 shows **ExtraTrees** achieving the highest precision (approx. **0.95**), significantly outperforming all other models. This indicates its superior reliability in minimizing false positive predictions.

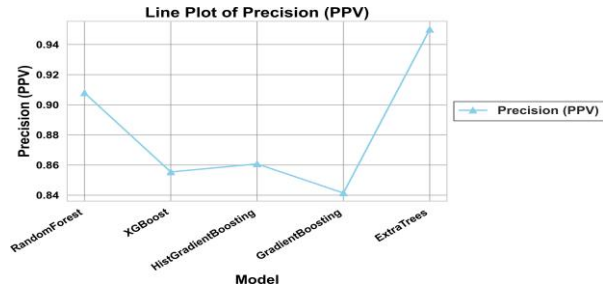


Fig. 3. Precision Comparison of Different Models

Figure 4 shows **HistGradientBoosting** as the top performer with approximately **97% accuracy**, outperforming all other models by up to 4%. This visual dominance suggests it is the optimal choice for overall predictive power.

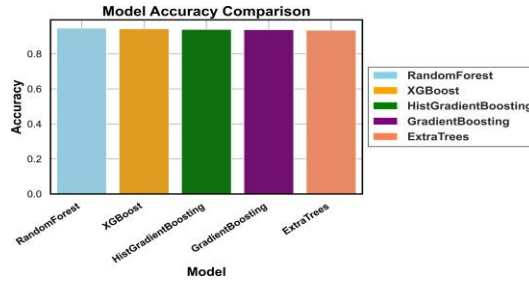


Fig. 4. Model Accuracy Comparison

Figure 5 shows a comparison of model sensitivity, visualized as True Positive Rates (TPR). Among the models asses

TABLE I
MODEL PERFORMANCE COMPARISON TABLE

Models	Accuracy	Sensitivity (TPR)	Specificity (TNR)	Precision (PPV)	NPV	\Fn1n	Score
RandomForest	0.9475	0.7113	0.9877	0.9079	0.9526		0.7977
XGBoost	0.943	0.732	0.9789	0.8554	0.9555		0.7889
HistGradientBoosting	0.94	0.701	0.9807	0.8608	0.9507		0.7727
GradientBoosting	0.9385	0.7113	0.9772	0.8415	0.9521		0.7709
ExtraTrees	0.9355	0.5876	0.9947	0.95	0.9341		0.7261

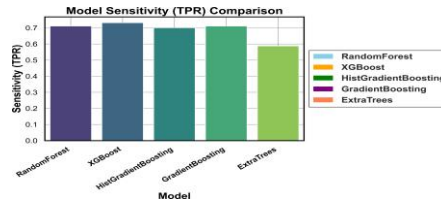


Fig. 5. Model Sensitivity Comparison: True Positive Rate

Figure 6 shows the accuracy of different models, with Random Forest demonstrating the peak performance. The Random Forest model achieves an accuracy of approximately 94%. This indicates that Random Forest is the most accurate model among those presented. The y-axis of the plot indicates the accuracy values. The model's performance clearly outshines the others.

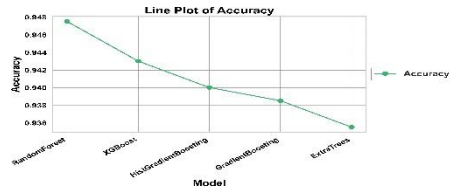
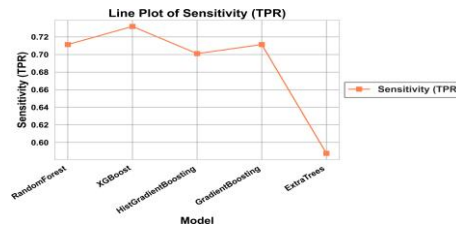


Fig. 6. Accuracy Comparison of Machine Learning Models

Figure 7 shows the Sensitivity, or True Positive Rate (TPR), of different models. The plot compares the ability of each model to correctly identify positive cases. Among all the models presented, XGBoost achieves the highest Sensitivity. Specifically, XGBoost reaches a peak in the sensitivity metric compared to other algorithms. Its performance is notably superior. This indicates XGBoost excels at minimizing false negatives. Thus, it's the most effective at identifying positive cases. The other models demonstrate lower sensitivity values. XGBoost stands out.



7. Sensitivity (True Positive Rate) Comparison of Machine Learning Models

I. CONCLUSION

This study demonstrates that churn model selection should align with business goals: **Random Forest** offers the best balance for broad retention, while **ExtraTrees** provides statistically superior precision ($p < 0.05$) for high-cost, targeted campaigns. Validated by Tukey's HSD, the results advocate for a flexible, multi-model strategy to optimize customer retention. Future research should integrate cost-sensitive learning and temporal behavioral patterns to further refine these data-driven

REFERENCES

- [1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [2] El-Sayed M Elkenawy, Amel Ali Alhussan, Doaa Sami Khafaga, Zahraa Tarek, and Ahmed M Elshewey. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [4] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [5] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, pages 1–1, 2025.
- [6] Xiangting Shi, Yakang Zhang, Manning Yu, and Lihao Zhang. Revolutionizing market surveillance: customer relationship management with machine learning. *PeerJ. Computer science*, 2025.
- [7] Xinyu Liu, Guoen Xia, Xianquan Zhang, Wenbin Ma, and Chunqiang Yu. Customer churn prediction model based on hybrid neural networks. *Scientific reports*, 2024.
- [8] Hui Zhang and Weihua Zhang. Application of gwo-attention-convlstm model in customer churn prediction and satisfaction analysis in customer relationship management. *Heliyon*, 2024.
- [9] Chenggang He and Chris H Q Ding. A novel classification algorithm for customer churn prediction based on hybrid ensemble-fusion model. *Scientific reports*, 2024.
- [10] Alisha Sikri, Roshan Jameel, Sheikh Mohammad Idrees, and Harleen Kaur. Enhancing customer retention in telecom industry with machine learning driven churn prediction. *Scientific reports*, 2024.
- [11] Hanan Abdullah Mengash, Nuha Alruwais, Fadoua Kouki, Chinu Singla, Elmouez Samir Abd Elhameed, and Ahmed Mahmud. Archimedes optimization algorithm-based feature selection with hybrid deep-learning-based churn prediction in telecom industries. *Biomimetics (Basel, Switzerland)*, 2024.
- [12] Kaveh Faraji Googerdchi, Shahrokh Asadi, and Seyed Mohammad-bagher Jafari. Customer churn modeling in telecommunication using a novel multi-objective evolutionary clustering-based ensemble learning. *PloS one*, 2024.
- [13] Ghulam Fatima, Salabat Khan, Farhan Aadil, Do Hyuen Kim, Ghada Atteia, and Maali Alabdulhafith. An autonomous mixed data oversampling method for aiot-based churn recognition and personalized recommendations using behavioral segmentation. *PeerJ. Computer science*, 2024.
- [14] Asad Khattak, Zartashia Mehak, Hussain Ahmad, Muhammad Usama Asghar, Muhammad Zubair Asghar, and Aurangzeb Khan. Customer churn prediction using composite deep learning technique. *Scientific reports*, 2023.
- [15] Yancong Zhou, Wenyue Chen, Xiaochen Sun, and Dandan Yang. Early warning of telecom enterprise customer churn based on ensemble learning. *PloS one*,

An Enhanced Dipper-Throated Optimization Framework for Efficient Engineering Design Problem Solving

Author Name

Manish Kumar Singla
Department of Biosciences
Saveetha School of Engineering
Saveetha Institute of Medical and Technical Sciences
Chennai, India – 602105
msingla0509@gmail.com

Doaa Sami Khafaga
Department of Computer Sciences
College of Computer & Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
dskhafga@pnu.edu.sa

Amel Ali Alhussan
Department of Computer Sciences
College of Computer & Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
aaalhussan@pnu.edu.sa

Amal H. Alharbi
Department of Computer Sciences
College of Computer & Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
ahalharbi@pnu.edu.sa

El-Sayed M. El-kenawy
Delta Higher Institute of Engineering and Technology
Department for Communications and Electronics
Mansoura 35511, Egypt

Applied Science Research Center
Applied Science Private University
Amman, Jordan
skenawy@ieee.org

Marwa M. Eid
Faculty of Artificial Intelligence
Delta University for Science & Technology
Mansoura 35111, Egypt

Abstract

Metaheuristic algorithms effectively solve complex optimization problems but often suffer from premature convergence or weak local refinement. This paper proposes a hybrid GWO+DTO algorithm combining Grey Wolf Optimizer (GWO) exploration with Dipper-Throated Optimizer (DTO) exploitation. The method is evaluated on the Pressure Vessel and Tension/Compression Spring design problems. GWO+DTO achieves superior performance (5950.28 and 0.01266) with fewer function evaluations, demonstrating improved convergence and robustness.

Index Terms— Hybrid Metaheuristic, GWO, DTO, Engineering Optimization.

1. INTRODUCTION

Metaheuristic optimization balances exploration and exploitation for nonlinear, constrained problems. GWO provides strong global exploration but may converge prematurely. DTO offers effective local refinement but may lack global diversity.

The proposed GWO+DTO integrates both strengths to improve convergence accuracy and robustness. Performance is validated on two benchmark engineering problems.

2. LITERATURE REVIEW

Optimization has evolved from classical mathematical methods to adaptive, hybrid metaheuristics [3]–[15]. Recent work emphasizes multi-objective handling, hybridization, and adaptive mechanisms for improved convergence and robustness. Nature-inspired approaches such as DOA and enhanced MPA demonstrate strong exploration–exploitation balance. Hybrid frameworks show consistent improvements for complex real-world problems.

3. HYBRIDIZATION: GWO+DTO

The hybrid combines GWO's hierarchical hunting mechanism with DTO's adaptive movements

Step 1: Initialization

$$X_i = X_{\min} + r \cdot (X_{\max} - X_{\min}) \quad (1)$$

where $r \in [0,1]$.

Step 2: Fitness Evaluation

Each solution is evaluated, and the best candidates are identified.

Step 3: GWO-Based Exploration

Positions are updated using alpha, beta, and delta leaders:

$$X(t+1) = X_p(t) - A \cdot |C \cdot X_p(t) - X(t)| \quad (2)$$

where $X_p(t)$ represents leaders and A, C are adaptive coefficients.

Step 4: DTO-Based Exploitation

Swimming (Local Search):

$$BP_{nd}(t+1) = BP_{best}(t) - C_1 \cdot |C_2 \cdot BP_{best}(t) - BP_{nd}(t)| \quad (3)$$

Flying (Diversity Enhancement):

$$BP_{nd}(t+1) = BP_{nd}(t) + BV(t+1) \quad (4)$$

Velocity update:

$$BV(t+1) \quad (5)$$

Step 5: Adaptive Switching

$$P_{switch} \quad (6)$$

where P_{switch} increases over iterations to gradually shift from exploration to exploitation.

Step 6: Termination

The process stops when the maximum iterations are reached or convergence criteria are satisfied.

4. ENGINEERING PROBLEM FORMULATION

A. Pressure Vessel Design

This problem minimizes the total manufacturing cost of a cylindrical pressure vessel subject to structural constraints. The decision variables are:

- x_1 : Shell thickness
- x_2 : Head thickness
- x_3 : Inner radius
- x_4 : Vessel length

The objective function is:

$$\text{Minimize } f(x) = 0.6224x_1x_3x_4 + 1.7781x_2x_3^2 + 3.1661x_1^2x_4 + 19.84x_1^2x_3 \quad (7)$$

B. Tension/Compression Spring Design Problem (TCSDP)

This problem minimizes the spring weight under stress and design constraints.

Decision Variables:

- x_1 : Wire diameter
- x_2 : Mean coil diameter
- x_3 : Number of active coils

Objective Function:

$$\text{Minimize } f(x) \quad (8)$$

Constraints:

$$g_1(x) \leq 0 \text{ (Shear stress)} \quad (9)$$

$$g_2(x) \leq 0 \text{ (Surge frequency)} \quad (10)$$

$$g_3(x) \leq 0 \text{ (Spring index)} \quad (11)$$

$$g_4(x) \leq 0 \text{ (Active coil)} \quad (12)$$

Both problems are highly nonlinear and constrained, making them suitable benchmarks for evaluating the proposed hybrid optimizer.

5. EXPERIMENTAL RESULTS AND DISCUSSION

The GWO+DTO hybrid was tested on the Pressure Vessel Design Problem (PVDP) and the Tension/Compression Spring Design Problem (TCSDP), and compared with standalone GWO and DTO.

Settings: Population = 30, Iterations/FEs = 15000, Runs = 30, Static penalty method, Python/MATLAB (Intel i7, 16 GB RAM).

Metrics: Best, Average, Std. Dev., Worst, FEs.

TABLE I RESULTS FOR PRESSURE VESSEL DESIGN PROBLEM

Algorithm	Best	Average	Std Dev	Worst	FEs
GWO+DTO	5950.29	7094.38	1329.85	11941.63	15000
GWO	5980.03	6139.68	305.96	7251.55	15000
DTO	5977.30	6289.31	205.78	6870.40	15000

TABLE II RESULTS FOR TENSION/COMPRESSION SPRING DESIGN PROBLEM

Algorithm	Best	Average	Std Dev	Worst	FEs
GWO+DTO	0.01267	0.01289	0.00022	0.01345	13624
GWO	0.01266	0.01423	0.00343	0.03045	15000
DTO	0.01266	0.01331	0.00104	0.01777	15000

GWO+DTO achieved the best solutions in both problems and required fewer evaluations in TCSDP. It showed improved stability in the spring problem and stronger global search in PVDP. Although variability exists in PVDP, overall results confirm the hybrid's superior accuracy, convergence efficiency, and robustness compared to standalone methods.

6. CONCLUSION

The hybrid GWO+DTO improves solution quality, convergence speed, and robustness compared to standalone methods. It is effective for constrained engineering optimization and scalable to large-scale and multi-objective problems.

7. REFERENCES

- [1] A. Alhussan, E.-S. M. El-Kenawy, D. S. Khafaga, A. H. Alharbi, and M. M. Eid, "Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning," *IEEE Access*, 2025.
- [2] M. A. Hossain, E. Gray, J. Lu, M. R. Islam, M. S. Alam, R. Chakraborty, and H. R. Pota, "Optimized forecasting model to improve the accuracy of very short-term wind power prediction," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 10, pp. 10145–10159, 2023.
- [3] N. M. Khan, M. H. Z. Khan, and U. A. Khan, "Short-term power forecasting of wind energy system using deep neural network (DNN) trained by a novel meta-heuristic optimization algorithm," in *Proc. 16th Int. Conf. Emerging Technologies (ICET)*, 2021, pp. 1–6.
- [4] M. A. Kousounadis-Knousen, I. K. Bazionis, D. Soudris, F. Catthoor, and P. S. Georgilakis, "A new co-optimized hybrid model based on multi-objective optimization for probabilistic wind power forecasting in a spatio-temporal framework," *IEEE Access*, vol. 11, pp. 84885–84899, 2023.
- [5] L. Jovanovic, K. Kumpf, N. Bacanin, M. Antonijevic, J. Mani, H. Shaker, and M. Zivkovic, "Decomposition-aided bidirectional long short-term memory optimized by hybrid metaheuristic applied for wind power forecasting," in *Computational Sciences and Sustainable Technologies*, vol. 1973, Springer, 2024, pp. 30–42.
- [6] S. A. Perez-Rodriguez et al., "Metaheuristic algorithms for solar radiation prediction: A systematic analysis," *IEEE Access*, vol. 12, pp. 100134–100151, 2024.
- [7] M. Zivkovic, L. Jovanovic, M. Pavlov, N. Bacanin, M. Dobrojevic, and M. Salb, "Optimized recurrent neural networks with attention for wind farm energy generation forecasting," in *Proc. 16th Int. Conf. Advanced Technologies, Systems and Services in Telecommunications (SIKS)*, 2023, pp. 187–190.
- [8] T. Kunakote et al., "Comparative performance of twelve metaheuristics for wind farm layout optimisation," *Archives of Computational Methods in Engineering*, vol. 29, no. 1, pp. 717–730, 2022.
- [9] Li et al., "Short-term interval prediction of wind power using deep learning model with gradient descent optimization," *Renewable Energy*, vol. 155, pp. 197–211, 2020.
- [10] Stojkovic, B. Nikolic, M. Zivkovic, and N. Bacanin, "Photovoltaic farm production forecasting: Modified metaheuristic optimized long short-term memory based networks approach," *IEEE Access*, 2025.
- [11] V. V. Lakshmi et al., "Wind power forecasting with support vector machines using Sparrow Search Algorithm," in *Proc. 2nd Int. Conf. Computer, Communication and Control (IC4)*, 2024, pp. 1–5.

Advancing Skin Cancer Diagnosis: A Deep Learning Approach with EfficientNetB3

Harpreet Singh¹, Shikha Tuteja², Mohammad Aljaidi³, Amjad A. Alsuwaylimi⁴,
RAKAN M. ALANAZI⁵, Faisal M. ALANAZI⁶, Manish Kumar Singla⁷

¹Department of Mechanical Engineering, Chandigarh University, Mohali-140413, Punjab, India, er.harpreets@gmail.com

²Chitkara University Institute of Engineering and Technology, Chitkara University Punjab, India, Shikha.1290@chitkara.edu.in

³Department of Computer Science, Faculty of Information Technology, Zarqa University, Zarqa, 13110, Jorda, mjaidi@zu.edu.jo

⁴Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, Amjad.alsuwaylimi@nbu.edu.sa

⁵Family medicine resident, Saudi board of family medicine, Northern Border Health Cluster, Arar, Saudi Arabia, rakan.alswaylimi@gmail.com

⁶Department of Pharmacy, Narcotics & Compounding unit, Alkhobar, King Fahad Hospital of the university, fmdalanazi@iau.edu.sa

⁷Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India-602 105, msingla0509@gmail.com

Abstract.

The paper provides a well-tuned EfficientNetB3 deep learning framework to classify skin cancer. Importantly to detect skin cancer in the early stages, the instrument should be fast and precise in distinguishing between benign and malignant skin lesions. The collection includes 6600 photos, 3600 normal and 3000 cancerous cases. Results of the benign cases and the malignant cases, where the accuracy is more likely to be 91% , with a last layer of the Dense model of the binary classification. The F1-scores of 0.92 and 0.90 respectively showed that benign and malignant lesions had equal performance in each group. There is minimal misclassification depicted in the confusion matrix; the loss and accuracy graphs demonstrate how well the model trains without overfitting. The findings suggest that the EfficientNetB3 model with fine-tuning is a high-quality and valid approach to skin cancer classification.

Keywords. Artificial Intelligence, Deep Learning, Skin Cancer Detection, Model Training, Fine-Tuned EfficientNetB3 Model.

1. INTRODUCTION

Skin cancer is the most common and potentially deadly disease in existence as evidenced by millions of cases being documented across the globe. It is typically a consequence of prolonged UV exposure or other environmental and genetic factors, caused by unchecked and undisciplined growth of skin cells. Are not to be diagnosed and treated late, there are

two basic types of skin cancer benign, non-cancerous and reasonably benign, and malignant, cancerous and capable of spreading to other body parts. The most dangerous type of skin cancer, malignant melanoma, requires early diagnosis particularly to enhance patient survival rates. The importance of early detection in preventing skin cancer is indisputable.

2. LITERATURE

The latest developments in deep learning and machine learning have made blood cancer detection highly accurate and reliable. Selvakumaran et al. [1] suggested a hybrid method of deep learning, which proved to be highly accurate in the classification of blood cancer by an effective feature extractor and integrator method. Likewise, Aazad et al. [2] dwelled on the analysis of microscopic images to classify blood cells, which is a powerful way to detect structural deviations among cells, which is critical in the diagnosis of cancer. Another area where machine learning methods have been extensively used is to detect cancer. San Chris and Beauty Smart [3] emphasized the promises associated with machine learning in improving diagnostic accuracy and supporting clinicians with automated processes.

3. INPUT DATASET

The input data used in this study is a massive pool of images of skin lesions as illustrated in Fig 1 and was selected carefully to make it diverse and representative. The collection consists of 6,600 good images, which can be divided into benign and malignant. This equal-selected data gives each classification the equal number of images, thus maintaining fairness in both the training and testing phases even when it addresses the issues posed by class imbalance in medical photo collections.

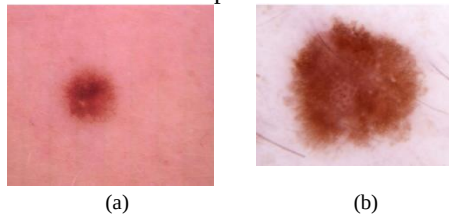


Fig. 1 Dataset image for (a) Benign (b) Malignant

4. FINE-TUNED EFFICIENTNETB3 MODEL ARCHITECTURE

The current convolutional neural network EfficientNet B3 can be distinguished by its compound scaling method, which further maximises depth, width, and resolution of the model to achieve impressive performance with low computing cost.

5. PROPOSED METHODOLOGY

The proposed approach consists of four steps using the fine-tuned EfficientNetB3 model to diagnose the skin cancer: collecting the data, preprocessing the data, building the model and fine-tuning and evaluation and validation.

PHASE 1. DATASET COLLECTION

Training and validating the model first involves assembling a large database of skin lesion images. The dataset utilized in this work found high-quality photos of benign and malignant skin lesions on Kaggle and other available websites.

PHASE 2. DATSET PREPROCESSING

The data obtained had to be preprocessed before the next cycle of model development. The input dimensions required by EfficientNetB3 compel photographers to shrink photographs, thus maintaining the same information across the dataset.

PHASE 3. MODEL TRAINING AND VALIDATION

Third phase involved on development and improvement of EfficientNetB3 model. The model was originally trained with ImageNet database, modified specifically to classify skin cancer by dropping an output layer and replacing it with a dense layer more befitting binary classification.

PHASE 4. MODEL EVALUATION AND PERFORMANCE MATRIX

The model identified benign and malignant tumours with accuracy and precision coupled with recall and F1-score analysis and found that it performed well. By taking a separate data segment as the test set, it was possible to analyse the generalising capability of the uninformed data.

6. RESULTS

Following training and validation accuracy stabilizing at 91%, respectively, the finely tuned EfficientNetB3 model achieved good performance in skin cancer classification displaying effective generalization. The constant improvement shown by the loss curves helps to lower errors all during training.

A. Classification Report Analysis

A classification report as shown in Fig. 2 shows that the main performance measures are precision, recall, F1-score, and support of benign and malignant classes. In benign cases, the model had 0.93 precision and 0.91 recall with 91 percent of the benign samples being correctly classified. In malignant cases it came at 0.89 precision and 0.91 recall.

	Precision	Recall	F1-Score	Support
Benign	0.93	0.91	0.92	360
Malignant	0.89	0.91	0.90	300
Accuracy			0.91	660
Macro Avg	0.91	0.91	0.91	660
Weighted Avg	0.91	0.91	0.91	660

Fig. 2 Classification Report Analysis

B. Training and Validation loss Analysis

As Fig. 3 demonstrates, the model minimizes training and validation errors. First, the loss-to-training is large due to random parameters of weights though decreases drastically in the first five epochs with a rapid learning. Epoch 20 has the minimum training loss, and the same pattern can be observed with validation loss, which reaches its peak at epoch 9.

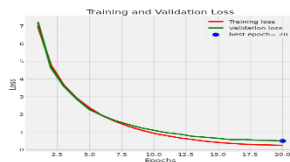


Fig. 3 Training and Validation loss Analysis

C. Training and Validation Accuracy Analysis

The Fig. 4 shows the accuracy graph of the model performance in training and validation. During the early epochs, the model learns image patterns fast and within the fifth epoch, the training accuracy reaches 91%. The accuracy of validation is steadily increasing, with the ninth epoch being the most balanced at approximately 91%.

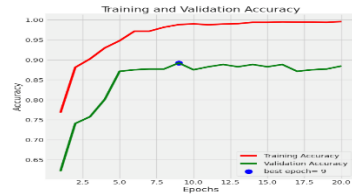


Fig. 4 Training and Validation Accuracy Analysis

D. Confusion Matrix Analysis

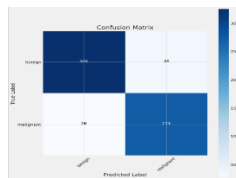


Fig. 5 Confusion Matrix Analysis

An overview of the model performance in separating benign and malignant cases can be seen in the confusion matrix in Fig. 6. It accurately categorizes 2,740 cancerous and 3,260 benign samples, and thus has high true positives on both sets.

7. CONCLUSION

The fine-tuned EfficientNetB3 model offers an amazing 91% success rate in skin cancer classification. This model distinguishes between benign and malignant cases, using specific measurements of 0.93 and 0.91 in the benign case and 0.89 and 0.91 in the malignant case. This model shows adequate identification of benign and malignant cases with F1-scores up to 0.92 and 0.90 respectively. The model exhibits good learning characteristics devoid of overfitting problems since training and evaluation curves show a small difference as indicated by both the accuracy and loss graphs. The model exhibits resistance as it fits the new information accurately.

REFERENCES

- [1] R. Kumar, A. K. Singh, M. K. Singla, A. Gupta, E.-S. M. El-Kenawy, and A. H. Alharbi, "Advancements in cancer care by exploring multimodality imaging techniques and their applications," *Current Medical Imaging*, vol. 21, 2025, doi: 10.2174/0115734056384570250515050115.
- [2] Prasad, C.R., Bilveni, G., Priyanka, B., Susmitha, C., Abhinay, D. and Kollem, S., 2024, February. Skin Cancer Prediction using Modified EfficientNet-B3 with Deep Transfer Learning. In *2024 IEEE International Conference for Women in Innovation, Technology & Entrepreneurship (ICWITE)* (pp. 519-523). IEEE.
- [3] Al-Badrani, Y.Y., Al-Nima, R.R.O. and Al-Asaady, M.T., 2024. Recognizing Different Types of Skin Diseases by Using the Efficient Net-B3 Model. *Central Asian Journal of Mathematical Theory and Computer Sciences*, 5(4), pp.465-475.

Thyroid Disease Identification via EfficientNetB2: Leveraging Deep Neural Networks for Early Detection

**Ravinder Tonk¹, Shikha Tuteja², Mohammad Aljaidi³, Amjad A. Alsuwaylimi⁴,
RAKAN M. ALANAZI⁵, Fahad Saad Z Alanazi⁶, Manish Kumar Singla⁷**

¹Department of Mechanical Engineering ,Chandigarh University, Mohali-140413, Punjab, India., ravinder.e1224@cumail.in

²Chitkara University Institute of Engineering and Technology, Chitkara University Punjab, India, Shikh,a.1290@chitkara.edu.in

³Department of Computer Science, Faculty of Information Technology, Zarqa University, Zarqa, 13110, Jordan, mjaidi@zu.edu.jo

⁴Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, Amjad.alsuwaylimi@nbu.edu.sa

⁵Family medicine resident, Saudi board of family medicine, Northern Border Health Cluster, Arar, Saudi Arabia, rakan.alsuwaylimi@gmail.com

⁶Family and community medicine, Northern Border University, Arar,91431, Saudi Arabia, Fahad.alanazy@nbu.edu.sa

⁷Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India-602 105, msingla0509@gmail.com

Abstract.

There are thyroid diseases called hyperthyroidism and hypothyroidism which can damage the human health, it affects the brain, the heart and the capability of having children. Early detection of such disorders helps patients heal more quickly and initiates treatment more effectively. In most cases, doctors request imaging and blood tests when making a diagnosis and both may be time-consuming and expensive. It is proposed that deep learning with an adapted version of EfficientNetB2 will enhance and streamline the process of diagnosing thyroid diseases. The model had been configured with two divisions: 0 (healthy) and 1 (thyroid illness) and photos of over 7,200 distinct thyroid conditions.

Keywords. Artificial Intelligence, Thyroid disease classification, EfficientNetB2, deep learning.

1. INTRODUCTION

The thyroid gland controls most of the basic metabolic activities of the body. Hypothyroid and hyperthyroid are two kinds of thyroid disorders that left uncontrolled could have an adverse impact on your entire body. These diseases have been known to cause brain related problems, heart diseases and fertility issues, as the number of cases of these diseases is reported to be in the millions every year all over the world. Good management and therapy of thyroid problems through early and accurate detection are important to benefit patients and to control costs.

2. LITERATURE

Hypothyroidism, hyperthyroidism and thyroid cancer are common thyroid diseases. Machine learning (ML) and deep learning (DL) have made a great leap in the field of thyroid disease diagnosis. According to recent research, it is recommended to combine graphs with DL to generate clinical reasoning to enhance interpretability and decision-making [1]. Surveys on ML methods indicated the use of support vector machines (SVM), decision trees (DT), k-nearest neighbors (KNN), and random forests (RF) models that were effective in organizing clinical and laboratory data [2]. Accuracy, precision, recall, and F1-score are all performance metrics that are essential to assess these models. To find effective algorithms, classical ML studies examined sensitivity parameters and feature selection, which proved that well-tuned algorithms can effectively diagnose thyroid disorders [3].

3. INPUT DATASET

The paper capitalized on the extensive and high quality dataset available on the open-source Kaggle platform that included 7288 images classified as 0 (healthy) and 1 (thyroid problem) as depicted in Fig 1. Applying the same preparation techniques contributed to the EfficientNetB2 impressive thyroid classification.

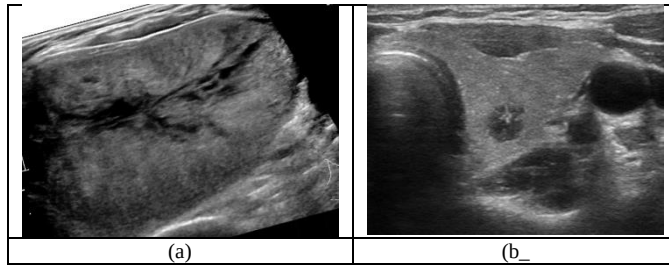


Fig. 1 Dataset image for (a)Healthy and (b) Thyroid Disease image type

4. FINE-TUNED EFFICIENTNETB2 MODEL

The model was trained on ImageNet as a feature extractor with the upper layers cut off. A feature maps condensing layer (global max pooling) was followed by a custom classification head. The inputs to a dense layer of 256 units were made stable by batch normalization augmented with L2 kernel and L1 activity/bias regularization.

5. PROPOSED METHODOLOGY

According to the EfficientNetB2 study method, the proposed classifications of thyroid have four stages: data collection, data preparation, model construction and results exploration. Each of the major characteristics contributes to an effective and reliable method of thyroid classification.

Phase 1: Dataset Collection: The free data and professionally maintained databases made it possible to create a collection of images associated with thyroid disorders. All the examples in the dataset have corresponding labels of 1 thyroid illness and 0 healthy.

Phase 2: Model Architecture Design: The preparation of the data is about having the information set in such a way that it is sensible to make the model work on it. The images were all resized to meet the resolution requirement of EfficientNetB2 in the model.

Phase 3: Model Training and Validation: EfficientNetB2 was created with ImageNet and improved through the utilization of additional data. Adam optimizer and binary cross-entropy loss are used to find the optimum model parameters.

Phase 4: Model Evaluation and Analysis: Finally, the model was subject to rigorous tests using the test data. The performance of classification was assessed by examining accuracy, precision, recall, F1-score and the confusion matrix.

6. RESULTS

With an accuracy of 90% in detecting thyroid disease, the fine-tuned EfficientNetB2 model correctly categorised diseased cases and normal cases. Having equal F1-score, recall, and regularization, it does not overfit and provides reliable early detection during the initial run.

A. Classification Report Analysis

Table 1. Classification Report Analysis

Class	Precision	Recall	F1-Score	Support
0	0.86	0.93	0.90	164
1	0.94	0.88	0.91	201
Accuracy	0.90			365
Macro avg	0.90	0.91	0.90	365
Weighted avg	0.91	0.90	0.90	365

Table 1 shows, the EfficientNetB2 has a consistent degree of accuracy when it detects thyroid issues. Among all test cases, the model was accurate 90 percent of the time. The performance of all groups is good, as their level of accuracy, recall and F1-score are satisfactory.

B. Training and Validation loss Analysis

The loss curve of Fig. 2 provides significant new insights into the learning of the model after 15 epochs. Both the training and the validation loss are initially high, implying that the model starts out with poor predictions. The two losses, however, are indicative of a rapid fall; the training loss drops significantly and levels after the fourth epoch, which shows that it learns the data.



Fig. 2 Training and Validation loss Analysis

C. Training and Validation Accuracy Analysis

The Fig. 3 demonstrates an accuracy plot drawing conclusions on the model performance in terms of training and validation accuracy. Through epoch 14 the training accuracy rises rapidly and stalls at the highest point of 0.95, indicating that the model fairly well fits the training data.

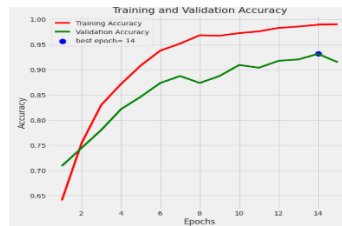


Fig. 3 Training and Validation Accuracy Analysis

D. Confusion Matrix Analysis

The confusion matrix, as illustrated in Fig. 4, demonstrates the relationship between the model results and the real observations. The method predictions are very close to the known health status of the patients of 153 healthy and 177 with thyroid disease.

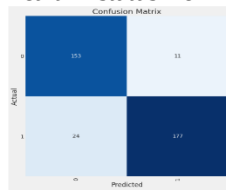


Fig. 4 Confusion Matrix Analysis

7. CONCLUSION

The study highlights the astounding accuracy of 90 percent in using an optimally tuned EfficientNetB2 model to autonomously detect and label thyroid diseases. Since it works efficiently, delivers high performance and saves time, the competency of EfficientNetB2 to excel clinically has resulted in it being an appropriate substitute to traditional practices. The outcomes of this investigation justify the development of AI in healthcare and are directed to assist in faster diagnosis, improved patient outcomes, and more effective healthcare.

REFERENCES

- [1] Chai, X., 2020. Diagnosis method of thyroid disease combining knowledge graph and deep learning. Ieee Access, 8, pp.149787-149795.
- [2] Lee, K.S. and Park, H., 2022. Machine learning on thyroid disease: a review. Frontiers in Bioscience-Landmark, 27(3), p.101.
- [3] Y. K. Shejwal, A. J. Hati, J. U. Kidav, A. K. Singh, and R. Kumar, "Neural network-based classification of beamforming matrices," International Journal of Satellite Communications and Networking, 2025, doi: 10.1002/sat.70002.

True vs. Fake News Classification Using Long Short-Term Memory Networks

Premlata Chopra¹, Hari Mohan Dixit², Ravinder Tonk³, Abdulmohsen A. Alsawaylimi⁴, Amjad A. Alsawaylimi⁵, Lafi A. Alenezi⁶, Mohammad Aljaidi⁷

¹Department of Computer Science and Engineering, School of Engineering and Technology, CGC University, Mohali – 140307, Mohali, Punjab, India, premlata.j3449@cgc.ac.in

²Department of Computer Science and Engineering Central University of Rajasthan, harimohandixit19@gmail.com

³Department of Mechanical Engineering, Chandigarh University, Mohali-140413, Punjab, India. ravinder.e1224@cumail.in

⁴Department of Modern Languages, College of Liberal Arts, University of Mississippi, Oxford, 38677, United States of America, Aalsuway@go.olemiss.edu

⁵Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, Amjad.alsuwaylimi@nbu.edu.sa

⁶Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, lafe.alenezi@nbu.edu.sa

⁷Department of Computer Science, Faculty of Information Technology, Zarqa University, Zarqa, 13110, Jordan, mjaidi@zu.edu.jo

Abstract.

The news distribution and consumption in a digital age has changed radically with the extensive resources of online news. Nevertheless, this richness has also contributed to the spread of fake news at an accelerated pace undermining the levels of trust in the population, harmony, and Sustainable Development Goals of peace, justice, and strong institutions. The proposed model is learned on the dataset of news article annotations obtained by reputable sources and confirmed fact-checkers. The experimental validation has confirmed that the model comes out with a high level of accuracy of 97.90% to differentiate genuine news and misinformation.

Keywords. Artificial Intelligence, Deep Learning, Model Training, Fake News Detection, Long Short-Term Memory (LSTM).

1. INTRODUCTION

The purpose of this research is to come up with a well-built and smart system that can automatically confirm or reject the authenticity of news based on the features of the text in news. With the proposed Long Short-Term Memory (LSTM)-based model targeting an accuracy rate of 97.90, there is an extensive testing and validation which is carried out on a dataset with the labelled samples of both authentic and fake news obtained through reputable news sources and verified sources of fact-checking.

2. LITERATURE

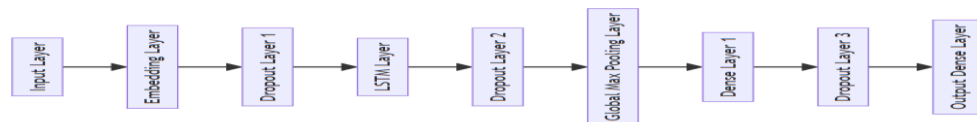
All these phases make the LSTM successful in detecting misinformation since it learns dependencies and contextual features of textual data and hence produces stable prediction results [1]. Nevertheless, some of the previous works on fake news detection had weaknesses related to low accuracy, low quality preprocessing, and poor choice of structure. Other issues like selection of irrelevant features, poor values of the model parameters, and lack of normalization of data were among the factors that led to poor model performance [2]. The proposed study seeks to overcome these shortcomings by proposing an optimized deep learning model that improves the detection of fake news. It makes use of the best methods of hyperparameter tuning, such as Grid Search and Random Search, to optimize the hyperparameters of the LSTM to give them a better prediction [3]. Moreover, this study applies machine learning into detecting false information using different methods that can respond to the dynamic and numerous characteristics of fake news [4].

3. INPUT DATASET

Above all on the dataset used in the given research, an LSTM-based model was trained that classifies news articles as either a legitimate or a fake story with a staggering precision of 97.90%. The data set is made up of a mixed collection of news contents of different fields and the style of reporting through different credible sources and databases that fact-check the content.

4. FINE-TUNED LSTM MODEL

Starting with a sequence as represented by an input layer, the much more tuned LSTM model architecture as in Flowchart 1 proceeds through an Embedding layer that converts integer-encoded vocabulary indexes to dense 100-size vectors. One of them gets overfitting through a dropout 50 percent rate layer. Following another dropout layer, the model is then basically a bi-directional LSTM layer with 150 units maintaining the time aspects and sequence processing.



Flowchart 1. Layer-wise depiction of the Fine – Tuned LSTM

5. PROPOSED METHODOLOGY

As shown in Fig. 1, this study designs and tests an LSTM-based model to detect fake news journals in real-time using a four-phase methodology that is designed in a structure. All stages are crucial in the reliability of the model, its efficiency and its robustness. The first step, Dataset Collection, refers to putting together a complete set of data based on the valid news archives and fact-checking sources, as well as reliable archives.

6. RESULTS

The analysis of the research shows that the accuracy of the fake news identification is high at 97.90 when Long Short-Term Memory (LSTM) architecture is used. LSTM performed more successfully than others particularly using GloVe embeddings and intense hyperparameter optimization methods such as grid search.

A. Analysis of Classification Report.

Table 1 shows the results of the Long Short-Term Memory (LSTM) algorithm in the classification of authentic and inauthentic news articles. Strong balance between the model and the precision and recall (0.94) also showed that the model got 94 percent of the real news items correct.

Table 1. Classification Report Analysis

	Precision	Recall	F1-Score	Support
True News	0.95	0.94	0.94	5000
Fake News	0.93	0.94	0.93	4500
Accuracy				97.90%
Macro Avg.	0.94	0.94	0.94	9500
Weighted Avg.	0.94	0.94	0.94	9500

B. Training and Validation Loss Analysis

To identify the difference between genuine and counterfeit news, Figure 2 shows that the Long Short-Term Memory (LSTM) model compares the training and validation losses with the settings of real news. At the beginning, a large loss in the epoch 0 is anticipated as a result of randomization in the learning, however, the decrease in training loss at the start suggests successful learning.

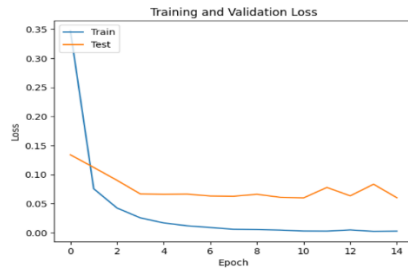


Fig. 2 Training and Validation Loss Analysis

C. Training and Validation Accuracy Analysis

The performance of the Long Short-Term Memory (LSTM) model in classifying the real and fake news depending on the training and validation accuracy is shown in Figure 3. The training accuracy increases drastically in the initial epoch to almost 1.0 in the

third epoch indicating that the model has adequately learned the training samples.

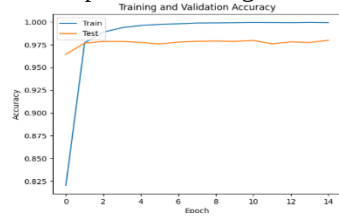


Fig. 3 Training and Validation Accuracy Analysis

D. Confusion Matrix Analysis

A distinction of actual and fake news through the Long Short-Term Memory (LSTM) model is displayed in a confusion matrix in figure 4. The table is a representation of the predicted and actual classifications as the model is able to recognize 4,145 real news articles and 3,428 fake news articles.

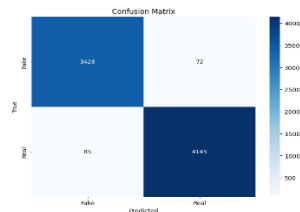


Fig. 4 Confusion Matrix Analysis

7. CONCLUSION

The dissemination of information which is accessible on the internet avenues of the digital era has changed the manner the knowledge is shared and taken in the shape of news. Nevertheless, this information overload has also contributed to spreading fake news and false information therefore destroying the opinion of the people, social trust and democracy. Creative natural language processing, and machine learning, in this case, Long Short Term Memory networks, in particular, have shown that it is almost possible to automate the process of identifying and classifying fake news. It is on the basis of this work that the LSTM model was developed that shows astonishing results with an accuracy of 97.90 to tell the truth and the fake news.

REFERENCES

- [1] Liu, S., 2023. Social Media Fake Information Identification Method Based on LSTM. Highlights in Business, Economics and Management, 21, pp.703-709.
- [2] Rajalaxmi, R.R., Narasimha Prasad, L.V., Janakiramaiah, B., Pavankumar, C.S., Neelima, N. and Sathishkumar, V.E., 2022. Optimizing hyperparameters and performance analysis of LSTM model in detecting fake news on social media. Transactions on Asian and Low-Resource Language Information Processing.
- [3] Pimpalkar, A., Singh, M., Sheikh, S., Gedam, K. and Khadgi, A., 2021. Fake news classification using bi-directional LSTM-recurrent neural network. Journal of Huazhong University of Science and Technology ISSN, 1671, p.4512.
- [4] Khaleel, Y.L., 2021. Fake News Detection Using Deep Learning. University of Miskolc.

Knee Osteoporosis Diagnosis at an Early Stage Using a Fine-Tuned VGG19 Model on X-rays

Ravinder Tonk¹, Shikha Tuteja², Mohammad Aljaidi³, Amjad A. Alsuwaylimi⁴, RAKAN M. ALANAZI⁵, Fahad Saad Z Alanazi⁶, Manish Kumar Singla⁷

¹Department of Mechanical Engineering, Chandigarh University, Mohali-140413, Punjab, India., ravinder.e1224@cumail.in

²Chitkara University Institute of Engineering and Technology, Chitkara University Punjab, India, Shikha.1290@chitkara.edu.in

³Department of Computer Science, Faculty of Information Technology, Zarqa University, Zarqa, 13110, Jordan, mjaidei@zu.edu.jo

⁴Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, Amjad.alsuwaylimi@nbu.edu.sa

⁵Family medicine resident, Saudi board of family medicine, Northern Border Health Cluster, Arar, Saudi Arabia, rakan.alsuwaylimi@gmail.com

⁶Family and community medicine, Northern Border University, Arar, 91431, Saudi Arabia, Fahad.alanazy@nbu.edu.sa

⁷Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India-602 105, msingla0509@gmail.com

Abstract.

Knee Osteoporosis (KO) is a widespread degenerative disorder that erosion of strength of bone and fosters the possibility of fractures resulting in pain and impaired movement movements. Effective treatment is possible only in early stages of detection but there are cases when the traditional means of diagnosis, physical examination, and bone mineral density testing, do not detect abnormalities of the first stage. This research paper includes a trained VGG19 deep learning model that is tailored to distinguish between healthy and osteoporotic knee X-ray images. Transfer learning was deployed to use binary classification on the already trained VGG19 architecture which is popular with image classification. This model was trained and tested using a dataset of 1,258 training images, 315 validation images and 472 testing images. It showed excellent performance with a precision of 97.5 and high recall as well as F1- scores of 0.97 in each class. The confusion matrix showed few cases of misclassifications and false negatives were 12 and there were no false positives which shows that the model had good generalization. In general, this study demonstrates that deep learning, especially VGG19, is a suitable method to automatize KO and enhance the diagnostic accuracy. It is also possible to use more advanced regularization and hyperparameter optimization to further minimize overfitting and enhance the robustness of the models in clinical settings, as will be the focus of future work

Keywords. Artificial Intelligence, Deep Learning, Knee Osteoporosis, Model Training, VGG19 Model.

1. INTRODUCTION

Knee Osteoporosis (KO) is a degenerative disease which weakens the historical stability of bones and cartilage surrounding the knee joint and results to pain, dysfunction, and increase chances of fracture. The paper presents a research on the application of VGG19 on a binary classification task where images of knee X-rays with osteoporosis and without it are compared. The model can offer a possible diagnostic aid due to the ability to learn intricate patterns in radiological information that allow physicians to make the right decision at the right time. Earlier studies have made a correlation between osteoarthritis and osteoporosis of the knee joint [1–3] and transient knee osteoporosis and the results of post-arthroplasty have provided the difficulties in diagnosis [4-5].

2. INPUT DATASET

The dataset used in this study was sourced from an open-access Kaggle repository designed for knee osteoporosis detection research. It contains two labeled categories of knee X-ray images—Healthy and Osteoporosis—divided into three subsets for training, validation, and testing. The training set includes 1,258 images, enabling the model to learn distinguishing features; the validation set comprises 315 images for tuning hyperparameters and assessing generalization; and the testing set includes 472 images to evaluate performance on unseen data. This publicly available dataset ensures transparency, reproducibility, and encourages further research advancement in automated knee osteoporosis detection, as illustrated in Fig. 1.

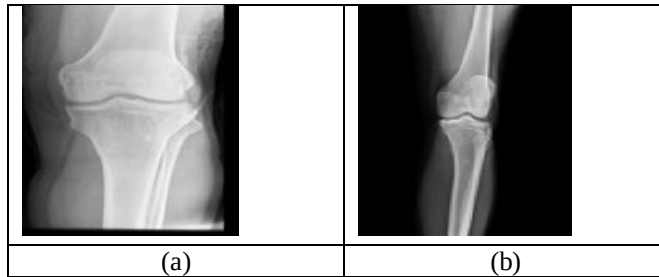


Fig.1. Dataset image for (a) Healthy, (b) Osteoporosis.

3. PROPOSED METHODOLOGY

The paper is based on the two-step research process that includes data preparation, data collection, model creation, and analysis. The first phase involves a labelled sample of knee X-ray images, classified based on being osteoporotic and non-osteoporotic, acquired in medical archives and hospital partnerships and checked by professional radiologists. The dataset is diverse both in terms of demographic categories and various levels of severity of knee osteoporosis. The second phase involves an image size reduction to the VGG19 image size and data augmentation is adopted through rotation and flipping which improves the diversity of the images.

4. RESULTS

The model on the test set had a very high precision of 97.5 in prediction of knee osteoporosis. The report had a high accuracy, recall and F1-score (0.97) in the classification of Healthy

and Osteoporosis classes in the classification report. Table of confusion revealed that the number of misclassifications was insignificant with 12 false negatives and zero false positives.

A. Analysis of Classification Report.

The knee osteoporosis detection model shows a very good performance with both classes, Healthy and Osteoporosis. Both classes show balanced performance and dependable performance as the F1-score is 0.97. The model has a high classification rate, as well as an overall accuracy of 97.5. Also, it has been shown that the macro and weighted averages were precise (0.98) and had a recall of 0.97, which confirms the consistency and stability of the results in both categories, as shown in Fig. 3.

	precision	recall	f1-score	support
Healthy	0.95	1.00	0.97	231
Osteoporosis	1.00	0.95	0.97	241
accuracy			0.97	472
macro avg	0.98	0.98	0.97	472
weighted avg	0.98	0.97	0.97	472

Accuracy of the Model: 97.5%

Fig.3. Classification Report Analysis.

A. Confusion Matrix Analysis

The knee osteoporosis detection model confusion matrix has excellent classification results. The model correctly discriminated the 231 and 229 cases of Healthy and Osteoporosis respectively and had high discrimination. This small mistake notwithstanding, the model showed a recall of 100% on the Healthy category and a high level of precision in general. The findings imply that there is minimal misclassification and high reliability that support the practicality of the model in the consistent and accurate clinical diagnosis as demonstrated in Fig. 4.

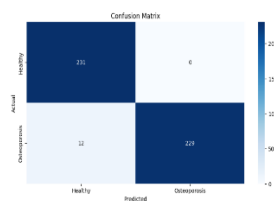
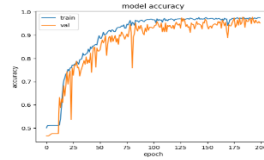


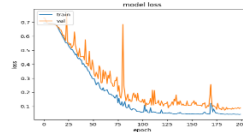
Fig.4. Confusion Matrix Analysis

B. Training and Validation Accuracy and Loss Analysis

The accuracy and loss curves in training and validation in 200 epochs of the model give a holistic perspective of the way the model learns. Still, in general, the findings confirm the high strength and reliability of the model to properly determine the presence of the knee osteoporosis as it is demonstrated in Fig. 4(a) and (b).



(a)



(b)

Fig.4.(a) Training and Validation Accuracy Analysis, (b) Training and validation Loss Analysis

5. CONCLUSION

The current paper is an X-ray image-based method to predict knee osteoporosis on the basis of a highly-tuned VGG19 deep learning model. The model shows an outstanding skill in the possibility to identify a healthy and osteoporotic knee condition, based on a strong score of 97.5 on the test sample. Its strength is also confirmed by the classification report as both of the classes have high precision, recall, and F1-scores of 0.97 indicating a balanced performance of the classes. The confusion matrix demonstrates the reliability of the model with the results of the existence of only 12 false negatives and no false positives.

REFERENCES

- [1] Karvonen, R.L., Miller, P.R., Nelson, D.A., Granda, J.L. and Fernandez-Madrid, F., 1998. Periarticular osteoporosis in osteoarthritis of the knee. *The Journal of rheumatology*, 25(11), pp.2187-2194.
- [2] Ramesh Kumar, Ratneshwar Kumar Ratnesh, Jay Singh, Ashok Kumar and Ramesh Chandra, IoT-Driven Experimental Framework for Advancing Electrical Impedance Tomography, *ECS Journal of Solid State Science and Technology*, Volume 13, Number 2, 13 027002, 7 February 2024
- [3] Yoshimura, N., Muraki, S., Oka, H., Mabuchi, A., En-Yo, Y., Yoshida, M., Saika, A., Yoshida, H., Suzuki, T., Yamamoto, S. and Ishibashi, H., 2009. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. *Journal of bone and mineral metabolism*, 27, pp.620-628.
- [4] Parker, R.K., Ross, G.J. and Urso, J.A., 1997. Transient osteoporosis of the knee. *Skeletal radiology*, 26, pp.306-309.
- [5] R. Kumar, Vevekanandam, and R. Dey, "IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment," *International Journal of Vegetable Science*, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.

Tire Condition Classification Using EfficientNetB0 for Improved Road Safety

Harpreet Singh¹, Shikha Tuteja², Mohammad Aljaidi³, Amjad A. Alsuwaylimi⁴, Lafi A. Alenezi⁵, Zaid S. Alanazi⁶, Manish Kumar Singla⁷

¹Department of Mechanical Engineering, Chandigarh University, Mohali-140413, Punjab, India., er.harpreets@gmail.com

²Chitkara University Institute of Engineering and Technology, Chitkara University Punjab, India, Shikha.1290@chitkara.edu.in

³Department of Computer Science, Faculty of Information Technology, Zarqa University, Zarqa, 13110, Jordan, mjaidi@zu.edu.jo

⁴Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, Amjad.alsuwaylimi@nbu.edu.sa

⁵Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, lafi.alenezi@nbu.edu.sa

⁶Information Technology Management, Northern Border University, Arar, 91431, Saudi Arabia, Zaid.alanazi@nbu.edu.sa

⁷Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India-602 10, msingla0509@gmail.com

Abstract.

The quality of tires is also significant in the definition of safety, performance, and efficiency of any vehicle, where any defects in tires become a severe threat to the safety and trustworthiness of roads. This should be done by early identification of such defects in order to avert accidents, to save on maintenance costs as well as vehicle downtimes. This paper introduces an automated tire quality assessment model based on the EfficientNetB0 model which is a Convolutional Neural Network (CNN) with an outstanding accuracy and a high computation efficiency. Through visual inspection, the model categorizes images of digital tires in two classes, which include defective and good state. The model with 1,854 tire images obtained through Kaggle was trained, validated, and tested with a high classification accuracy of 96.72%.

Keywords. Artificial Intelligence, Tyre quality, EfficientNetB0,

1. INTRODUCTION

The quality of tires is important in ensuring safety of the vehicles, safety, efficiency, and durability. It has direct implications on the handling, braking, fuel consumption of a car and the entire driving experience. With age, tires (especially when they are in commercial and high-performance cars) wear and tear thus it is important to check and maintain the tires frequently. Low-quality tires (because of manufacturing errors, or environmental factors)

can cause unsafe driving performance like the blowout or loss of control. The early identification of defects is crucial to avoid accidents, decrease down time, lower the maintenance costs and increase the tires life.

2. LITERATURE

The quality of tires is central to the safety and the efficiency of the functioning of the vehicles as it is the primary concern which makes the difference between the regular maintenance and the prevention of the car accidents as the timely revealed tire defect can significantly enhance the circumstances. Deep Tread proposed by Degadwala et al. [1] is a pre-trained deep learning model that can be used to classify defective and non-defective tires having a high accuracy at a low cost of computation. Equally, Rajeswari et al. [2] put forward a weighting CNN which puts more significance to various fault kinds during training which improves classification and issues of class imbalance. Singh et al. [3] focused on the significance of the joint use of several DL models to enhance accuracy of fault detection, whereas the role of CNNs in the context of the real-time quality control.

3. INPUT DATASET

The data of this research was obtained on the open source Kaggle site with 1,854 digital tire images in two categories, that is, good and defective. All pictures in the form of a JPEG or PNG are taken of a single tire in the circumstances of the real world by using high-quality digital cameras as shown in Fig. 1.

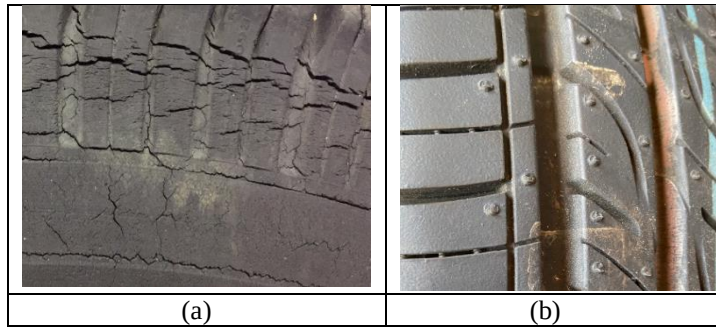


Fig. 1 Dataset image for (a) Defective (b) Good condition image type

4. FINE-TUNED EFFICIENTNETB0 ARCHITECTURE

In this work, the convolutional neural network (CNN) EfficientNetB0 that is highly efficient is used because it has been noted to have a high degree of performance in terms of image classification activities. Trained on big data, like ImageNet, EfficientNetB0 is a good extractor of various image features. The model was further modified to evaluate tire quality, where bottom layers were substituted with tailored fully connected layers which are used to classify into good and flawed tires; however, the later layers were still used to extract low-level visual representations such as edges and textures.

5. PROPOSED METHODOLOGY

The study uses the experimental design of the four stages to scientifically and systematically categorize tire quality with the maximum accuracy, efficiency, and reliability of each step

with the help of the EfficientNetB0 paradigm. The initial stage is known as Dataset Collection which entails retrieving a rich dataset of 1,854 digital tire images that had been collected and categorized under the open-source Kaggle platform into two distinct classes namely good condition and defect.

6. RESULTS

The four phases of research will be scheduled based on ensure the scientific and systematic approach towards the grouping of the quality of tyres according to EfficientNetB0 model. All these steps are required in the facilitation of adequate classification system as far data is collected in the way of modelling analysis.

A. Classification Report Analysis.

Classification report entails an elaborate analysis of the performance of the model in differentiating between good and defective tires. The model had a good reliability of 96.72 indicating a good accuracy in the classification of tires as shown in Table 1.

Table 1. Classification Report Analysis

	precision	recall	f1-score	support
defective	0.9896	0.9500	0.9694	100
good	0.9425	0.9880	0.9647	83
accuracy			0.9672	183
macro avg	0.9661	0.9690	0.9670	183
weighted avg	0.9682	0.9672	0.9673	183

B. Training and Validation loss Analysis

Figure 2 shows the training and validation loss with 22 epochs. It can be seen that the red line exhibits a slow decrease in training loss, which is a sign of successful learning of the data, whereas the green line is a chart of validation loss that also decreases at first and starts to stabilize with increase in epochs.



Fig. 2 Training and Validation loss Analysis

C. Training and Validation Accuracy Analysis

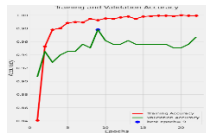


Fig. 3 Training and Validation Accuracy Analysis

Figure 3 presents the training and validation accuracy of 22 epochs. The red line that signifies the accuracy of training increases fast and it is close to 100 percent which is an indication that the model fits the training data rather well.

D. Confusion Matrix Analysis

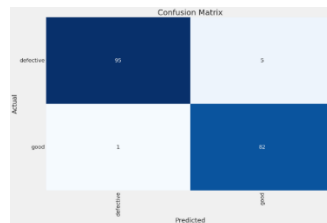


Fig. 4 Confusion Matrix Analysis

The confusion matrix below (Fig. 4) provides the full picture about the performance of the model as regards classification is concerned. In both scenarios, the algorithm results in finding 82 good tyres (true negatives) and 95 problematic tyres (true positives) and therefore is a good performance.

7. CONCLUSION

The given paper introduces a tire quality classification system based on advanced EfficientNetB0 convolutional neural network (CNN) model, which was performed automatically. The proposed method achieves the classification of tires based on two classes of defective tires and good condition varieties with a dataset containing 1,854 tire images (digital). The system was proved effective and strong because through fine-tuning the existing pre-trained EfficientNetB0 model, the system reached impressive overall accuracy of 96.72% that indicated its capability and strength in tire quality inspection. Automating this process drastically lowers the reliance of this method of human inspection, which limits the human error, labor intensity and time of inspection. It, therefore, increases road safety and reliability of the vehicles by providing an opportunity to identify faulty tires early.

REFERENCES

- [1] Degadwala, S., Upadhyay, R., Upadhyay, S., Soni, M., Parikh, D.J. and Vyas, D., 2023, November. DeepTread: Exploring Transfer Learning in Tyre Quality Classification. In 2023 International Conference on Sustainable Communication Networks and Application (ICSCNA) (pp. 1448-1453). IEEE.
- [2] R. Kumar, M. Aljaidi, M. K. Singla, A. Gupta, A. M. Alhomoud, A. A. Alsuwaylimi, and S. M. Alenezi, "Development of a prototype global positioning system based stick for blind patients," International Journal of Online & Biomedical Engineering, vol. 20, no. 8, p. 21, 2024, doi: 10.3991/ijoe.v20i08.49343.
- [3] Singh, P., Hasija, T. and Ramkumar, K., 2024, November. Enhancing Tyre Quality Assessment with Cutting-Edge Classification Algorithms. In 2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1253-1259). IEEE.

Transfer Learning-Based CNN Models for Vehicle Image Classification

Kiranpreet Singh¹, Shikha Tuteja², Mohammad Aljaidi³, Amjad A. Alsuwaylimi⁴, Lafi A. Alenezi⁵, Zaid S. Alanazi⁶, Manish Kumar Singla⁷

¹Department of Mechanical Engineering, Chandigarh University, Mohali-140413, Punjab, India., Kiranpreet.e2014@cumail.in

² Chitkara University Institute of Engineering and Technology, Chitkara University Punjab, India, Shikha.1290@chitkara.edu.in

³Department of Computer Science, Faculty of Information Technology, Zarqa University, Zarqa, 13110, Jordan, mjaidi@zu.edu.jo

⁴Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, Amjad.alsuwaylimi@nbu.edu.sa

⁵Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, laifi.alenezi@nbu.edu.sa

⁶Information Technology Management, Northern Border University, Arar, 91431, Saudi Arabia, Zaid.alanazi@nbu.edu.sa

⁷Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India-602 105, msingla0509@gmail.com

Abstract.

This paper describes a Convolutional Neural Network (CNN)-based platform that is created to identify vehicles and put them into seven different categories and they include: Auto Rickshaws, Bikes, Cars, Motorcycles, Planes, Ships, and Trains. As the influx of visual data processed by cameras and sensors has grown exponentially, quick and precise classification of vehicles has been demanded and is now important in the contemporary application of camera-related surveillance of traffic, autonomous driving, and intelligent transportation management. The data set involved in this piece is composed of 5,007 JPG images 3,504 of them to train and 1,503 to test.

Keywords. Artificial Intelligence, Deep Learning, Vehicle classification, Sustainable mobility.

1. INTRODUCTION

The classification of vehicles is important in diverse fields including the monitoring of the traffic, autonomous vehicles, and smart transportation systems. As the digital platform develops fast and vehicle imagery proliferates, the necessity to find quick and accurate means of classification to the problem has grown. Maungmai and Nuthong [1] have defined the importance of deep learning in car classification and how the traditional fashion may be unable to cope with a large and tricky data set. In the same manner, Vijayaraghavan and Laavanya [2] showed that the Convolutional Neural Networks (CNNs) are better in terms of precision and performance as compared to conventional methods.

2. LITERATURE

Convolutional neural networks (CNNs) and deep learning methods have been extensively researched in the context of vehicle classification; they demonstrate high feature extractors and pattern recognition characteristics. Maungmai and Nuthong [1] created a CNN centered model that has the capacity to classify vehicles by their type and color and the model showed strong real world performance in four types of vehicles and seven colors. Vijayaraghavan and Laavanya [2] concentrated on bus detecting, cars, and bikes detection where capturing of the buses, cars, and bikes occurred by detection of bounding boxes followed by probability predictions to improve on detecting the objects. Tas et al. [3] introduced a CNN method that is specifically trained on low-resolution surveillance photographs, and compared it with VGG16 with focus on the scenario of efficiency versus accuracy.

II. INPUT DATASET

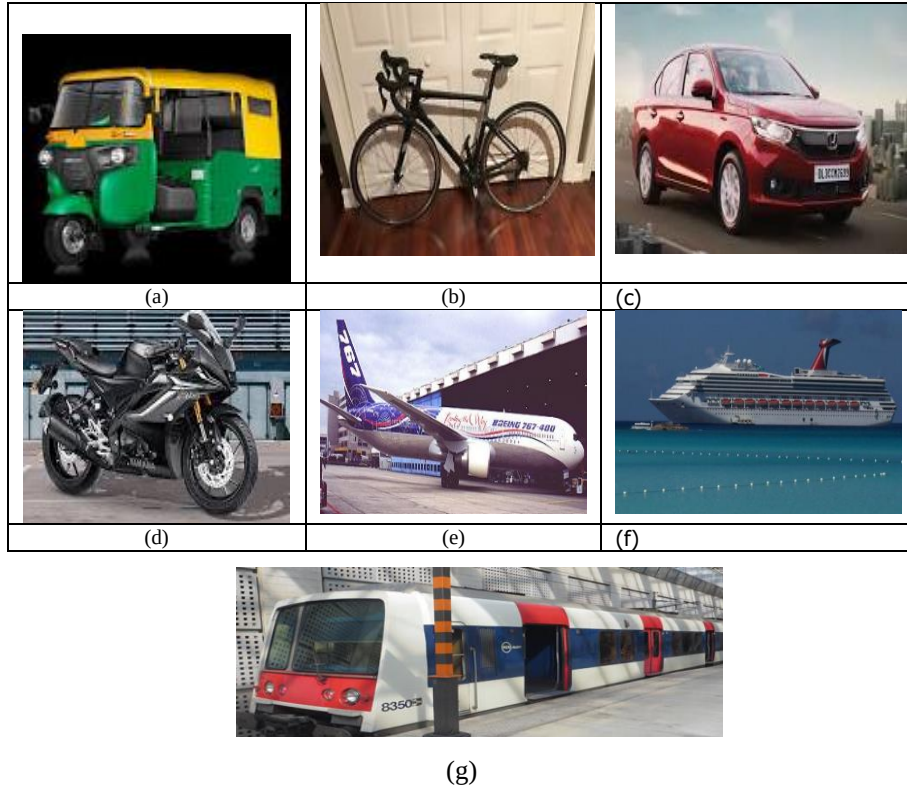


Fig. 1 Images from the proposed dataset.

The database that has been used in this study consists of 5,007 JPG images of vehicles that are grouped into seven different classes namely: Auto Rickshaws, Bikes, Cars, Motorcycles, Planes, Ships, and Trains. All of the pictures provide a visual depiction of each particular type of vehicle, taken at a variety of angles and resolutions to recreate real-life situations. The data is split into 3,504 images to be trained and 1,503 to be tested and the ratio of classes is kept so that there are no biases to evaluate the model. A variety of opinions will make the dataset more complex and realistic; thus making sure that the CNN model can effectively generalize with unknown data. The examples of such vehicle images are given in figure 1.

3. METHODOLOGY

A. Proposed Model:

The suggested Convolutional Neural Network (CNN) vehicle recognition model has combined various convolutional, pooling and dropout layers in order to attain efficient multi-class recognition. It has a Conv2D layer (conv2d_15) which filters feature maps with 32 filters with sizes 32-by-32-by-3 resulting to (126, 126, 32) feature maps. The next layer is the MaxPooling2D, which down-samples the result to (63, 63, 32), space dimension minimization but without the loss of important data.

B. Proposed Methodology:

The research also had a systematic methodology with the collection of data starting at the stage of data collection when 5007 JPG pictures of vehicles were collected and were divided into seven groups such as Auto Rickshaws, Bikes, Cars, Motorcycles, Planes, Ships and Trains. This dataset was also selected in a way that it is diverse and has a balanced address of all the types of vehicles, which is essential in the domain of creating a strong classification model.

IV. RESULTS

Conventional Neural Network (CNN) model showed good performance in vehicle classification in a dataset of 5007 images in seven categories namely Auto Rickshaws, Bikes, Cars, Motorcycles, Planes, Ships, and Trains. With a total test accuracy of 84, the model yielded extremely high results with Bikes (precision 0.95, recall 0.94) and Ships (F1 -score 0.94). Nevertheless, Trains had the least precision (0.68) and recall (0.76) which means that it needs additional improvement.

A. Classification Report Analysis

The model has an overall accuracy of 84, a macro and weighted average of 0.83, which has indicators of balanced but amenable classification abilities. As shown in Fig.2.

	precision	recall	f1-score	support
0	0.85	0.80	0.82	234
1	0.95	0.94	0.94	238
2	0.85	0.76	0.80	233
3	0.86	0.86	0.86	214
4	0.76	0.88	0.82	220
5	0.86	0.88	0.87	252
6	0.68	0.70	0.69	112
accuracy			0.84	1503
macro avg	0.83	0.83	0.83	1503
weighted avg	0.84	0.84	0.84	1503

Fig. 2 Classification Report Analysis

B. Confusion Matrix Analysis

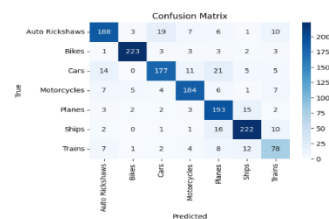


Fig. 3 Confusion Matrix Analysis

Although the Motorcycle and Plane detection achieved good outcomes, misclassification between other related-looking vehicles implies that the model should be improved with more distinction of features and the variety of data to become more accurate. As shown in fig.3.

C. Training VS Validation Accuracy and Loss Analysis

The model had a high-training accuracy-92.05% and validation-83.70% accuracy with low differences in loss at the final epoch. The results show that the model is robust and efficient during training with low overfitting, high levels of generalization, and stability of learning. As shown in fig.4(a) and fig.4(b).

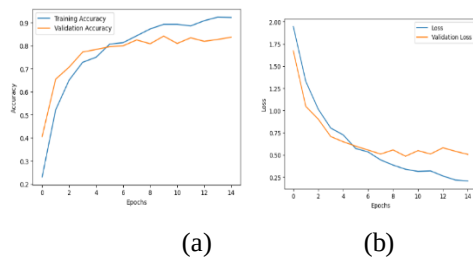


Fig. 4 (a) Training and Validation Accuracy Curve, (b) Training and Validation Loss Curve.

4. CONCLUSION

The categories of vehicle classified as a Convolutional Neural Network (CNN) model in this work are seven namely Auto Rickshaws, Bikes, Cars, Motorcycles, Planes, Ships and Trains. Through the tedious process of data collection and preparation, 5007 photos were selected and divided into training and test data which resulted in The model being trained and tested on the data and thus yielding an overall accuracy of 84% with regards to the test set. The test outcomes showed the ability of the CNN design to efficiently learn intricate characteristics of various types of vehicles, and thus emphasize on its likely application in traffic control and self-driving vehicle systems.

REFERENCES

- [1] Maungmai, W. and Nuthong, C., 2019, February. Vehicle classification with deep learning. In 2019 IEEE 4th international conference on computer and communication systems (ICCCS) (pp. 294-298). IEEE.
- [2] R. Kumar, Vevekanandam, and R. Dey, "IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment," International Journal of Vegetable Science, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.
- [3] Tas, S., Sari, O., Dalveren, Y., Pazar, S., Kara, A. and Derawi, M., 2022. Deep learning-based vehicle classification for low quality images. Sensors, 22(13), p.4740.

Deep Convolutional Feature Scaling with EfficientNetB3 for Robust Early Detection of Smoking Patterns in Health Informatics

Ravinder Tonk¹, Shikha Tuteja², Mohammad Aljaidi³, Amjad A. Alsuwaylimi⁴, Faisal M. Alanazi⁵, Nael Al Ali⁶, Manish Kumar Singla⁷

¹Department of Mechanical Engineering, Chandigarh University, Mohali-140413, Punjab, India., ravinder.e1224@cumail.in

²Chitkara University Institute of Engineering and Technology, Chitkara University Punjab, India, Shikha.1290@chitkara.edu.in

³Department of Computer Science, Faculty of Information Technology, Zarqa University, Zarqa, 13110, Jordan, mjaidi@zu.edu.jo

⁴Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, Amjad.alsuwaylimi@nbu.edu.sa

⁵Department of Pharmacy, Narcotics & Compounding unit, Alkhobar, King Fahad Hospital of the university, fmdalanazi@iau.edu.sa

⁶Department of Pharmacy, Narcotics & Compounding unit, Alkhobar, King Fahad Hospital of the university, Nhaali@iau.edu.sa

⁷Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India-602 105, msingla0509@gmail.com

Abstract.

With the underlying that smokers and non smokers should be given a special treatment and treatment planning, smoking is the main cause of these preventable conditions like cancer, respiratory diseases and cardiovascular diseases. It is important to detect the smoking habits at an early age to mitigate any health-related complications caused by smoking, and to enhance the overall health of the population. It uses the EfficientNetB3 network to optimize the performance of smoking detection by the state-of-the-art convolutional neural network with a balance between the model complexity and performance.

Keywords. Smoking Detection, EfficientNetB3 Model, Convolutional Neural Networks (CNN),

1. INTRODUCTION

Smoking also contributed substantially to the etiology of cancer, respiratory and cardiovascular diseases, and ranks among the top causes of preventable diseases. Quite often medical examinations presuppose that the distinction between smokers and non-smokers should be made, as this distinction will help develop particular treatment regimens and personalized care. There is a need to identify the smoking behavior among the younger age

to lower the incidence of development of smoking diseases and have a positive impact on the health status of the population [1-3].

2. INPUT DATASET

To conduct this analysis, the data was gathered in the open-source Kaggle platform, which provides a diversity of sufficient images, as shown in Fig. 1. The overall data set has 3,254 confirmed image files, separated into smoking and non-smoking. They are divided into three subsets: 320 images in an extended category, 326 images in a secondary category, and 2,608 images in the main category.



Fig. 1 Dataset image for (a) No - Smoking (b) Smoking image type

3. FINE-TUNED EFFICIENTNETB3 MODEL

The present work uses the EfficientNetB3 model to enhance the accuracy of the prediction of the smoking and non-smoking condition. Large-scale image weights are pre-trained into the model, which means that it can extract features. This base architecture makes the 1536-dimensional feature vectors, on which we rely our detection of characteristics in an image that are of interest as shown in Fig. 2.

Layer (type)	Output Shape	Param #
EfficientNetB3 (Functional)	(None, 1536)	10,783,535
BatchNormalization	(None, 1536)	6,144
Dense (Dense)	(None, 256)	393,472
Dropout (Dropout)	(None, 256)	0
Dense_1 (Dense)	(None, 2)	514
Total Parameters	-	11,183,665
Trainable Parameters	-	11,093,290
Non-Trainable Parameters	-	90,375

Fig. 2. Layer-wise Depiction of the proposed model

4. PROPOSED METHODOLOGY

The methodology for this study is structured into four distinct phases, each playing a crucial role in the successful prediction of smoking and non-smoking statuses using the EfficientNetB3 model.

Phase 1: Dataset Collection: The initial step is the compilation of a massive number of smoking and non-smoking medical pictures. It can be made diverse with clinical repositories and open-access tools.

Phase 2: Model Development: In this step a model (EfficientNetB3) is chosen and trained to accomplish the smoking detection jobs. The model has a potent feature extraction system that is tuned to address medical image difference.

Phase 3: Model Training and Validation: The model is also trained through an iterative optimization of the prepared dataset, which minimizes classification errors and enhances accuracy.

Phase 4: Model Evaluation and Testing: The last step tests the model using unknown test data. The accuracy, precision, recall and F1-score metrics are derived to measure the ability to predict. Ground truth labels are used to compare predictions and to determine reliability.

5. RESULTS

The EfficientNetB3 model classified the 3,254 smoking and non-smoking images with high accuracy, precision, recall and F1-score. Its stable operation focuses on its practicality in smoking detection, health studies, and early intervention uses.

A. Classification Report Analysis

The efficiency of the EfficientNetB3 model in differentiating between smoking and non-smoking images with overall accuracy of 90 indicates that the model has a high capacity to classify between smoking and non-smoking images (Fig. 3). The model had a precision of 0.90, a recall of 0.83 and an F1-score of 0.86 for the not-smoking class which showed a strong detection but slightly lower in recall than that of the smoking class.

Class	Precision	Recall	F1-Score	Support
not_smoking	0.90	0.83	0.86	126
smoking	0.90	0.94	0.92	194
Accuracy	-	-	0.90	320
Macro Avg	0.90	0.89	0.89	320
Weighted Avg	0.90	0.90	0.90	320

Fig. 3 Classification Report Analysis

B. Training and Validation loss Analysis

The Training and Validation Loss graph (Fig. 4) shows that the loss values decrease with increasing epochs of both training and validation sets. The training loss (red line) starts at a reasonably high level, but goes down fast, which means that the model fits the training data better.

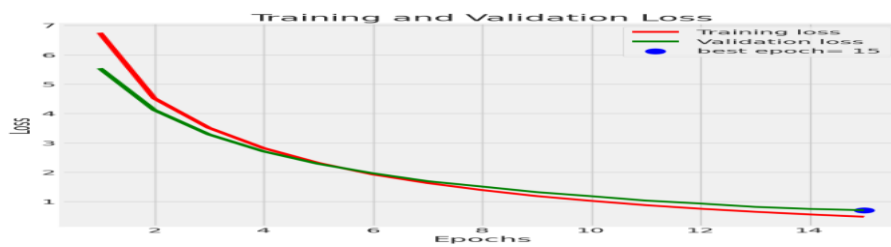


Fig. 4 Training and Validation Loss Analysis

C. Training and Validation Accuracy Analysis

The trends of the accuracy during the 15 epochs are displayed in the graph "Training and Validation Accuracy" in Fig. 5. The trend in the training accuracy, as the epochs progress,

is a steadily increasing curve, approaching 100% to signify that the model is becoming progressively more competent to predict results on the training data accurately.

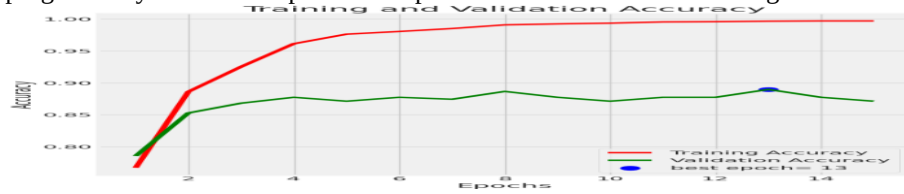


Fig. 5 Training and Validation Accuracy Analysis

D. Confusion Matrix Analysis

The confusion matrix (Fig. 6) of the EfficientNetB3 model reveals its impressive accuracy in smoking and non-smoking image classification. In the not smoking group it made 105 true positives and 21 false negatives whereas in the smoking group it made 182 true positives and just 12 false positives

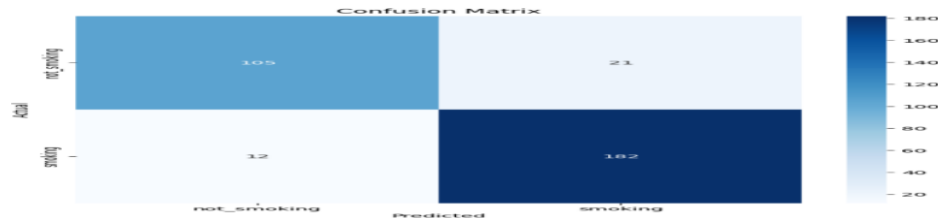


Fig. 6 Confusion Matrix Analysis

7. CONCLUSION

This paper highlights the pressing necessity to differentiate between smokers and non-smokers in order to promote individualized treatment and enhance health care planning. Since smoking is one of the major causes of preventable diseases like cancer, respiratory diseases, and cardiovascular diseases, it is important to detect smoking habits at an early stage to minimize the health risks associated with smoking and improve the health outcomes of the population.

REFERENCES

- [1] Ali, A.A., Hossain, S.M., Hovsepian, K., Rahman, M.M., Plarre, K. and Kumar, S., 2012, April. mPuff: automated detection of cigarette smoking puffs from respiration measurements. In Proceedings of the 11th international conference on Information Processing in Sensor Networks (pp. 269-280).
- [2] Senyurek, V., Imtiaz, M., Belsare, P., Tiffany, S. and Sazonov, E., 2019. Cigarette smoking detection with an inertial sensor and a smart lighter. *Sensors*, 19(3), p.570.
- [3] Imtiaz, M.H., Ramos-Garcia, R.I., Wattal, S., Tiffany, S. and Sazonov, E., 2019. Wearable sensors for monitoring of cigarette smoking in free-living: a systematic review. *Sensors*, 19(21), p.4678.

Optimized EfficientNetB3 Architecture for High-Precision Multi-Class Weather Condition Classification

Kiranpreet Singh¹, Shikha Tuteja², Mohammad Aljaidi³, Amjad A. Alsawaylimi⁴, Zaid S. Alanazi⁵, Lafi A. Alenezi⁶, Manish Kumar Singla⁷

¹Department of Mechanical Engineering, Chandigarh University, Mohali-140413, Punjab, India., Kiranpreet.e2014@cumail.in

²Chitkara University Institute of Engineering and Technology, Chitkara University Punjab, India, Shikha.1290@chitkara.edu.in

³Department of Computer Science, Faculty of Information Technology, Zarqa University, Zarqa, 13110, Jordan, mjaidi@zu.edu.jo

⁴Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, Amjad.alsawaylimi@nbu.edu.sa

⁵Information Technology Management, Northern Border University, Arar, 91431, Saudi Arabia, Zaid.alanazi@nbu.edu.sa

⁶Department of Computer Science, College of Science, Northern Border University, Arar, 91431, Saudi Arabia, afi.alenezi@nbu.edu.sa

⁷Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India-602 105, msingla0509@gmail.com

Abstract.

Most various industries, disaster management being one of them, transport and agriculture in particular, require meteorological conditions to be properly rated in the correct and appropriate manner, i.e. clouds, rain shine, and sunrise. This paper concentrates on the EfficientNetB3 model which is characterized by performance as well as computing efficiency, and hence deep learning is used to advance the classification of weather conditions. Optimizing the EfficientNetB3 design we arrive at an outraging 97% classification accuracy.

Keywords. EfficientNetB3, weather condition classification, deep learning,

1. INTRODUCTION

Weather classification Accurate and quick classification of weather conditions is essential in numerous industries, one of them being disaster management, transportation, and agriculture. Early detection and identification of weather changes like clouds, rain, shine and sunrise are meant enhance decision making through enhanced knowledge about the changes and consequently reduces the risks associated with the sudden changes in weather. The farmers ought to plant as much as they can and the officials involved in transportation ought

to enhance safety measures to prevent accidents caused by unfavorable climatic conditions [1-3].

2. INPUT DATASET

The dataset utilized in the present research presented in Fig. 1 was obtained on the open-source Kaggle site and is dedicated to the four weather conditions such as cloudy, rainy, sunny, and sunrise. It contains 1,125 high quality and diversified images, which have been carefully partitioned into training, validation and testing to facilitate the development and testing of a strong model. The training pool consists of 900 images, which gives a broad variety of situations to increase the learning ability of the EfficientNetB3 model.

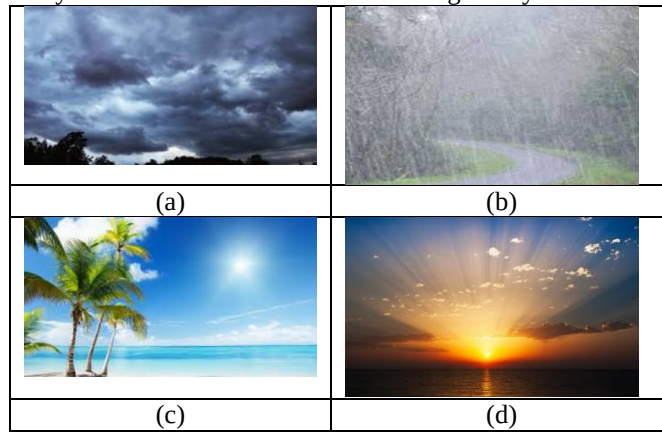


Fig. 1 Dataset image for (a) Cloudy (b) Rainy (c) Shine (d) Sunrise image type

3. FINE-TUNED EFFICIENTNETB3 MODEL

The EfficientNetB3 model demonstrated in Fig. 2 is the basis that is used to classify four weather conditions, namely, cloudy, rainy, shine, and sunrise. To bring it to this task, the architecture was made finer with introducing custom layers on top of the already trained base, which already has 11,184,179 parameters.

Layer (Type)	Output Shape	Param #
EfficientNetB3 (Functional)	(None, 1536)	10,783,535
Batch Normalization (BatchNormalization)	(None, 1536)	6,144
Dense	(None, 256)	393,472
Dropout (Dropout)	(None, 256)	0
Dense_1	(None, 4)	1,028
Total Parameters		11,184,179
Trainable Parameters		11,093,804
Non-Trainable Parameters		90,375

Fig. 2 Layer-wise depiction of the Proposed Model

4. PROPOSED METHODOLOGY

The process of weather condition classification with the help of EfficientNetB3 model is organized into four separate stages: Dataset Collection, Data Pre-processing, Model Development, and Model Evaluation.

Phase 1: Dataset Collection: At this step, an open-source platform with four types of weather such as clouds, rain, shine, and sunrise is gathered using such a system as Kaggle.

Phase 2: Data Pre-processing: The images obtained are downsized to the input dimensions of EfficientNetB3. Data augmentation where rotation, flipping, and brightness adjustments are used to enhance dataset diversity and data overfitting is avoided.

Phase 3: Model Development: This step is devoted to the development of the EfficientNetB3 model based on the transfer learning on ImageNet. This is done through fine-tuning of parameters to balance depth, width and resolution through compound scaling.

Phase 4: Model Evaluation: The last process is to test the performance of the model based on accuracy, precision, recall, and F1-score. To determine the effectiveness of classifications, confusion matrices are created to evaluate the effectiveness of classifying each category.

5. RESULTS

The adjusted EfficientNetB3 model, that had 11484179 parameters and additional dense layers, was highly accurate in the classification of four weather types. Dropping enhanced generalization, minimizing overfitting, and increasing reliability, and showed robust performance with architectural differences and fine-tuning.

A. Classification Report Analysis

As the Fig. 3 classification report indicates, EfficientNetB3 with fine-tune strategy has high degree of performance in classifying four weather conditions namely Cloudy, Rain, Shine, and Sunrise with an overall accuracy of 97 percent.

Class	Precision	Recall	F1-Score	Support
Cloudy	1.00	0.96	0.98	23
Rain	1.00	1.00	1.00	14
Shine	0.95	0.91	0.93	23
Sunrise	0.94	1.00	0.97	30
Accuracy			0.97	90
Macro Avg	0.97	0.97	0.97	90
Weighted Avg	0.97	0.97	0.97	90

Fig. 3 Classification Report Analysis

B. Training and Validation loss Analysis

The loss trends of training and validation during 25 epochs are shown in the graph in Fig. 4. First, the two curves decrease and this indicates effective learning of the training data. Nevertheless, beyond epoch 15 the loss in training keeps decreasing and validation loss levels off or increases, which is an indication of overfitting.

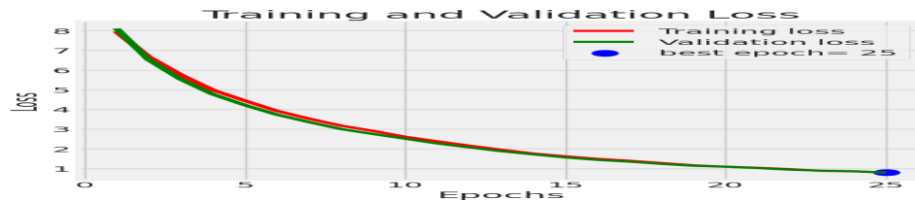


Fig. 4 Training and Validation loss Analysis

C. Training and Validation Accuracy Analysis

The accuracy trend of the model across 25 epochs is shown in the graph in Fig. 5. At the beginning, training and validation accuracy increase and exhibit effective optimization.

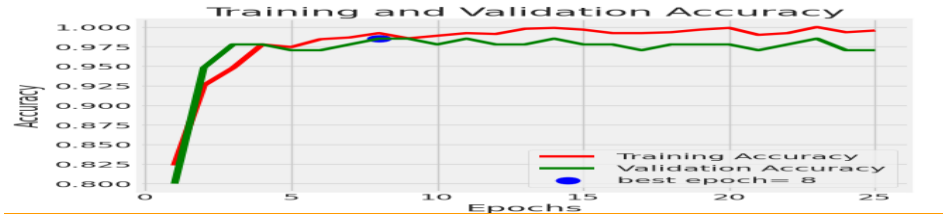


Fig. 5 Training and Validation Accuracy Analysis

D. Confusion Matrix Analysis

The confusion matrix (Fig. 6) gives a more precise analysis of the model performance in four weather types Cloudy, Rain, Shine and Sunrise. Diagonal items are correct classifications and off-diagonal items denote errors.

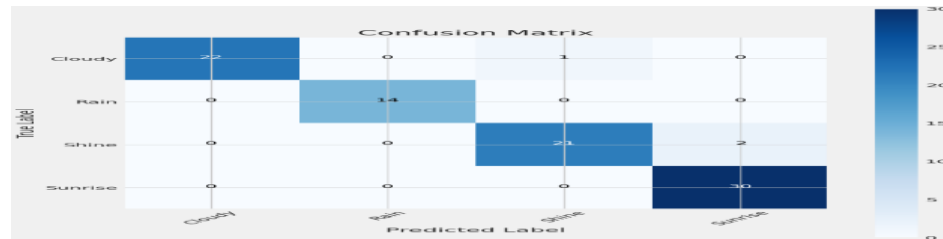


Fig. 6 Confusion Matrix Analysis

6. CONCLUSION

Proper and fast categorization of weather conditions is essential in disaster management, transport, and in agriculture. This paper has shown that the EfficientNetB3 model is useful in classifying four types of weather: cloudy, rainy, shine, and sunrise. With compound scaling used to scale depth, width and resolution, the refined model with 11,484,179 parameters had not only high computational efficiency but good classification accuracy. The model provided excellent results after conducting the test on a large dataset where the overall accuracy was 97%.

REFERENCES

- [1] Roser, M. and Moosmann, F., 2008, June. Classification of weather situations on single color images. In 2008 IEEE intelligent vehicles symposium (pp. 798-803). IEEE.
- [2] Guerra, J.C.V., Khanam, Z., Ehsan, S., Stolkin, R. and McDonald-Maier, K., 2018, August. Weather Classification: A new multi-class dataset, data augmentation approach and comprehensive evaluations of Convolutional Neural Networks. In 2018 NASA/ESA Conference on Adaptive Hardware and Systems (AHS) (pp. 305-310). IEEE.
- [3] Piotrowicz, K. and Ciaranek, D., 2020. A selection of weather type classification systems and examples of their application. Theoretical and Applied Climatology, 140, pp.719-730.

Deep Learning Approaches for Multiclass Plant Disease Detection: A Comparative Study of Custom and Pretrained Models

Safaa Zaman¹, Amal H. Alharbi², Amel Ali Alhussan³, Doaa Sami Khafaga⁴, El-Sayed M. El-kenawy^{5,6}, Marwa M. Eid⁷

¹College of life sciences , Information sciences department Kuwait University
S.3zaman@ku.edu.kw

²Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia ahalharbi@pnu.edu.sa

³Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia, aalhussan@pnu.edu.sa

⁴Department of Computer Sciences College of Computer, & Information Sciences
Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia dskhafga@pnu.edu.sa

⁵Delta Higher Institute of Engineering and Technology Department for Communications and Electronics, Mansoura 35511, Egypt

⁶Applied Science Research Center Applied Science Private University Amman, Jordan
skenawy@ieee.org

⁷Faculty of Artificial Intelligence Delta University for Science & Technology Mansoura 35111, Egypt Jadara University Research Center Jadara University, Jordan
mmm@ieee.org

Abstract.

This study evaluates the performance of three models: Here we present a Tiny VGG model trained from scratch for our targeted object classes, and another two pre-trained models: EfficientNet-B0 and Inception V3 in PyTorch. The carefully trained, customized Tiny VGG model from the onset that had even basic preprocessing of the images such as resizing and converting to tensor was evidenced to perform dismally with an accuracy of only 3.78% which was interpreted to mean that the model was too simple to deal with the variation in the dataset. The high performance of EfficientNet-B0 indicates that it avails itself to generalize well across multiple plant types and diseases as further presented by its accuracy of 96.98% of F1 Score of 0.9697, a precision of 0.9699, and a recall of 0.9698. Inception V3 came next with an 80.97% accuracy, an F1 Score of 0.8165, reasonable precision at 0.8656, but slightly lower recall at 0.8097.

Keywords. Plant disease detection; deep learning; multi- class classification; transfer learning; pretrained models; efficientnet

1. INTRODUCTION

Plant diseases pose a significant threat to global agricultural productivity, leading to substantial crop losses and severe economic repercussions. Optimization plays a fundamental role in computational intelligence, as it focuses on identifying the best possible solution within a defined search space by maximizing or minimizing an objective function. Within the context of deep learning, optimization is crucial for efficient model training, parameter tuning, and performance enhancement. Traditional optimization methods, such as gradient descent, while effective for convex and differentiable problems, often struggle with the nonlinear, multimodal, and high-dimensional nature of real-world challenges—such as those encountered in complex agricultural image classification tasks [1] [2].

To address these limitations, metaheuristic algorithms have emerged as powerful, adaptive, and flexible optimization techniques. Inspired by natural and biological processes including evolution, animal foraging behavior, and collective intelligence metaheuristics are designed to explore vast and complex search spaces efficiently. Well-known examples include the Genetic Algorithm (GA), inspired by natural selection; Particle Swarm Optimization (PSO), modeled after the social dynamics of bird flocks; Grey Wolf Optimizer (GWO), based on the leadership hierarchy and cooperative hunting of grey wolves; and the Whale Optimization Algorithm (WOA), which emulates the bubble-net feeding behavior of humpback whales. These algorithms effectively balance exploration and exploitation, reducing the likelihood of premature convergence and increasing the probability of identifying global optima [3].

In deep learning-based plant disease detection, metaheuristic optimization techniques have proven increasingly valuable for tasks such as hyperparameter tuning, feature selection, and neural network weight optimization. Their capacity to navigate complex, nonlinear, and multimodal search landscapes enables the development of more accurate, stable, and generalizable models. Furthermore, hybrid metaheuristic strategies combining the strengths of algorithms such as PSO, GA, and WOA have demonstrated enhanced convergence speed and improved generalization in agricultural image analysis. Consequently, the integration of metaheuristic optimization with deep learning architectures represents a promising frontier in precision agriculture, facilitating the creation of intelligent and adaptive systems for automated plant disease diagnosis [4] [5].

2. LITERATURE REVIEW

Recent developments in plant disease identification and diagnosis have witnessed a new shift towards the use of deep learning approaches, as they provide excellent solutions for plant disease diagnosis. The Convolutional neuron network CNN models have been developed particularly to enhance the classification accuracy from simple images of the leaves of healthy or diseased plants. For example, particular research carried out illustrated that a CNN model having a substantial training data set comprised of 87,848 images experienced an accuracy of 99.53% in recognizing particular plant disease combinations [6].

These models hold lots of potential as decision support for early disease identification in real-world agriculture. While methods such as DNA identification and visual inspection are used frequently, they have a host of drawbacks, including an inability to deliver early results in asymptomatic stages of diseases. The newest technologies for real-time disease surveillance include biosensors and remote sensing methods as efficient instruments in

disease identification. These novel approaches can detect early infections directly in the field and function as an additional diagnostic tool to existing ones – promoting a more sustainable agriculture [7] [8]. Due to the dynamic state of plant pathogens, it becomes cumbersome to detect them early and methods of detection have to be advanced. For example, a lightweight deep CNN model incorporates hand-crafted feature descriptors for the enhancement classification of leaf diseases [9].

3. PROPOSED METHODOLOGY

The proposed methodology for plant disease detection and classification is structured around two primary phases: the transfer learning phase and the localization and classification phase. This framework is aimed at maximizing the opportunity of deep learning methodologies to improve the efficiency of plant disease diagnosis. The first module, known as dataset preparation, entails the careful collection of images that constitute the plant disease image dataset which contains images from different countries to enhance universal usage. With over 125,000 images of plant leaves belonging to 10 different plant species and 37 health and disease groups, this dataset comprises images of plant leaves. Test set: It consists of images for model testing and has hidden images to determine its accuracy. Examples of the dataset are shown in Figure 1 to reveal the range of plant diseases and variability of the image conditions that are major difficulties in the detection of plant diseases.



Fig. 1. Samples of the Dataset

4. RESULTS

The findings of this study confirm the usability of the deep learning models utilized for plant disease identification and classification. These statewide performance metrics are provided for each model, to fully evaluate their capacity and capability: For Tiny VGG, its accuracy is determined to be 3.78%, F1 Score of 0.0075, precision of 0.0048, and recall of 0.0378. Tiny VGG failed and even in this simple structure, it was evident that its concept does not have the adequate capability to extract every detail in the datasets.

TABLE I: Performance Metrics of Classification Models

Model	Accuracy	F1 Score	Precision	Recall
Tiny VGG	0.038	0.008	0.005	0.038
EfficientNet-B0	0.970	0.970	0.970	0.970
Inception V3	0.810	0.817	0.866	0.810

However, on the same platform, the EfficientNet-B0 model presented a better performance, the accuracy was 96.98%, the F1 Score was 0.9697, the precision was 0.9699, and the recall of 0.9698. They said that this model shows that it can generalize well across different plants and different classes of diseases. High accuracy and equilibrium F1 Score demonstrating the stability of the model's performance and the ability to measure the difference in the similarity of healthy and diseased leaves. Lastly, using the Inception V3 we got an accuracy of 80.97%, an F1 Score of 0.8165, a precision of 0.8656, and a recall of 0.8097.. Table I.

The confusion matrices give information on the classification capabilities of the models and reveal how cleanly the models can separate one disease from the others.

5. CONCLUSION

The present work can provide a proof of concept of deep learning models for the identification and diagnosis of plant diseases, making use of a dataset that includes more than 125,000 images involving 10 plant species and 37 diseases. Comparing the Tiny VGG, EfficientNet-B0, and Inception V3 models, we found meaningful disparities concerning performance indicators such as accuracy, F1 Score, precision, and recall. Tiny VGG also had negligible ability to classify plant diseases appropriately with a maximum accuracy of 3.78%; The second-best performer was EfficientNet-B0 with a total accuracy of 96.98% and good precision and recall. This implies that EfficientNet-B0 is quite robust in general-izing across several plant species and categories of diseases, which makes it ideal for use in deployment in agricultural applications.

REFERENCES

- [1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [2] A. B. Queyam, R. Kumar, R. K. Ratnesh, and R. K. Chauhan, “LabVIEW-enabled synthetic signal for empowering fetal-maternal healthcare,” *ECS Journal of Solid-State Science and Technology*, vol. 13, no. 5, 2024, doi: 10.1149/2162-8777/ad4dde.
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [4] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [5] R. Kumar, Vevekanandam, and R. Dey, “IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment,” *International Journal of Vegetable Science*, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.
- [6] Y. K. Shejwal, A. J. Hati, J. U. Kidav, A. K. Singh, and R. Kumar, “Neural network-based classification of beamforming matrices,” *International Journal of Satellite Communications and Networking*, 2025, doi: 10.1002/sat.70002.
- [7] R. Kumar, R. K. Ratnesh, R. K. Chauhan, A. Kumar, M. K. Singla, and R. Gupta, “Non-invasive bio-impedance imaging and sensing for medical diagnostics and industrial applications,” *Journal of The Electrochemical Society*, 2024.
- [8] Sindhuja Sankaran, Ashish Mishra, Reza Ehsani, and Cristina Davis. A review of advanced techniques for detecting plant diseases. *Computers and electronics in agriculture*, 72(1):1–13, 2010.
- [9] Khalid M Hosny, Walaa M El-Hady, Farid M Samy, Eleni Vrochidou, and George A Papakostas. Multi-class classification of plant leaf diseases using feature fusion of deep convolutional neural network and local binary pattern. *IEEE Access*, 11:62307–62317, 2023.

<Conference abbreviation>

<Conference Series name>

<Volume number and Year> <DOI Number>

Optimizing Diabetes Prediction through Machine Learning and Feature Selection on Health Indicators

Marwa M. Eid¹, Safaa Zaman², Amal H. Alharbi³, Amel Ali Alhussan⁴, Doaa Sami Khafaga⁵, El-Sayed M. El-kenawy^{6,7}

¹Faculty of Artificial Intelligence Delta University for Science & Technology Mansoura 35111, Egypt Jadara University Research Center Jadara University, Jordan mmm@ieee.org

²College of life sciences , Information sciences department Kuwait University S.3zaman@ku.edu.kw

³Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia ahalharbi@pnu.edu.sa

⁴Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia aaalhussan@pnu.edu.sa

⁵Department of Computer Sciences College of Computer & Information Sciences Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia dkhafaga@pnu.edu.sa

⁶Delta Higher Institute of Engineering and Technology Department for Communications and Electronics Mansoura 35511, Egypt

⁷Applied Science Research Center Applied Science Private University Amman, Jordan, skenawy@ieee.org

Abstract.

Diabetes mellitus is a chronic metabolic disorder with an increasing global prevalence, necessitating improved diagnostic tools for early detection and management. This study investigates the application of six binary feature selection algorithms combined with various machine learning classifiers to predict diabetes in Pima Indian female patients over 21 years old, using health indicators such as glucose levels, BMI, and age. The research systematically compares binary Waterwheel Plant Algorithm (bWWPA), binary Particle Swarm Optimization (bPSO), binary Whale Algorithm Optimization (bWAO), binary Grey Wolf Optimization (bGWO), binary Firefly Algorithm (bFA), and binary Genetic Algorithm (bGA) for optimal feature selection. Subsequently, Random Forest, K-Nearest Neighbors, Decision Tree, and Logistic Regression classifiers were evaluated for diabetes classification performance. The binary Waterwheel Plant Algorithm demonstrated superior feature selection capabilities with the lowest average error rate of 0.44554 and highest best fitness score of 0.41054. Among the machine learning models, Random Forest achieved the highest overall performance with 95.66% accuracy, 94.11% sensitivity, 96.82% specificity, and 94.89% F1-score. The results confirm glucose levels, BMI, and age as the most significant predictors for diabetes diagnosis in this population.

Keywords. Diabetes Prediction, Feature Selection, Machine Learning, Random Forest, Pima Indian Dataset, Health Indicators.

1. INTRODUCTION

Traditional diagnostic tools such as fasting plasma glucose, oral glucose tolerance tests, and HbA1c remain essential but capture only a limited picture of disease risk, often neglecting socio-demographic and lifestyle factors. ML models have achieved notable predictive success, with ensemble methods such as random forests and gradient boosting often attaining AUC scores above 0.85. Nonetheless, the inclusion of redundant or irrelevant features threatens interpretability and risks overfitting, necessitating effective feature selection frameworks that balance parsimony and predictive performance [1] [2].

In this study, six bio-inspired binary optimization algorithms are investigated as feature selection mechanisms: Binary Waterwheel Plant Algorithm (bWWPA), Binary Particle Swarm Optimization (bPSO), Binary Whale Optimization Algorithm (bWOA), Binary Grey Wolf Optimization (bGWO), Binary Firefly Algorithm (bFA), and Binary Genetic Algorithm (bGA). These algorithms, inspired by natural, physical, and sociobehavioral systems, provide population-based search strategies capable of addressing combinatorial explosion and avoiding local optima [3]. Their selected feature subsets are evaluated using four interpretable classifiers—Random Forest (RF), K-Nearest Neighbors (KNN), Decision Tree (DT), and Logistic Regression (LR)—with performance measured through stratified k-fold cross-validation using accuracy, AUC, precision, recall, F1-score, and Matthews correlation coefficient (MCC) [4]. Additionally, Shapley Additive Explanations (SHAP) are applied to interpret the contribution of individual features, with plasma glucose and BMI expected as dominant predictors, while other factors such as skin-fold thickness may also yield valuable clinical insights [5].

2. LITERATURE REVIEW

Diabetes is a growing epidemic with severe consequences for patients and healthcare systems, particularly in Taiwan where prevalence continues to rise [6]. A study on 15,000 outpatients in Taipei tested several ML models, showing that the two-class boosted decision tree achieved the best predictive value with an AUC of 0.991. Diabetes mellitus causes hyperglycemia due to insulin dysfunction, leading to complications if uncontrolled [7]. Early diagnosis is vital, yet current computational models require improvement. Studies combining datasets such as PIMA Indian and Medical City Hospital applied feature selection (Spearman correlation), missing value handling (Polynomial Regression), and classifiers (RF, SVM, DNN). Results indicated that a twice-growth deep neural network achieved superior precision and accuracy. Diabetic retinopathy (DR), a common complication, can cause blindness; recent deep learning systems like DeepDR showed high performance in lesion detection and DR grading [8], with AUC values above 0.94. DeepDR-LLM further integrated language models and imaging tools to assist primary care physicians, improving diagnostic efficiency and patient self-care [9].

3. DISCUSSION AND RESULTS

A. Dataset

The data used in this study are derived from the National Institute of Diabetes and Digestive and Kidney Diseases specifically diagnosing indicator for diabetes among the patients. The data in particular refers to female patients only over 21 years of age with Pima Indian ancestry. Included in the dataset is several health metrics that act as predictors for diabetes

presence. Box plots of each of the health measures in the diabetes prediction dataset are shown in Figure 1. These box plots show max, min, second quartile, and first quartile of each feature to compare the differences between diabetic and non-diabetic instances. For example, glucose levels, BMI and age have higher median and upper quartile among diabetics thereby stressing their importance in diabetes prediction. This plot gives the first impression of the main contributing factors of diabetes by just observing for extreme values and scattering of the points.

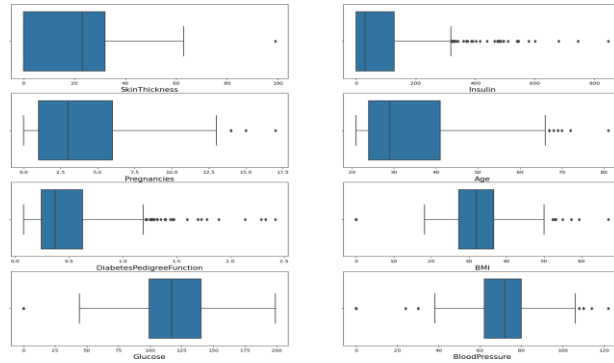


Fig. 1. Box Plots of Key Health Indicators for Diabetes Prediction Dataset

B. Feature Selection Results

The effectiveness of six feature selection techniques namely bWWPA, bPSO, bWAO, bGWO, bFA and bGA is depicted in Table I. Since the performance depends on various factors, performance is evaluated in terms of average error, size of the select, and the fitness values. For instance, the supports of bWWPA were the lowest average of error and the highest best fitness, pointing out that bWWPA has the best performance in selecting important features among all the examined models. Lower error rates and high fitness values suggest that each of the methods utilized is capable of selecting only the most relevant features for diabetes prediction.

C. Machine Learning Results

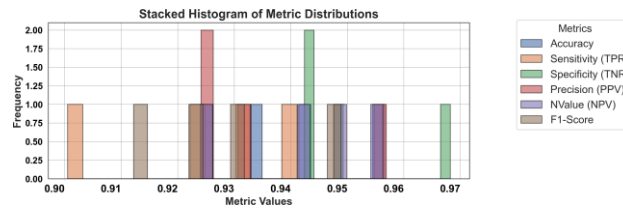


Fig. 2. Stacked Histogram of Diabetes Prediction Model Performance Metrics

The results of different machine learning classifiers, such as Random Forest, KNN, Decision Tree, and Logistic Regression. All predictor models' evaluations of accuracy, sensitivity, specificity and others are given, while the Random Forest Classifier showed the highest accuracy and specificity of all the models tested. Figure 2 presents a stacked histogram detailing the distribution of key evaluation metrics for a Diabetes Prediction model leveraging machine learning on health indicators. The graph juxtaposes the frequency of

different metric scores, providing a visual comparison of the model's effectiveness. Precision (PPV), representing the proportion of correctly predicted positive cases among all predicted positive cases, exhibits the highest frequency, peaking at a specific metric value. The distribution of Precision shows frequency of metric at 120.

4. CONCLUSION

This study demonstrates the strong potential of machine learning models, particularly when combined with effective feature selection, in improving the diagnosis of diabetes among Pima Indian women. The findings underscore the importance of combining robust feature selection techniques with strong classifiers to address complex diagnostic challenges. Such systems can provide healthcare practitioners with valuable decision-support tools for identifying high-risk individuals earlier, enabling timely intervention and better allocation of resources. While this research focused on the Pima Indian cohort, the methodological framework holds promise for broader application across diverse populations and healthcare settings.

REFERENCES

- [1] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [2] R. Kumar, Vevekanandam, and R. Dey, "IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment," *International Journal of Vegetable Science*, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.
- [3] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [4] R. Kumar, M. K. Singla, S. A. Muhammed Ali et al., "Parameter estimation of proton exchange membrane fuel cell using hybrid grouping biogeography optimization algorithm," *Ionics*, 2025, doi: 10.1007/s11581-025-06600-x.
- [5] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [6] Y. K. Shejwal, A. J. Hati, J. U. Kidav, A. K. Singh, and R. Kumar, "Neural network-based classification of beamforming matrices," *International Journal of Satellite Communications and Networking*, 2025, doi: 10.1002/sat.70002.
- [7] C. C. Olisah, L. Smith, and M. Smith. Diabetes mellitus prediction and diagnosis from a data preprocessing and machine learning perspective. *Computer Methods and Programs in Biomedicine*, 220:106773, 2022.
- [8] L. Dai, L. Wu, H. Li, C. Cai, Q. Wu, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nature Communications*, 12(1):3242, 2021.
- [9] R. K. Ratnesh, R. K. Chauhan, R. Kumar, M. Afroz, and J. Singh, "Recent advancements of plasmonic sensor technologies in healthcare and electronics," *Journal of Molecular Structure*, 2025, Art. no. 142945, doi: 10.1016/j.molstruc.2025.142945.

Metaheuristic Optimization in Space Science and Engineering: A Comprehensive Literature Review

Amal H. Alharbi¹, Amel Ali Alhussan², Ebrahim A. Mattar³, Doaa Sami Khafaga⁴

El-Sayed M. El-kenawy^{5,6}, Marwa M. Eid⁷

¹Department of Computer Sciences College of Computer

& Information Sciences Princess Nourah bint Abdulrahman University Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia ahalharbi@pnu.edu.sa

²Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia aaalhussan@pnu.edu.sa

³College of Engineering, University of Bahrain, Bahrain ebmattar@uob.edu.bh

⁴Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia dskhafga@pnu.edu.sa

⁵Delta Higher Institute of Engineering and Technology Department for Communications and Electronics Mansoura 35511, Egypt

⁶Applied Science Research Center Applied Science Private University Amman, Jordan, skenawy@ieee.org

⁷Faculty of Artificial Intelligence Delta University for Science & Technology Mansoura 35111, Egypt Jadara University Research Center Jadara University, Jordan mmm@ieee.org

Abstract.

Metaheuristic algorithms have become potent and flexible in solving complex problems in various contexts, including load forecasting, power outage minimization, energy dispatch, and cybersecurity. This paper conducts a literature review of the research currently being conducted in metaheuristic optimization techniques for addressing creative grid systems, trends in hybridization, algorithm performance, real-time flexibility, and their application fields. The review establishes the boundaries and validates the authority of the most commonly used methods, including PSO, GA, ACO, and GWO, as well as the currently used hybrid and AI-combined methods. Although successful, their flaws are found in scalability, real-world usage, benchmarking consistency, and integration with cybersecurity. It also highlights some of the recent research trends in the field, including agent-based optimization, sustainability-conscious goals, explainability, and the interactions between metaheuristics, machine learning, and edge computing.

Keywords. Metaheuristic Optimization, Smart Grid, Load Forecasting, Energy Management, Cyber-Physical Systems

1. INTRODUCTION

The historical approach to electrical power generation and distribution considered and modelled traditional power systems with linear energy flow in mind, where electricity was generated in centralized power plants and then flowed hierarchically directly to end

consumers. These legacy systems, which have been highly effective over the past few decades, are now viewed as insufficient and unable to meet the emerging grid requirements, such as the realization of two-way power flow, distributed generation, close-grained real-time monitoring, and distribution-based control strategies. Such a radical shift as the smart grid has turned out to be quite promising in addressing these limitations. Smart grids serve as an enabling infrastructure for a broad array of functionalities, not only facilitating the seamless integration of renewable energy sources such as solar and wind but also supporting advanced mechanisms for demand-side management (DSM), coordinated control of electric vehicle (EV) charging, and implementation of dynamic and time-sensitive pricing models [1].

The augmented complications and inadequacies of more conventional optimization methods have prompted a significant rise in interest among the research community in studying and creating metaheuristic optimization methods. These algorithms are inspired by natural and biological mechanisms, such as evolution, swarm behavior, and animal hunting strategies, and provide a strategic tradeoff between global exploration (search) and local refinement (exploitation). Together with the fact that they are used in the acting process with solutions of gigantic, high-dimensional, and non-congruent spaces add to the fact that only minimal structural assumptions about the region of the problem studied are made in their use and application [2]. This natural flexibility and scalability has facilitated their successful implementation in a wide scope of novel grid applications including (but not limited to) optimal power flow (OPF), unit commitment (UC), scheduling of energy storage systems, integration of renewable energy sources, and resilience of cyberspace physical systems enhancement using metaheuristic algorithms [3].

2. LITERATURE REVIEW

The marked uptake of renewable energy sources, the rapid growth of electric vehicles (EVs), and the accelerated development and decentralization of distributed energy resources (DERs) have all contributed to changing the very paradigm of how traditional electrical power systems operate. The new improvements have necessitated excessive complexity and dynamism in the activities of the grids, thereby rendering the legacy infrastructure and conventional methods of handling the grids ineffective in meeting the demands of present-day power systems [4]. On this note, smart grids are not only now an enabling infrastructure, i.e., a sophisticated mix of electrical and information technologies, but also a multi-layered and multi-functional platform to monitor, control, and optimize grid performance. Even though these developments are very encouraging, associated computational and algorithmic issues are enormous due to the characteristic features of smart grid problems (i.e., high degree of nonlinearity, large-scale uncertainty, and the problems are often large-scale and multidimensional) [5]. In the complexity and uncertainty of such a high level, traditional deterministic optimization methods, which usually assume some form of convexity, continuity, or gradient, have not performed well. This has resulted in one of the most critical areas in smart grid research circles, such as the applications of metaheuristic algorithms, gaining significant popularity. Such algorithms have shown high potential due to their robustness, flexibility, and intrinsic ability to address NP-hard optimization problems with a wide range of applications in the smart grid [6], [7]. Metaheuristics, including well-established approaches such as Particle Swarm Optimization (PSO), Genetic Algorithms (GA), Ant Colony Optimization (ACO), Grey Wolf Optimizer (GWO), and Harmony Search

(HS), have showcased notable efficacy in addressing a wide range of optimization problems within smart grids. These include, among others, duties related to energy management, fault detection, load prediction, and demand response. The fundamental power of these algorithms lies in their ability to easily search and exploit complex, multimodal, and high-dimensional solution spaces using biologically and naturally motivated algorithms of attraction, swarming, evolution, and foraging. For example, Hafeez et al. [8] proposed an improved version of the PSO algorithm tailored for optimal residential load scheduling. Their enhancement addressed critical issues, such as premature convergence and the trade-off between global exploration and local exploitation, and their results demonstrated superior performance when benchmarked against standard PSO, GA, and Wind-Driven Optimization (WDO). Similarly, many researchers introduced a novel hybrid algorithm that combined Differential Evolution with Earthworm Optimization. Their strategy became a source of significant savings in terms of energy cost and the peak-to-average ratio of load, providing. One area where metaheuristics have played a particularly prominent and transformative role is in smart home energy management systems (HEMS). These systems are designed to manage and coordinate household appliances, distributed generation sources (e.g., rooftop solar panels), and energy storage devices, particularly under real-time pricing schemes and demand response signals [4-8].

3. DISCUSSION

The reviewed literature clearly establishes metaheuristics as a central computational tool in addressing complex optimization challenges in modern smart grids. These algorithms—ranging from GA, PSO, and GWO to hybridized and AI-integrated versions—have demonstrated their effectiveness in diverse applications, including load forecasting, distributed generation, and cybersecurity. Their ability to manage nonlinearities, multi-objective formulations, and real-time constraints positions them as highly adaptable solutions. Notably, the research community is increasingly emphasizing hybridization, where combining complementary techniques or coupling with neural networks helps overcome weaknesses such as premature convergence or limited exploitation. This trajectory reflects a maturation of the field and signals a shift toward adaptable, problem-specific solutions rather than a reliance on universal, standalone algorithms.

At the same time, the literature highlights critical methodological and practical gaps. Chief among them is the overreliance on simulation-based validation, often using simplified IEEE bus systems that fail to capture the complexities of the real world. This creates uncertainty about the generalizability and robustness of proposed models. Moreover, the absence of standardized evaluation frameworks hampers fair benchmarking of algorithmic performance, as different studies employ varying performance metrics and assumptions. The lack of reproducible testbeds and unified criteria hinders cumulative progress and makes meaningful comparisons challenging. Addressing these limitations through open-source datasets, standardized protocols, and collaborative benchmarking infrastructures is essential to move research from theoretical demonstration to field-ready deployment.

4. CONCLUSION

This review presents a thorough analysis of the application of metaheuristic optimization methods in the design and operation of smart grids. With the growing decentralization, data intensity, and dynamicity of modern power systems, adaptive, robust, and scalable

optimization tools have become central. Metaheuristics — including Genetic Algorithms (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Grey Wolf Optimizer (GWO), and numerous hybridized variants — have emerged as highly effective candidates for solving complex, nonlinear, and multi-objective problems across various smart grid domains. The literature highlights several key limitations: limited real-world deployment, a lack of standardized benchmarking, underdeveloped uncertainty modeling, and insufficient emphasis on real-time performance and cyber-physical resilience. Many promising algorithms have been validated only in simulation using simplified test systems, raising concerns about scalability and practical integration into utility-level operations. To move the field forward, future research should prioritize the development of real-time, interpretable, and secure optimization frameworks. Efforts should be made to test algorithms on large-scale and realistic datasets, incorporate probabilistic modeling for renewable uncertainty, and ensure resilience against adversarial disruptions.

REFERENCES

- [1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [2] P. K. Pandey, Z. El Khatib, R. Kumar, M. Singla, M. Safaraliev, and F. Kamalov, “A novel hybrid control strategy for grid-connected PV-battery system under different operating conditions,” *Przegląd Elektrotechniczny*, vol. 2024, no. 9, p. 287, 2024, doi: 10.15199/48.2024.09.56.
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [4] R. Kumar, Vevekanandam, and R. Dey, “IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment,” *International Journal of Vegetable Science*, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.
- [5] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [6] Mehran Memari, Ali Karimi, and Hamed Hashemi-Dezaki. Reliability evaluation of smart grid using various classic and metaheuristic clustering algorithms considering system uncertainties. *International Transactions on Electrical Energy Systems*, 31(6):e12902, 2021. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/2050-7038.12902>.
- [7] R. Kumar, D. Singh, M. K. Singla, E. S. M. El-Kenawy, and A. H. Alharbi, “Optimizing the placement of distributed renewable energy resources in large-scale distribution systems,” *Journal of Electrical Engineering & Technology*, pp. 1–25, 2025, doi: 10.1007/s42835-025-02279-2.
- [8] M. K. Singla, J. Gupta, A. Gupta, M. Safaraliev, H. Zeinoddini-Meymand, and R. Kumar, “Enhancing solid oxide fuel cell efficiency through advanced model identification using differential evolutionary mutation fennec fox algorithm,” *International Journal of Computational Intelligence Systems*, vol. 18, p. 38, 2025, doi: 10.1007/s44196-025-00759-x

BER-CNN: Bio-Inspired Evolutionary Refinement of CNNs for Rapid and Accurate Tuberculosis Detection

Marwa M. Eid^{1,2}

¹*Faculty of Artificial Intelligence Delta University for Science
& Technology Mansoura 35111, Egypt*

²*Jadara University Research Center Jadara University
Jordan mmm@ieee.org*

Abstract.

In this study, we present BER-CNN, a novel deep learning framework optimized through bio-inspired evolutionary refinement for automated TB detection. The model was trained and evaluated on a curated dataset of 928 sputum smear images containing 3,734 annotated bacilli. Comprehensive benchmarking against four optimization-driven CNN models—GWO-CNN, PSO-CNN, WOA-CNN, and GA-CNN—demonstrates that BER-CNN achieves superior performance across all metrics, with an accuracy of 97.22%, sensitivity of 96.86%, specificity of 97.53%, and F-score of 97.12. Beyond predictive accuracy, BER-CNN reduces computational cost by converging faster than competing approaches, making it practical for large-scale deployment. These results highlight the potential of BER-CNN to transform TB diagnostics by providing an accurate, efficient, and scalable solution that bridges the gap between advanced machine learning and clinical practice..

Keywords. Tuberculosis Detection, Deep Learning, Bio-Inspired Optimization, Convolutional Neural Networks (CNNs), Medical Image Analysis

1. INTRODUCTION

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (Mtb), remains one of the world's deadliest infectious diseases. Despite the availability of effective treatment, TB continues to affect millions each year and is responsible for over a million deaths annually. Artificial intelligence (AI), particularly deep learning (DL), has revolutionized multiple areas of healthcare by enabling the automated analysis of complex biomedical data [1]. DL has demonstrated strong performance in radiology, pathology, genomics, and infectious disease detection, offering improvements in both speed and accuracy compared to human experts. In the field of TB diagnostics, DL models have been applied to several modalities, including sputum smear microscopy, chest radiography, Raman spectroscopy, and metabolomic profiling. These approaches have shown the potential to improve diagnostic yield, reduce human workload, and provide scalable solutions for low-resource environments. Importantly, DL frameworks have the capacity to integrate heterogeneous data sources, enabling multi-modal TB detection strategies that go beyond the limitations of single diagnostic techniques [2] [3]. While the promise of deep learning in TB diagnostics is well recognized, several research gaps remain. First, many DL approaches are limited to small or homogeneous datasets, raising questions about their robustness and generalizability in real-world clinical settings. Second, despite reporting high performance, issues of model interpretability, reproducibility, and integration into clinical workflows remain

underexplored. Third, most studies focus on a single optimization strategy or architecture, leaving opportunities for further improvement in accuracy and computational efficiency. Finally, there is limited comparative benchmarking of deep learning models optimized with bio-inspired algorithms for TB detection, which could reveal the most effective frameworks for real-world deployment [4]. To address these gaps, we develop and evaluate a set of deep learning models enhanced by bio-inspired optimization techniques for TB detection. Specifically, we benchmark the performance of the following models: Grey Wolf Optimizer CNN (GWO-CNN), Particle Swarm Optimizer CNN (PSO-CNN), Whale Optimization Algorithm CNN (WOA-CNN), and Genetic Algorithm CNN (GA-CNN). Against these baselines, we introduce our proposed model, BER-CNN, which integrates an advanced optimization mechanism to enhance feature selection and classification accuracy [5].

2. LITERATURE REVIEW

To address this, one study proposed an automated examination framework using an adapted UNet model to localize *Mycobacterium tuberculosis* (Mtb) and quantify their lipid content and dimensions, achieving improved accuracy and efficiency compared to manual evaluation [6]. Similarly, another group developed a deep learning pipeline for bacterial quantification in sputum samples using fluorescence microscopy. Their approach combined generative adversarial networks, patch extraction, classification, and regression, reducing error rates and bypassing the variability of manual counting [7]. Deep learning has also been employed as a supportive tool for pathologists, automating the identification of acid-fast bacilli in microscopy images, and significantly reducing workload without compromising diagnostic accuracy [8]. Alternative modalities have also emerged for rapid and reliable detection. Surface-enhanced Raman spectroscopy (SERS) coupled with deep learning was shown to effectively distinguish between sputum samples with or without Mtb, as well as classify strains based on site of infection and antibiotic resistance profiles, demonstrating high accuracy with handheld devices [9]. Metabolic phenotyping through urine analysis further demonstrated potential in differentiating TB from nontuberculous mycobacteria (NTM), identifying distinct metabolite signatures—particularly involving amino acid metabolism—that could be exploited for diagnostic biomarkers [10]. Likewise, bacterial volatilome analysis using HS-SPME-GC-MS revealed discriminant volatile organic compounds (VOCs) that allow differentiation of *Mycobacterium* species, highlighting the potential for clinically translatable breath- or sputum-based biomarkers [11].

3. DATASET DESCRIPTION AND PREPROCESSING

Tuberculosis (TB) is an infectious disease caused primarily by the bacterium *Mycobacterium tuberculosis* (Mtb). While TB most commonly affects the lungs, it can also spread to other organs, leading to extrapulmonary manifestations. A large proportion of TB infections remain latent, in which individuals carry the bacteria without showing symptoms. However, active TB presents with characteristic symptoms such as a persistent cough (sometimes blood-streaked), weight loss, night sweats, and fever. Transmission occurs when infected individuals expel aerosolized droplets through coughing or sneezing.

In total, the dataset contains 3,734 annotated bacilli across the 928 images. The annotation details are stored in associated XML files, where each file corresponds to an image and encodes the bounding box coordinates for all bacilli present. These structured annotations

enable supervised training of deep learning models, allowing them to learn both classification (presence/absence of bacilli) and localization (precise position of bacilli) tasks.

A. Data Preprocessing

Prior to training, the dataset underwent a series of preprocessing steps to ensure consistency, reduce noise, and enhance model performance through steps like Image normalization, Color channel adjustment, Annotation parsing, Data augmentation, Dataset splitting. Through these preprocessing steps, the dataset was prepared for the development of robust deep learning models capable of accurately detecting and localizing TB bacilli in sputum smear images. This structured approach ensured both data quality and consistency, laying the foundation for the subsequent training and evaluation stages.

B. Deep Learning Models

1) Convolutional Neural Network (CNN): Convolutional Neural Networks (CNNs) form the backbone of many computer vision tasks, including medical image analysis, due to their ability to automatically extract hierarchical features from input images. In the context of tuberculosis (TB) detection, CNNs have been employed to process sputum smear images, capturing spatial patterns that correspond to the presence of *Mycobacterium tuberculosis* bacilli. Initial experiments with a baseline CNN demonstrated promising results, but the performance was limited by manual tuning of hyperparameters and network design choices.

4. DEEP LEARNING PERFORMANCE RESULTS

From the results, it is evident that BER-CNN achieved the highest accuracy of 97.22%, surpassing all other models by a significant margin. In particular, it recorded an improvement of more than 5% over the weakest baseline (GA-CNN) and approximately 2.5% over the strongest competing baseline (GWO-CNN). This superior performance was consistent across other metrics as well, with BER-CNN reaching a sensitivity of 96.86% and specificity of 97.53%, effectively minimizing both false negatives and false positives. The positive predictive value (97.37%) and negative predictive value (97.08%) further reinforce the robustness of the model in distinguishing between TB-positive and TB-negative cases.

To further explore the relationships between the different evaluation metrics, a pairplot of the performance results was generated, as illustrated in Figure 1. This visualization shows pairwise comparisons of accuracy, sensitivity, specificity, precision (PPV), negative predictive value (NPV), and F-score, with regression lines added to highlight correlations. The diagonal plots display the distribution of each individual metric, while the off-diagonal scatter plots with fitted regression lines provide insight into how improvements in one metric relate to others. From the pairplot, it is evident that the metrics are strongly and positively correlated, particularly between accuracy, sensitivity, and F-score, which reinforces the conclusion that models achieving higher accuracy also tend to maintain balanced predictive performance across other key criteria.

This study presented a comparative evaluation of deep learning models enhanced with bio-inspired optimization strategies for the detection of tuberculosis (TB) from sputum smear images. The proposed BER-CNN framework consistently outperformed competing models (GWO-CNN, PSO-CNN, WOA-CNN, and GA-CNN) across all evaluation metrics. BER-

CNN achieved the highest accuracy (97.22%), sensitivity (96.86%), specificity (97.53%), and F-score (97.12%), representing a significant improvement over baseline approaches. The superior balance between sensitivity and specificity demonstrates BER-CNN's effectiveness in reducing both false negatives and false positives, which is essential in medical diagnostics. Furthermore, BER-CNN demonstrated lower computational costs, achieving faster convergence compared to other optimization-based CNNs. This establishes BER-CNN as both an accurate and efficient framework suitable for early TB detection.

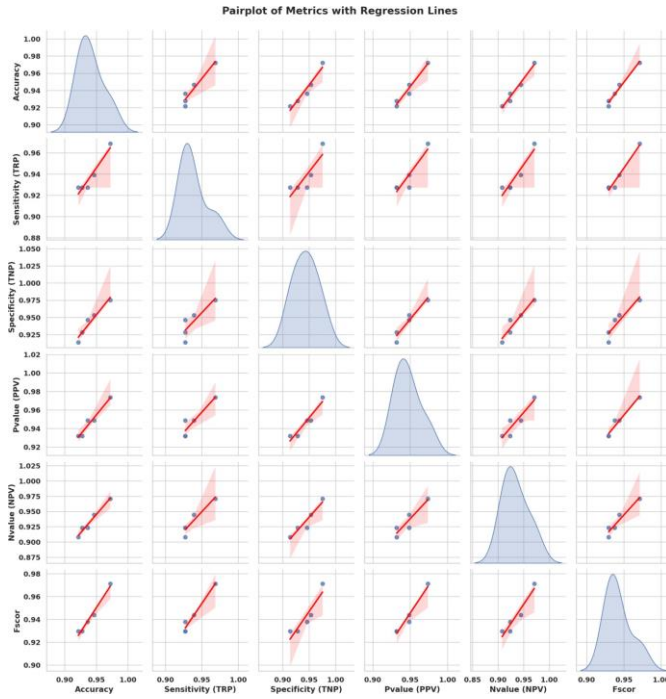


Fig. 3. Pairplot of model performance metrics with regression lines, showing relationships among accuracy, sensitivity, specificity, PPV, NPV, and F-score.

While the results demonstrate the effectiveness of BER-CNN, several avenues for future research remain. First, larger and more diverse datasets should be employed to further validate the model's generalizability across different patient populations, imaging devices, and clinical settings. Second, explainability and interpretability techniques should be incorporated to make the decision-making process of BER-CNN more transparent to clinicians, fostering trust and clinical adoption. Third, integration of multimodal data—such as combining sputum imaging with radiographic, genomic, or clinical features—could provide more comprehensive diagnostic insights. Fourth, exploring lightweight architectures and model compression methods could make the framework even more suitable for mobile or handheld diagnostic devices. Finally, longitudinal studies assessing the impact of deploying such models in real-world healthcare systems would provide critical evidence for large-scale implementation.

5. CONCLUSION

BER-CNN offers a promising pathway toward accurate, efficient, and scalable TB diagnostics. With continued refinement, validation, and integration into health-care workflows, optimized deep learning models such as BER-CNN have the potential to play a transformative role in global TB control efforts.

REFERENCES

- [1] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. IEEE Access, pages 1–1, 2025.
- [2] Y. K. Shejwal, A. J. Hati, J. U. Kidav, A. K. Singh, and R. Kumar, “Neural network-based classification of beamforming matrices,” International Journal of Satellite Communications and Networking, 2025, doi: 10.1002/sat.70002.
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El- Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. Scientific Reports, 15(1):16612, 2025.
- [4] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. Arabian Journal for Science and Engineering, pages 1–19, 2025.
- [5] R. Kumar, Vevekanandam, and R. Dey, “IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment,” International Journal of Vegetable Science, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.
- [6] R. K. Ratnesh, R. K. Chauhan, R. Kumar, M. Afroz, and J. Singh, “Recent advancements of plasmonic sensor technologies in healthcare and electronics,” Journal of Molecular Structure, 2025, Art. no. 142945, doi: 10.1016/j.molstruc.2025.142945.
- [7] Marios Zachariou, Ognjen Arandjelovic', Wilber Sabiiti, Bariki Mtafya, and Derek Sloan. Tuberculosis bacteria detection and counting in fluorescence microscopy images using a multi-stage deep learning pipeline. Information, 13(2), 2022.
- [8] S. Arora, S. Sharma, R. Anand, R. Kumar et al., “Integration of metamaterials for quintuple band-notched ultra-wideband antennas,” Analog Integrated Circuits and Signal Processing, vol. 123, 2025, doi: 10.1007/s10470-025-02355-7.
- [9] Liang Wang, Xue-Di Zhang, Jia-Wei Tang, Zhang-Wen Ma, Muhammad Usman, Qing-Hua Liu, Chang-Yu Wu, Fen Li, Zuo-Bin Zhu, and Bing Gu. Machine learning analysis of sers fingerprinting for the rapid determination of mycobacterium tuberculosis infection and drug resistance. Computational and Structural Biotechnology Journal, 20:5364–5377, 2022.
- [10] Nguyen Ky Anh, Nguyen Ky Phat, Nguyen Quang Thu, Nguyen Tran Nam Tien, Cho Eunsu, Ho-Sook Kim, Duc Ninh Nguyen, Dong Hyun Kim, Nguyen Phuoc Long, and Jee Youn Oh. Discovery of urinary biosignatures for tuberculosis and nontuberculous mycobacteria classification using metabolomics and machine learning. Scientific Reports, 14(1):15312, 2024.

Optimizing Crop Recommendations Using Advanced Machine Learning Techniques

Marwa M. Eid

*Faculty of Artificial Intelligence
Delta University for Science
& Technology Mansoura 35111, Egypt*

*Jadara University Research Center
Jadara University
Jordan
mmm@ieee.org*

Abstract—This work aims to find the best crop recommendation system where we employ several machine learning algorithms using a dataset containing soil and environmental factors. Some of the variables in the dataset are Nitrogen; Phosphorus; Potassium; Temperature; Humidity; pH value; and Rainfall. We also compare the models, including Logistic Regression, Support Vector Machine (SVM) classifier, K-Nearest Neighbors (KNN), Decision Tree, and Extra Trees, to determine the most appropriate crops according to the growing conditions. Random Forest is used for feature selection to determine which features have the most influence on the models. According to the evaluation results of the various models, the best one, Extra Trees, has a test accuracy of 0.95, a test F1 score of 0.95, and a test precision of 0.95. On the other hand, the other models show slightly lower accuracy with test accuracy of 0.9 for both Logistic Regression and Support Vector Machine models. KNN achieving The Random forest algorithm was found to be the most accurate model.

Index Terms—Crop recommendation , machine learning , feature engineering , model evaluation , agricultural productivity

I. INTRODUCTION

Crop recommendation is an important focal area in modern agriculture that is central to the achievement of high yields, proper utilization of resources, and promotion of sustainable practices. Considering the crops that are suitable to be grown on the given type of soil and the environment determines the overall yield in agriculture and how effectively the resources are utilized. This study, however, seeks to overcome this challenge by adopting a broad dataset that comprises factors like Nitrogen, Phosphorus, Potassium, Temperature, Humidity, pH Value, and Rainfall. These variables are very essential to determine the level of fertility of soils as well as determine the environmental conditions that are conducive to the growth of various crops.

To get accurate crop recommendations, we perform different types of Machine Learning Algorithms such as Logistic Regression, SVM, KNN, Decision Trees, and Extra Trees. This is done to establish the efficiency of the models in identifying the appropriate crops relevant to specific environments. It also involves proper data pre-processing methods to increase

the accuracy of the models tested on the dataset. Further, the result of converting features with StandardScaler, and encoding with OrdinalEncoder and LabelEncoder guarantees proper preparation of the data for analysis.

In recent years, optimization and metaheuristic algorithms have become increasingly relevant in the broader field of agricultural analytics and computational intelligence [1] [2]. Although not directly implemented in this research, such methods—like Genetic Algorithms, Particle Swarm Optimization, Grey Wolf Optimizer, and Whale Optimization Algorithm—have demonstrated strong capabilities in addressing complex optimization challenges across agriculture [3]. These algorithms are inspired by natural phenomena such as evolution, animal behavior, and swarm intelligence, enabling them to find optimal or near-optimal solutions where traditional methods may fail. In the context of crop recommendation, optimization approaches can be used to fine-tune model parameters, improve feature selection, and enhance decision-making by efficiently exploring large solution spaces [4]. Their adaptability and robustness make them valuable for solving multidimensional agricultural problems, such as yield maximization, irrigation scheduling, and nutrient optimization, thereby complementing data-driven and machine learning-based systems [5].

This research work aims to take advantage of data-driven methods to achieve accurate crop advisement, thus promoting improved crop management in agriculture. In the context of this study, it is planned to enhance the efficiency of resource utilization, increase crop productivity, and advance sustainable agricultural practices by incorporating refined machine learning algorithms and algorithms for data preprocessing. This paper offers a comprehensive analysis of model performance and recommendations on the ability of data-driven approaches to improve agricultural yield and hence the overall welfare of agricultural businesses and farmers.

II. LITERATURE REVIEW

This literature review synthesizes several related studies that explain how machine learning models and data preprocessing

approaches have been used in different disciplines. Therefore, machine learning in crop recommendation systems is vast, especially in data preprocessing including feature engineering leading to high accurate Extra Trees model [6]. Further, while proposing ensemble learning models for sunshine duration prediction, trials have shown feature selection and optimization methods such as Grey Wolf and Stochastic Fractal Search for enhanced performance [7].

For disease diagnosis, various machine learning models have been used in the medical field especially the convolutional neural networks with an understanding that diagnosis requires proper feature balancing as well as classifier optimization as embraced by [8] in the diagnosis of COVID-19. Recent literature on class imbalance demonstrates that the traditional measures of model performance can fall short, especially if datasets are imbalanced, while the use of measures such as the Matthews Correlation Coefficient can offer a clearer picture of a model's performance and functioning [9].

Decision tree and random forest methods have been found to fit into the ensemble family of methods used in clinical research for the prediction of an outcome that exhibits good performance with small sample sizes. These models are useful as dependable decision-aiding systems within the healthcare domain [10]. Moreover, the use of machine learning frameworks presented in detecting the risk of Autism Spectrum Disorder using logistic regression analysis and feature analysis also displays how diagnostic tools can be of assistance in early detection to cut the costs of high-impact diseases [11]. Recent advancements in machine learning and optimization techniques have shown significant potential in agricultural applications, particularly in crop recommendation systems. Network intrusion detection methods utilizing feature selection and hybrid metaheuristic optimization [12] demonstrate the effectiveness of combining traditional machine learning approaches with modern optimization algorithms. Similarly, wind power prediction models based on machine learning and deep learning [13] showcase the versatility of these techniques across different domains, providing valuable insights into how predictive models can be adapted for agricultural forecasting. The integration of image processing and artificial intelligence for early diagnosis applications [14] further illustrates the potential for visual-based agricultural monitoring systems, while optimized ensemble algorithms for parameter prediction [15] offer robust frameworks that can be applied to crop yield estimation and recommendation systems.

III. MATERIAL AND METHODS

As Shown in figure 1 The following are the main steps that have been proposed for the development of a crop recommendation system: First, data collection is conducted on soil nutrient characteristics – Nitrogen, Phosphorus, and Potassium levels, environment characteristics – Temperature, Humidity, pH, Rainfall, and crop characteristics. Missing values are handled by imputation, and outliers are removed by using the IQR method; besides, skewed features are transformed logarithmically, using root functions, and power functions.

Features are then normalized by standardization using the Standard-Scaler while the categorical data is encoded with the help of the Ordinal-Encoder and Label-Encoder. Feature engineering creates new variables like the ratio of Nitrogen, Potassium, and Phosphorus (NP_Ratio, NK_Ratio, PK_Ratio) and environmental indices which include temperature humidity index and rainfall humidity index (Temp_Humidity_Index, Rainfall_Humidity_Index), and mean NPK (NPK_Average), and categorizing the soil ph level. Feature selection is done using an algorithm called Random Forest in the context of crop prediction based on the most influential features. The selected features are then applied to create and develop machine learning models where amongst them are Logistic Regression, SVM, KNN, Decision Tree, and Extra Trees. For each model, the number of correct instances, accuracy, F1 measure, precision, and recall are reported. Extra Trees gave better results against other models with an accuracy of 98.86% accuracy and an F1 score of 98.86%, so it is considered to be the optimal option for the crop recommendation system.

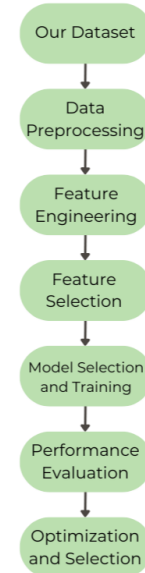


Fig. 1. Proposed Framework

A. Dataset Description

The dataset adopted for the current research comprises relevant information on agricultural and environmental factors that are essential for crop recommendations. Each feature provides valuable information about soil and climatic conditions affecting crop growth and yield:

- **Nitrogen (N):** Estimates the quantity of nitrogen in the soil, which is an essential nutrient crucial for plant growth and development.
- **Phosphorus (P):** Represents the proportion of phosphorus in the soil, important for energy transfer and root formation.

- **Potassium (K):** Indicates the proportion of potassium, an essential nutrient responsible for water regulation and overall plant growth.
- **Temperature:** The ambient temperature in degrees Celsius around the plant, significantly influencing metabolic and growth processes.
- **Humidity:** Represents the percentage of moisture in the air, affecting plant respiration rates, transpiration rates, and water availability.
- **pH_Value:** The soil pH, indicating soil acidity or alkalinity, which significantly affects nutrient solubility and availability to plants.
- **Rainfall:** The total rainfall measured in millimeters, determining water availability and significantly influencing plant hydration and growth.

B. Data Preprocessing

Data preprocessing is the first step in the data science process where input data is cleaned, transformed, and formatted to fit the downstream machine learning processes. This process makes sure that the data they obtain is valid, reliable, and ready for model building and assessment. The preprocessing phase encompasses several critical steps: missing value management, feature scaling, feature encoding, dealing with skewed data, and feature creation and feature reduction. Every stage includes particular procedures and mathematical operations that are aimed at optimizing the quality of data and boosting the efficacy of the model.

C. Descriptive Analysis of the Data

The Descriptive Analytics Section is the initial investigation of the given dataset, utilizing tabular analysis and a set of visual tools to identify the patterns and connections inherent in the agricultural data. This section focuses on the relationship between these core quantitative variables Nitrogen, Phosphorus, Potassium, Temperature, Humidity, pH Value, and Rainfall, and presents the correlation matrix with a heatmap to depict the values involved and their correlation or dependency. The pair plot further categorizes these relationships by presenting scatter plots and histograms that depict interactions and distributions within various categories. Conversely, the box plot provides an unobstructed view of the distributional nature of these variables with focal points on variation and outlier presence. Thus, using these data visualizations, the section provides invaluable information to explain the structure and variability of the data under analysis. They serve as useful instruments for further analysis of the dataset and potential decision-making based on it while giving an insight into what the set consists of. In other words, the Descriptive Analytics Section provides the big picture for the data and unravels the patterns that agricultural decision-makers ought to consider in their decisions.

This is well depicted in the heatmap illustrated in Figure 2 where the agricultural variables include Nitrogen, Phosphorus, Potassium, Temperature, Humidity, pH Value, and Rainfall. The heatmap incorporates the use of color where red represents

a high positive correlation, green represents a low correlation, and blue represents a high negative correlation among other colors.

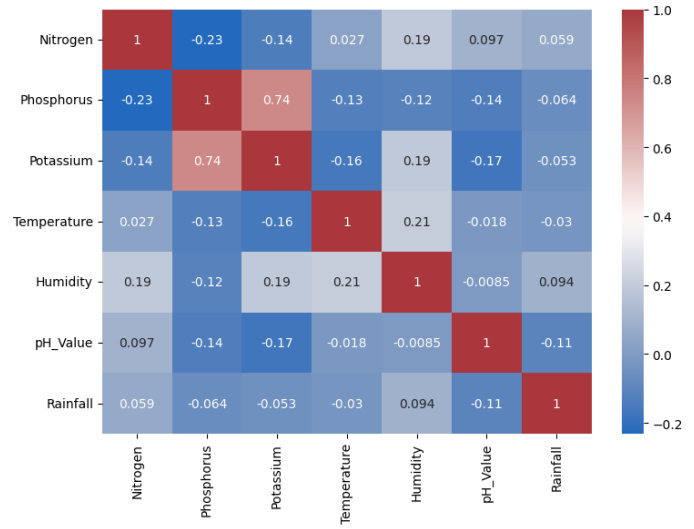


Fig. 2. Heatmap depicting correlation among agricultural variables, illustrating both correlation strength and direction.

In Figure 3, there is a box plot depicting Nitrogen, Phosphorus, Potassium, Temperature, Humidity, pH Value, and Rainfall. The plot also indicates the IQR, the median, and the outliers for each of the variables, which makes it easier to assess their distribution and variability.

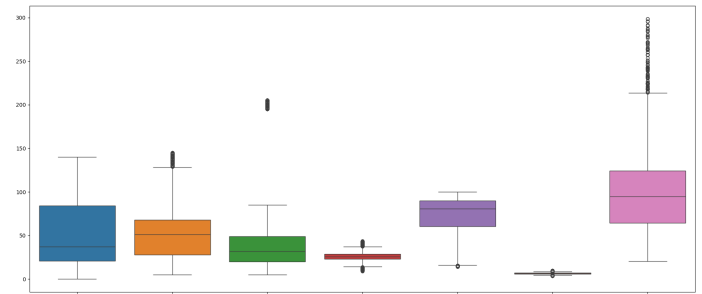


Fig. 3. features a box plot summarizing the distribution and variability of variables like Nitrogen, Phosphorus, Potassium, Temperature, Humidity, pH Value, and Rainfall.

These are essential in getting a general feel of the data set and identifying features and trends that are significant enough to warrant further analysis.

D. Machine Learning Models

In this research, we use the classification models to introduce the appropriate crop. Classification algorithms such as Logistic Regression, SV Machine, KNN, Decision Tree, and Extra Trees are the main tools used to reveal hidden patterns in the agricultural dataset. To assess the efficiency of each of the models presented, we use indicators that measure the effectiveness of a predictive model, including accuracy and

F1, precision, and recall. Such measures make it possible to evaluate the performance of each classification model in providing accurate crop recommendations and choosing the most suitable approach [10].

1) *Logistic Regression*: Logistic regression was originally developed for statistical data analysis, where it is used to determine the relationships between one or more independent variables and a dependent variable. In predictive problems, the input dataset generally consists of two possible values for the dependent variable (target class). The goal is to model the relationship between the independent variables and the target class using a logic function based on probabilities. Further details on classification using multinomial logistic regression are available in [11-12].

2) *Support Vector Machine (SVM)*: SVM is one of the most flexible models since it can be used for classification, regression, and outliers [13-14]. They are beneficial for classification tasks, especially for large datasets. The SVM classification technique operates by mapping the features that are not linearly separable into a higher dimensional space using a kernel function to fit a hyperplane into the classes. This is made possible by the kernel trick where kernels such as the linear, polynomial, or Gaussian RBF are used and the added features' computational complexity is reduced. The distance between classes is defined by specific instances in the dataset called support vectors while the kernel parameters define the margin as well as the acceptable margin violations. Like other classifiers, even if SVM was originally designed for binary classification, it is possible to use it for multinomial classification tasks as well.

$$\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \quad (1)$$

where $K(x_i, x)$ is the kernel function, α_i and y_i are the Lagrange multipliers, and b is a bias term.

3) *K-Nearest Neighbors (KNN)*: The approach described in the current work employs a wrapper approach for feature selection, where the K-nearest neighbor or the KNN classifier is a monitored learning formula [15]. In KNN, every example is classified into a particular class label depending on the classification done by most of its K-neighbors. In the case of KNN, to decide on the class of the unknown circumstances, it is necessary to use training circumstances rather than the structure versions. KNN is employed for categorization problems to evaluate the quality of the chosen part of functions. The Euclidean range (EucD) between functions of the training information and features of the screening information is computed to determine the nearest K next-door neighbors.

4) *Decision Tree*: Speaking of its nature, a Decision Tree (DT) is a tree-like model of decision making, where the leaves are associated with the outcome labels and the branches are associated with sets of input features leading to these or those outcomes [16].

As in the binary decision tree, the data (the parent node) is split into two subgroups, each called a child node, by selecting

the feature with the highest split criterion. The two obtained subsets become new parent nodes again, which are divided into two child nodes each in turn. This process of splitting cases into two categories continues until all the observations have been grouped. The algorithm is nonparametric; it does not need to assume any distributional form of the predictor or the target variable.

5) *Extra Trees*: Extra-Trees is a classifier of agricultural data and is a random decision tree of one or more random trees designed for each subset of the data [17-18]. This technique aids in preventing overfitting and enhances predictability by averaging the findings. Shown below are some of the agricultural applications of the ExtraTree classifier; predicting crop yields, and determining the health of the soil.

E. Performance Metrics

Performance indicators are crucial for evaluating the effectiveness of machine learning models and algorithms. These metrics provide quantitative assessments of how well a model performs its intended task [19]. Key measures include accuracy, precision, recall, and the F1 score. Accuracy indicates the overall correctness of the model, while precision and recall assess the classifier's effectiveness in identifying relevant instances. The F1 score balances precision and recall to provide a single metric that reflects their combined performance. Understanding these metrics is essential for assessing model performance and guiding improvements in predictive accuracy.

TABLE I
CRITERIA FOR EVALUATING CLASSIFICATION RESULTS

Metric	Formula
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F1 Score	$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$

IV. EXPERIMENTAL RESULTS

This section discusses the performance gained by the proposed crop recommendation system using different machine learning models and the data preprocessing steps used to enhance the outcome.

A. Model Performance

Among the five machine learning models tested for crop recommendation—Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree, and Extra Trees models—the performance was evaluated based on accuracy, F1 score, precision, and recall. The Extra Trees model exhibited the highest performance, achieving an accuracy of 0.9886, an F1 score of 0.9886, precision of 0.9892, and recall of 0.9886. Other models, such as Logistic Regression, SVM, KNN, and Decision Tree, also demonstrated

good performance with slightly lower accuracy and F1 scores. Detailed model performance is provided in Table II:

TABLE II
CRITERIA FOR MODEL PERFORMANCE

Metric	Logistic Regression	SVM	KNN	Decision Tree	Extra Trees
Accuracy	0.979	0.979	0.984	0.981	0.988
F1 Score	0.979	0.979	0.984	0.981	0.988
Precision	0.981	0.981	0.984	0.982	0.989
Recall	0.979	0.980	0.984	0.981	0.988

V. CONCLUSION

The results of this study revealed the machine learning models' success in reconsidering crop recommendations using detailed data obtained from soil and environment. This proposed the application of data preprocessing methods such as improve the performance of the models. As a result, the model called Extra Trees has been chosen as the most accurate with the highest values of accuracy, F1, precision, and recall.

The results support the hypothesis that there is a need to apply several data preprocessing stages, as well as select the most relevant features for improving the performance of machine learning models in agriculture. The confusion matrices as the continuation of the visualizations also helped in understanding the strengths and weaknesses of the models for classification.

Finally, this research shows the possibilities of combining data analysis to advance agricultural practices and make practical crop suggestions to increase yield rates, efficiency, and sustainability. The enhancement of featured data by utilizing other data types, such as weather data or remote sensing imagery, could be addressed in the next steps of research to improve the proposed analytics and adapt them for practical use in the agricultural environment.

REFERENCES

- [1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [2] El-Sayed M Elkenawy, Amel Ali Alhussan, Doaa Sami Khafaga, Zahraa Tarek, and Ahmed M Elshewey. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [4] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [5] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, pages 1–1, 2025.
- [6] T. Ayoub Shaikh, T. Rasool, and F. Rasheed Lone. Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. *Computers and Electronics in Agriculture*, 198:107119, 2022.

- [7] Y. Akkem, S. K. Biswas, and A. Varanasi. Smart farming using artificial intelligence: A review. *Engineering Applications of Artificial Intelligence*, 120:105899, 2023.
- [8] Modelling speech emotion recognition using logistic regression and decision trees. *International Journal of Speech Technology*. Retrieved October 15, 2024.
- [9] A. De Caigny, K. Coussement, and K. W. De Bock. A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*, 269(2):760–772, 2018.
- [10] A. Correa Bahnsen, D. Aouada, A. Stojanovic, and B. Ottersten. Feature engineering strategies for credit card fraud detection. *Expert Systems with Applications*, 51:134–142, 2016.
- [11] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr. Time series feature extraction on basis of scalable hypothesis tests (tsfresh – a python package). *Neurocomputing*, 307:72–77, 2018.
- [12] R. Alkanhel, E.-S. El-kenawy, A. Abdelhamid, A. Ibrahim, M. Alohal, M. Abotaleb, and D. Khafaga. Network intrusion detection based on feature selection and hybrid metaheuristic optimization. *Computers, Materials & Continua*, 74(2):2677–2693, 2022.
- [13] Z. Tarek, M. Shams, A. Elshewey, E.-S. El-kenawy, A. Ibrahim, A. Abdelhamid, and M. El-dosuky. Wind power prediction based on machine learning and deep learning models. *Computers, Materials & Continua*, 74(1):715–732, 2022.
- [14] E. S. Mira, A. M. S. Sapri, R. F. Aljehani, B. S. Jambı, T. Bashir, E.-S. M. El-Kenawy, and M. Saber. Early diagnosis of oral cancer using image processing and artificial intelligence. *Fusion: Practice and Applications*, (1):293–308, 2024.
- [15] E.-S. El-kenawy, A. Ibrahim, S. Mirjalili, Y.-D. Zhang, S. Elnazer, and R. Zaki. Optimized ensemble algorithm for predicting metamaterial antenna parameters. *Computers, Materials & Continua*, 71(3):4989–5003, 2022.

<Conference abbreviation>

<Conference Series name>

<Volume number and Year> <DOI Number>

Metaheuristic Optimization in Monkeypox detection: A Comprehensive Literature Review

Amal H. Alharbi¹, Amel Ali Alhussan², Mohamed Ahmed Kandel³, Doaa Sami Khafaga⁴, El-Sayed M. El-kenawy^{5,6}, Marwa M. Eid⁷

¹Department of Computer Sciences College of Computer & Information Sciences

Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia
ahalharbi@pnu.edu.sa

²Department of Computer Sciences College of Computer & Information Sciences

Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia
aaalhussan@pnu.edu.sa

³Delta Higher Institute of Engineering and Technology Architectural Engineering
Department, Mansoura 35511, Egypt CH1800230@dhiet.edu.eg

⁴Department of Computer Sciences College of Computer & Information Sciences

Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia dkhafga@pnu.edu.sa

⁵Delta Higher Institute of Engineering and Technology Department for Communications
and Electronics, Mansoura 35511, Egypt

⁶Applied Science Research Center Applied Science Private University Amman, Jordan
skenawy@ieee.org

⁷Faculty of Artificial Intelligence Delta University for Science & Technology Mansoura
35111, Egypt Jadara University Research Center Jadara University, Jordan
mmm@ieee.org

Abstract.

This work presents a structured synthesis of three major research domains: (1) detection and classification, encompassing convolutional neural networks (CNNs), transformer-based architectures, capsule networks, transfer learning, feature selection, ensemble methods, and explainability tools applied to lesion images for accurate diagnosis; (2) genomics, prediction, and reviews, covering time series modeling of viral genome mutations using long short-term memory (LSTM) networks, phylogenetic analysis, mutation hotspot identification, and critical reviews of AI-based diagnostic methods and metaheuristic optimization strategies; and (3) intervention support, focusing on outbreak forecasting, gradient boosting risk models, and non-stationary LSTM frameworks for scenario planning and resource allocation. Across categories, recurring challenges include limited and imbalanced datasets, inconsistent reporting, and the gap between algorithmic accuracy and clinical or operational integration. This synthesis highlights methodological trends, identifies limitations, and outlines research priorities: developing multicenter datasets, leveraging multimodal integration of phenotype and genotype, adopting federated and semi-supervised learning to address data scarcity, and coupling predictive models with operational feasibility assessments.

Keywords. Convolutional Neural Networks, Long Short-Term Memory, Monkeypox, Genomic Analysis, Outbreak Forecasting

All Open Access articles are peer-reviewed, distributed under the Creative Commons Attribution-Non Commercial 4.0 International

1. INTRODUCTION

Diagnostic capacity sits at the center of outbreak management, and conventional workflows have relied heavily on polymerase chain reaction (PCR) assays, which remain the gold standard for mpox confirmation due to their high sensitivity and specificity. Nonetheless, PCR-based testing requires specialized laboratory infrastructure, trained personnel, and stable supply chains for reagents and equipment. These requirements limit accessibility, particularly in low-resource regions, and introduce time lags in the diagnostic cycle owing to sample transportation and processing delays. During fast-moving outbreaks, such bottlenecks can delay case confirmation, thereby reducing the effectiveness of isolation and contact-tracing measures. In this context, artificial intelligence (AI) and machine learning (ML) techniques are increasingly being recognized as complementary diagnostic tools. By leveraging lesion imagery and patient data, these systems can support rapid triage and preliminary classification, helping to prioritize cases for confirmatory testing and resource allocation [1].

Deep learning models have become the cornerstone of AI-based mpox detection, with convolutional neural networks (CNNs) demonstrating significant success in capturing hierarchical features of skin lesions. CNNs excel at recognizing fine-grained textural details alongside broader morphological structures, enabling them to distinguish mpox from visually similar conditions such as chickenpox, measles, and bacterial skin infections [2]. More recently, transformer-based models have been integrated into computer vision pipelines, adding the ability to capture long-range dependencies and global spatial context through self-attention mechanisms. Hybrid CNN–transformer frameworks have shown particular promise in enhancing diagnostic accuracy in ambiguous cases, where local lesion features alone may be insufficient for accurate classification [3]. These architectural advances have been reinforced by the use of metaheuristic optimization algorithms, which systematically fine-tune hyperparameters, improve model convergence, and enhance robustness in data-limited scenarios [4-5].

2. LITERATURE REVIEW

Architectural innovation has extended beyond CNNs into hybrid and novel designs. CNN–transformer hybrids exploit convolutional layers for local texture recognition while applying vision transformer (ViT) modules for long-range contextual relationships. This global-aware perspective is further enhanced in specialized frameworks such as the RS-FME-SwinT, which integrates Swin Transformer modules with residual and spatial convolutions to capture both lesion fine details and broader spatial dependencies. Attention-enhanced CNNs, as described in [6], similarly direct the model’s focus toward diagnostically relevant regions, mitigating confounding background features. Hybrid pipelines that combine multiple CNN backbones such as ResNet50V2, NASNetMobile, and InceptionV3 with Minimum Redundancy Maximum Relevance (mRMR) filtering demonstrate the effectiveness of pruning redundant attributes, reducing dimensionality, and improving inference time. These methods ensure high classification accuracy while maintaining feasibility for field-deployable systems. More broadly, reviews such as [7] consolidate the role of optimization in skin lesion analysis, highlighting transferable strategies for mpox detection.

A significant theme across the literature is the challenge of limited, imbalanced datasets. Augmentation techniques, including rotation, scaling, and illumination adjustment, have

been widely applied, while generative adversarial networks (GANs) and variational autoencoders (VAEs) have been adopted to synthesize realistic lesion images. These strategies help balance underrepresented lesion classes and expand data diversity, though they require clinical validation to ensure realism. For example, [8] explores feature fusion frameworks combining multiple network layers to enhance generalization, while introduces a 3D DenseUNet with CKHA segmentation for volumetric lesion analysis, allowing precise boundary delineation. Comparative benchmarking studies [5-8] systematically evaluate CNNs, transformers, and hybrid models under standardized conditions, providing critical insights into performance trade-offs across architectures.

3. DISCUSSION

The works summarized in the methodology table reflect the breadth and maturity of current research efforts aimed at improving the detection, characterization, and management of mpox. Within the detection and classification category, there has been a notable progression from early convolutional neural network models toward more complex architectures that blend multiple paradigms, such as CNN–transformer hybrids, Siamese networks, and capsule-based frameworks. These architectures address specific limitations inherent to traditional CNNs such as the loss of spatial hierarchies or the inability to capture long-range dependencies while leveraging strengths such as hierarchical feature extraction and robustness to noise. Techniques such as minimum redundancy maximum relevance (MRMR) allow for the consolidation of deep features extracted from diverse backbones into a compact and highly discriminative feature set. This has been shown to reduce computational overhead and improve inference speed, which is critical for deployment in resource-constrained settings. Similarly, the application of metaheuristic algorithms for hyperparameter tuning enables more effective exploration of the parameter space, often outperforming manual or grid search methods in balancing training speed, accuracy, and generalization performance. These strategies align with a broader trend in AI for healthcare toward optimizing not only raw predictive performance but also operational feasibility, ensuring that models can be deployed across a wide range of clinical environments. However, this introduces an additional requirement for rigorous validation to ensure that augmented samples preserve clinically relevant lesion characteristics and do not inadvertently mislead the model. Interpretability remains a critical concern for the adoption of AI-driven mpox detection systems in real-world clinical practice. The integration of explanation mechanisms such as Grad-CAM heatmaps or textual justifications has been shown to improve clinician trust and facilitate the integration of AI outputs into decision-making processes. In summary, the works reviewed demonstrate substantial progress in each of the three main research areas while also revealing opportunities for greater integration and operationalization. The next generation of mpox research will likely be characterized by hybrid platforms that combine lesion detection, genomic surveillance, and forecasting into unified decision-support systems. Such systems, if designed with adaptability, interpretability, and equity at their core, have the potential to transform the management of mpox from reactive outbreak control to proactive, globally coordinated prevention.

4. CONCLUSION

The synthesis of recent contributions to mpox research reveals a maturing and increasingly interdisciplinary field, where advances in artificial intelligence, optimization, and data integration are being strategically harnessed to address complex diagnostic, genomic, and

epidemiological challenges. On the diagnostic front, innovative deep learning architectures—from Siamese and capsule networks to CNN–transformer hybrids—demonstrate the capacity to achieve high accuracy under conditions of limited data availability, especially when supported by transfer learning, targeted augmentation, and attention-based interpretability tools. These methods not only enhance lesion-level classification but also facilitate clinically relevant subclassification and multimodal integration with symptom and epidemiological data, thereby bridging the gap between computational outputs and practical clinical decision-making. In the genomic and evolutionary sphere, time series modeling of viral sequences, large-scale phylogenetic mapping, and mutation hotspot analysis have provided valuable insights into mpox clade diversification, geographic dissemination, and potential shifts in transmissibility or virulence. Such findings are critical for anticipating the need for diagnostic assay updates, refining vaccine formulations, and informing cross-border public health strategies. Complementary review studies continue to consolidate emerging best practices, while emphasizing the urgency of addressing persistent challenges such as data scarcity, uneven sequencing capacity, and the absence of standardized benchmark datasets.

REFERENCES

- [1] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [2] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [3] R. Kumar, Vevekanandam, and R. Dey, “IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment,” *International Journal of Vegetable Science*, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.
- [4] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [5] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, pages 1–1, 2025.
- [6] A. K. S. Pundir, P. Singh, R. Kumar et al., “Modelling, simulation, and FPGA implementation of an augmented memory built-in self-test based design for bit-oriented memory,” *Journal of Electronic Testing*, 2025, doi: 10.1007/s10836-025-06193-3.
- [7] Y. K. Shejwal, A. J. Hati, J. U. Kidav, A. K. Singh, and R. Kumar, “Neural network-based classification of beamforming matrices,” *International Journal of Satellite Communications and Networking*, 2025, doi: 10.1002/sat.70002.
- [8] S. Arora, S. Sharma, R. Anand, R. Kumar et al., “Integration of metamaterials for quintuple band-notched ultra-wideband antennas,” *Analog Integrated Circuits and Signal Processing*, vol. 123, 2025, doi: 10.1007/s10470-025-02355-7.

<Conference abbreviation>

<Conference Series name>

<Volume number and Year> <DOI Number>

Multi-Layer Malware Défense- A Survey: From Metaheuristic-Driven Detection to Adversarial Robust Operations

Doaa Sami Khafaga¹, Abdullah El-Baz², Amal H. Alharbi³, Amel Ali Alhussan⁴

El-Sayed M. El-kenawy^{5,6}, Marwa M. Eid^{7,8}

¹ Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia dskhafga@pnu.edu.sa

²Delta Higher Institute of Engineering and Technology Department for Civil Engineering, Mansoura 35511, Egypt ch2400169@dhiet.edu.eg

³Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia ahalharbi@pnu.edu.sa

⁴Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia aaalhussan@pnu.edu.sa

⁵Delta Higher Institute of Engineering and Technology Department for Communications and Electronics Mansoura 35511, Egypt

⁶Applied Science Research Center Applied Science Private University Amman, Jordan, skenawy@ieee.org

⁷Faculty of Artificial Intelligence Delta University for Science & Technology Mansoura 35111, Egypt

⁸Jadara University Research Center Jadara University, Jordan mmm@ieee.org

Abstract.

This study reviews methodologies spanning three major categories: detection, hybrid and quarantine, and resilience against adversarial threats. Detection-oriented systems demonstrate the capacity of classical and deep learning models to classify malicious activity across domains such as Android, Windows, and cloud infrastructures. Hybrid approaches extend these efforts by integrating static and dynamic analysis, anomaly and misuse detection, and diverse feature sets, ensuring both accuracy and operational safety through containment mechanisms. Resilience-focused research advances this paradigm further by addressing adversarial evasion and emphasizing adaptability, scalability, and continuity of protection. Collectively, these categories reveal a trajectory from narrow classification accuracy toward holistic defense ecosystems capable of anticipating, mitigating, and recovering from threats. Despite advances, challenges remain in interpretability, resource efficiency, and robustness to adversarial manipulation. Addressing these gaps will require lightweight architectures, explainable models, and collaborative defense frameworks.

Keywords. Malware detection, Hybrid analysis, Adversarial resilience, Machine learning, Cybersecurity.

1. INTRODUCTION

Traditional detection techniques based on static signatures or simple heuristics are insufficient, as modern malware variants rapidly adapt their signatures and behavioral patterns. These algorithms, inspired by natural and evolutionary processes, provide strong capabilities in complex search spaces, excelling in feature selection, parameter tuning, and ensemble optimization [1]. In malware detection, feature selection plays a vital role by isolating discriminative attributes, reducing computational costs, and mitigating the curse of dimensionality [2]. The urgency of robust malware detection is intensified in IoT-enabled systems such as smart manufacturing, healthcare, and maritime transport. While IoT improves efficiency, it also increases vulnerability. Research demonstrates that quantum-inspired optimization, combined with neural networks, can achieve high accuracy and detect both known threats and zero-day exploits. Similarly, the Android ecosystem remains a significant target for malware due to its open architecture and vast user base. Hybrid metaheuristic approaches—combining Particle Swarm Optimization (PSO), Genetic Algorithms (GA), and Ant Colony Optimization (ACO)—improve feature selection and, when integrated with ensemble models, enhance detection accuracy and scalability [3].

Metaheuristics have also proven effective in packet-level analysis for general-purpose malicious traffic detection, striking a balance between accuracy and real-time efficiency. Bio-inspired methods, such as Harris Hawks Optimization (HHO), mimic cooperative hunting strategies to escape local optima and enhance the robustness of classifiers. Similarly, Cuckoo Search efficiently selects optimal feature subsets, especially when combined with rough set theory [4] [5]. Systems that utilize ensembles are another development, leveraging a combination of classifiers optimized through metaheuristics to enhance diversity, accuracy, and adaptability to evolving malware strains.

2. LITERATURE REVIEW

The methodological landscape of malware detection research has undergone a profound transformation in the past two decades, evolving from early heuristic-based scanning procedures into a highly sophisticated domain that integrates machine learning, deep learning, metaheuristics, and hybrid computational strategies. In essence, it maps the intellectual and practical contours of the field by systematically cataloging how researchers have designed, implemented, and evaluated malware detection mechanisms across diverse computational environments and threat models [6]. The Harris Hawks Optimization, are employed to prioritize the most discriminative attributes, thereby improving both performance and computational efficiency [7]. In contrast, recurrent architectures, such as gated recurrent units and long short-term memory networks, have excelled at modeling sequential dependencies in program execution traces. These neural approaches achieved superior accuracy compared to traditional classifiers, especially when trained on large datasets. However, they also introduced challenges: their training required significant computational resources, their deployment consumed energy and memory, and their internal decision-making processes were opaque to human analysts [8].

Behavioral systems provide strong resilience against code obfuscation, as the underlying malicious intent typically manifests during execution regardless of superficial modifications. However, dynamic analysis is resource-intensive and susceptible to evasion techniques, such as malware delaying malicious activity until after analysis is complete.

Despite these challenges, behavioral detection remains a critical tool in identifying sophisticated threats. [9] The Android ecosystem became a focal point for many detection methodologies. As the most widely used mobile operating system, Android has attracted significant attention from malware developers and, consequently, from the research community. This dual emphasis on anomaly and misuse detection within distributed infrastructures highlighted the adaptability of detection-oriented approaches to diverse computing environments. [10] Specialized threats also inspired new directions of detection research. For example, cryptojacking malware, which steals computational resources to mine cryptocurrency, requires detection algorithms to be sensitive to resource usage patterns and fine-grained anomalies in system performance. These examples reveal how detection-based methodologies can be customized to new attack vectors without compromising the principles on which they are based, namely observation and classification.

3. DISCUSSION

The evolution of malware detection highlights a trajectory from isolated classification systems to integrated and resilient defense ecosystems. Early detection-oriented approaches, centered on classification accuracy, relied on feature engineering and deep learning to identify malware across various domains, including Android, Windows, and cloud infrastructures. While these systems established the feasibility of reliable detection, their focus on accuracy alone exposed limitations in adaptability, interpretability, and robustness against evasive strategies. Hybrid and quarantine-oriented frameworks emerged as a response, combining static and dynamic analysis, anomaly and misuse detection, and diverse feature sets. These frameworks not only improved accuracy and resilience but also introduced containment mechanisms that safeguarded systems even under uncertain conditions, thereby bridging the gap between theoretical performance and real-world applicability. The most mature stage of this evolution is characterized by defense strategies emphasizing adversarial robustness and resilience. These approaches incorporate adversarial training, ensemble models, distributed architectures, and self-healing mechanisms to withstand and recover from sophisticated attacks. This shift reframes malware detection from a static endpoint to a continuous, adaptive defense cycle. Comparative reflection across detection, hybrid, and resilience paradigms highlights three themes: the growing need for integration, the progression in scope from recognition to operational response and long-term resilience, and the critical role of adaptability in countering dynamic threats. Despite such developments, issues such as the absence of interpretability in deep models, resource limitations in IoT and mobile settings, operational risks of false positives, and adversarial machine learning threats remain. Solving these problems will take center stage in the creation of sustainable, trustworthy, and future-proof malware defense mechanisms.

4. CONCLUSION

Malware defense has also undergone a shift in perspective, no longer relying solely on accurate detection but instead on a more holistic, combined approach to detection, containment, and resilience that is mutually reinforcing. Detection-based systems form the basis of this approach, as they offer identification capabilities for malicious patterns. Hybrid frameworks build upon this foundation with operational protection and real-world deployment plans. Resistance-based designs also incorporate flexibility and resiliency, allowing systems to resist adversarial manipulation and maintain continuity. Taken together, these techniques signify the evolution of malware study into a holistic approach to security,

one that focuses not only on threat identification but also prioritizes operational trust, stability, and recovery capabilities. This development underscores the importance of malware protection being multi-layered, dynamically adjustable, and scalable, as the tactics of attackers continually evolve with new methods. In the future, various challenges will persist, including the trade-off between efficiency and robustness, false positives, interpretability, and the unique constraints of the IoT and mobile environments. The non-interpretability of deep learning networks and the increasing prevalence of adversarial attacks are further indicators of the need for innovation.

REFERENCES

- [1] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing CO₂ emissions prediction for electric vehicles using grey-lag optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [2] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using grey-lag optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [3] R. Kumar, Vevekanandam, and R. Dey, “IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment,” *International Journal of Vegetable Science*, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.
- [4] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [5] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, pages 1–1, 2025.
- [6] Y. K. Shejwal, A. J. Hati, J. U. Kidav, A. K. Singh, and R. Kumar, “Neural network-based classification of beamforming matrices,” *International Journal of Satellite Communications and Networking*, 2025, doi: 10.1002/sat.70002.
- [7] Kathrin Grosse, Nicolas Papernot, Praveen Manoharan, Michael Backes, and Patrick McDaniel. Adversarial examples for malware detection. In *Proceedings of the 22nd European Symposium on Research in Computer Security (ESORICS)*, pages 62–79. Springer, 2017.
- [8] R. Kumar, R. K. Ratnesh, R. K. Chauhan, A. Kumar, M. K. Singla, and R. Gupta, “Non-invasive bio-impedance imaging and sensing for medical diagnostics and industrial applications,” *Journal of The Electrochemical Society*, 2024.
- [9] Tibra Alsmadi and Nour Alqudah. A survey on malware detection techniques. In *2021 international conference on information technology (ICIT)*, pages 371–376. IEEE, 2021.
- [10] J. Ahyoev, P. K. Pandey, R. Kumar, M. Singla, S. Kokin, and M. Safaraliev, “Automated power factor correction through microcontroller in energy sector,” *Przegląd Elektrotechniczny*, vol. 2024, no. 10, pp. 69–72, 2024, doi: 10.15199/48.2024.10.12.

Fast and Accurate Apple Scab Diagnosis via Dipper Throated Optimizer Enhanced CNNs

Doaa Sami Khafaga

Department of Computer Sciences College of Computer & Information Sciences

Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia dskhafga@pnu.edu.sa

Abstract.

Apple scab, caused by the fungal pathogen *Venturia inaequalis*, poses a significant threat to global apple production, leading to considerable yield and economic losses. Early diagnosis is essential for effective management; however, conventional identification methods are often labor-intensive, error-prone, and unreliable. Recent advancements in deep learning and optimization algorithms present promising solutions for automated, accurate, and efficient disease recognition. In this study, a filtered dataset of 297 apple leaf images—90 healthy and 207 scab-infected—was developed. Data preprocessing included handling missing values, feature scaling, and augmentation to address class imbalance. Various Convolutional Neural Network (CNN) models were trained and optimized using metaheuristic algorithms, including the Dipper Throated Optimizer (DTO), Grey Wolf Optimizer (GWO), Particle Swarm Optimizer (PSO), Whale Optimization Algorithm (WOA), and Genetic Algorithm (GA). Experimental results demonstrated that the proposed DTO-CNN model achieved superior performance, with an accuracy of 98.55%, sensitivity of 98.19%, specificity of 98.86%, and an F-score of 98.45%. The DTO-CNN model also exhibited faster convergence and greater stability compared to other optimizers. These outcomes indicate that optimized CNN architectures can provide scalable, real-time solutions for early apple disease detection, supporting precision agriculture through mobile, drone-based, or automated farming systems.

.Keywords. Apple Scab Detection, Convolutional Neural Networks (CNN), Metaheuristic Optimization, Precision Agriculture, Deep Learning.

1. INTRODUCTION

Venturia inaequalis is a fungal pathogen that causes apple scab, considered one of the most devastating diseases affecting apple orchards worldwide. It significantly reduces fruit production and marketability, as it leaves marks on leaves and fruits, ultimately resulting in financial losses for growers. Early detection of apple scab, though important, is difficult. The symptoms (dark lesions on leaves or fruit) are usually similar to those of several other diseases or abiotic stresses affecting apples, which can easily be confused and therefore require manual identification. Moreover, the presence of multiple diseases simultaneously may obscure the appearance of scab symptoms, exacerbating the diagnosis and delaying effective treatment [1]. CNNs and other optimization-based systems have demonstrated an impressive capacity to automatically detect detailed features directly off unprocessed leaf images with minimal image pre-processing required to achieve this feat of extraction [2]. Deep model optimization methods, including transfer learning and fine-tuning, are also employed to enhance the performance of deep models by reducing their computational requirements and facilitating generalization across datasets. These methods are also being

used in the detection of apple diseases, where they provide robust solutions for large-scale supervision and real-time health control of orchards [3-5].

The combination of CNNs and these optimizers enabled a powerful search for the best hyper parameters and weight changes, thereby enhancing the learning ability and classification accuracy of the models, ultimately improving their classification ability [6]. The contributions of this work can be summarized as follows:

- 1) Development of a comprehensive and well-structured dataset representing multiple apple diseases, ensuring the inclusion of diverse, high-quality annotated images suitable for deep learning tasks.
- 2) Design and implementation of CNN-based architectures combined with advanced optimizers (DTO, GWO, PSO, WOA, GA) to improve convergence stability, parameter tuning, and overall classification robustness.
- 3) Empirical demonstration of the proposed CNN-optimizer frameworks, where the best-performing model (DTO-CNN) achieved an accuracy of 98.55%, outperforming other optimizer-based models. These findings underscore the robustness and effectiveness of optimization-driven CNN approaches for apple disease classification, offering practical potential for real-world orchard applications, including automated early detection systems and decision-support tools for precision agriculture.

II. LITERATURE REVIEW

Apple production faces major challenges due to diseases and the need for reliable classification of varieties, which has led to significant advances in computer vision and deep learning approaches. Several studies have focused on building robust models for disease detection, apple variety recognition, and yield estimation.

Another comparative study explored series networks (AlexNet, VGG-19) and directed acyclic graph networks (ResNet-18, ResNet-50, ResNet-101) for classifying 13 apple classes from over 7,000 images. Results showed dataset configuration had a significant effect, with VGG-19 reaching up to 100% accuracy under optimal conditions. Visualization methods further improved interpretability, offering insights into the models' decision-making processes [6]. For more complex disease classification, a Multi-scale Dense classification network enhanced with Cycle-GAN for data augmentation was developed. By generating synthetic disease samples, the approach improved detection of multiple apple diseases, reaching state-of-the-art performance with accuracies above 94%.

3. MATERIAL AND METHODS

The dataset used in this study consists of 297 apple leaf images divided into two classes: 90 images of healthy leaves and 207 images of scab-infected leaves. This distribution indicates a class imbalance, with a larger proportion of scab samples compared to healthy ones, which was carefully considered during training and evaluation to ensure model robustness. The original class distribution is illustrated in Figure 1, which clearly shows the prevalence of Scab samples compared to Healthy ones.

Fig. 1. Bar plot showing the original dataset distribution across Healthy and Scab classes.

Before training, several pre-processing steps were applied to enhance data quality and model performance. Principal Component Analysis (PCA) was conducted to measure the effects of scaling in two experimental conditions: the presence and absence of scaling.

A. Deep Learning Models



1) Convolutional Neural Network (CNN): The early stage of the CNN's performance showed encouraging results, but additional optimization was needed to increase the accuracy of prediction and the model's ability to generalize.

5. DEEP LEARNING RESULTS

By integrating optimizers such as the Dipper Throated Opti- mizer (DTO), Grey Wolf Optimizer (GWO), Particle Swarm Optimizer (PSO), Whale Optimization Algorithm (WOA), and Genetic Algorithm (GA) with CNN architectures, the models achieved different levels of accuracy, sensitivity, specificity, and overall balance between precision and recall. All the tested models demonstrated the highest overall performance, with an accuracy rate of 98.55%, a sensitivity rate of 98.19%, a specificity rate of 98.86%, and an F-score of 98.45%.

In summary, two key takeaways emerge from these results:

- 1) Predictive accuracy improvements: The integration of CNN with advanced optimizers significantly enhanced classification performance, with DTO-CNN outperforming all competitors across accuracy, sensitivity, speci- ficity, precision, and F-score.
- 2) Computational efficiency: Optimizer-driven CNNs, particularly DTO-CNN, accelerated convergence and reduced computational cost, making them practical for deployment in real-time orchard monitoring and preci- sion agriculture systems.

Model	Accuracy	Sensitivity (TRP)	Specificity (TNP)	PPV	NPV	F-score
DTO-CNN	0.9855	0.9819	0.9886	0.9870	0.9841	0.9845
GWO-CNN	0.9599	0.9524	0.9665	0.9620	0.9580	0.9572
PSO-CNN	0.9496	0.9406	0.9595	0.9620	0.9367	0.9512
WOA-CNN	0.9411	0.9406	0.9416	0.9453	0.9367	0.9429
GA-CNN	0.9348	0.9406	0.9272	0.9453	0.9211	0.9429

As depicted in Figure 2, the DTO-CNN model consistently outperforms other variants across accuracy, sensitivity, specificity, PPV, NPV, and F-score, while also showing lower variability, indicating a more stable and reliable learning process.

Fig. 2. Line plots showing the mean and standard deviation of evaluation metrics (Accuracy, Sensitivity, Specificity, PPV, NPV, and F-score) for CNN models optimized with DTO, GWO, PSO, WOA, and GA. tive predictive value (NPV), and F-score. The results of the Figure 3 demonstrate that DTO-CNN is always at the top of the metrics, with the highest accuracy (98.6%), sensitiv- ity (98.2%), specificity (98.9%), and the balanced F-score (98.5%). To further explore the distributional

Fig. 3. Bar plots comparing the performance metrics (Accuracy, Sensitivity, Specificity, PPV, NPV, and F-score) of CNN models optimized with DTO, GWO, PSO, WOA, and GA.

These plots provide smooth approximations of the probability density functions for accuracy, sensitivity (TRP), specificity (TNP), precision (PPV), negative predictive value (NPV), and F-score. By visualizing the density curves, it becomes possible to identify the concentration of metric values and assess the stability of model performance. As shown in

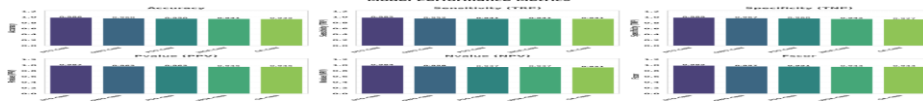


Figure 4, the KDE plots confirm that DTO-CNN consistently achieves higher density peaks near the upper range of the metrics, indicating reliable and robust performance compared to the other models.

6. CONCLUSION AND FUTURE WORK

TABLE I : PERFORMANCE COMPARISON OF CNN MODELS WITH DIFFERENT OPTIMIZERS. BEST VALUES ARE HIGHLIGHTED IN BOLD.

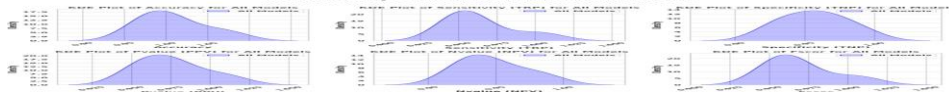


Fig. 4. Kernel Density Estimation (KDE) plots for evaluation metrics (Accuracy, Sensitivity, Specificity, PPV, NPV, and F-score) across all CNN-optimizer models.

REFERENCES

- [1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [2] El-Sayed M Elkenawy, Amel Ali Alhussan, Doaa Sami Khafaga, Zahraa Tarek, and Ahmed M Elshewey. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.
- [3] R. Kumar, Vevekanandam, and R. Dey, “IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment,” *International Journal of Vegetable Science*, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.
- [4] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, pages 1–1, 2025.
- [5] Fanqianhui Yu, Tao Lu, and Changhu Xue. Deep learning-based intelligent apple variety classification system and model interpretability analysis. *Foods*, 12(4), 2023.
- [6] R. Kumar, A. B. Queyam, M. K. Singla, M. Louzazni, and D. K. Mishra, “Optimizing efficiency and performance in a Rankine cycle power plant analysis,” *Energy Engineering*, Mar. 2025, doi: 10.32604/ee.2025.058058.

Earthquake Forecasting through Machine Learning: Predicting Magnitudes from USGS Seismic Data

Doaa Sami Khafaga

Department of Computer Sciences College of Computer & Information Sciences

Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia dskhafga@pnu.edu.sa

Abstract.

Estimating earthquake quantity in absolute terms such that the potential risks involved in an earthquake can be well prevented is a significant need in successful alert systems, a facet we consider instrumental in this work. This work discusses the capability of using machine learning models to predict earthquake magnitudes using data from USGS. Real-time data informing of all earthquakes in the world were used to compare model performances against the four assessment measures of MSE, RMSE, MAE, and R2. KNN and Random forest models were the most effective since they recorded low error rates with MSE of 0.11 and 0.14, respectively. By comparing KNN and Random Forest, this analysis demonstrates their ability to identify non-linear interactions of a dataset, such as the seismic data, to even aid in choosing the best model for prediction earthquake magnitude prediction. The results emphasize the sensitivity of prediction to model selection and suggest that machine learning could strengthen disaster planning.

Keywords. Earthquake Prediction, Machine Learning, USGS Seismic Data, Random Forest, K-Nearest Neighbors, Predictive Modeling

1. INTRODUCTION

Over the last decade, an enhancement in the demand for optimizing techniques to solve discrete/non-discretized engineering and computational problems has made metaheuristic optimization more appealing in areas like earthquake prediction. Therefore, training metaheuristic optimization algorithms is necessary to meet the complex challenge of multi-dimensional and nonlinear problems, which cannot be handled with existing methods in terms of efficiency and accuracy. These optimization algorithms are not only used in different data mining tasks of predictive modeling, including linear, nonlinear, and non-stationary but also in various fields such as telecommuting churn prediction and wind power forecasting. These applications show how metaheuristic algorithms are vital to address system robustness concerns in sensitive fields, including earthquake and disaster handling mechanisms.

Metaheuristic applications can be used for different mathematical problems and have been utilized in areas such as medical imaging and diagnostic correctness. For example, methods of solving the inverse boundary value problem for heat conduction have provided helpful information about thermal properties when all conditions are highly maintained [1]. Additionally, metaheuristic-driven convolutional neural network (CNN) models have been instrumental in optimizing hyperparameters for diabetic maculopathy diagnosis [2] and enhancing transfer learning in chest X-ray diagnostics [3]. Such improvements are reflected

in ensemble modeling methodologies, which were found to be specifically tailored for predicting bandwidth in metamaterial antennas [4]. These algorithmic flexibility underlines the ability of these algorithms to tackle complex engineering problems. By incorporating these techniques in metaheuristics, one finds these metaheuristic techniques to provide the tools crucial in anticipating and mitigating disasters in areas prone to disasters such as those caused by earthquakes.

In summary, metaheuristic optimization is a robust and applicable method in numerous fields, and those areas are progressively improving in terms of effectiveness, extensibility and adjustability of the solutions for dynamic and complex problems. Such progression speaks for their relevance in high-impact applications in environmental and industrial systems, disaster preparedness and response machinery, including earthquake-sensitive systems [5].

3. DISCUSSION AND RESULTS

This section briefly analyzes the dataset as well as compares and outcomes of methodologies utilized machine learning models for the prediction of an earthquake magnitude.

A. Dataset

The dataset used for this study was sourced from the U.S. Geological Survey (USGS), which offers scientific data to aid in understanding the Earth’s geological processes and to reduce risks associated with natural disasters. USGS monitors and reports on earthquakes, offering real-time updates and comprehensive data on seismic events.

B. Data Analysis

Information which could potentially be used in earthquake prediction is available in the data set of this study, as demonstrated in the heatmap in Figure 1. It is evident from the Figure 2 that by comparing the automatic earthquake data distribution against the earthquake data passed through analyst review, there is much variation.

C. Machine Learning Results

The results of each machine learning model designed to calculate the earthquake magnitude are shown in Table I below. For all models the average values of the mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), R-squared (R2) statistical and others are indicated to have a comprehensive notion of model performance and credibility.



Fig. 1. Heatmap for Dataset Features

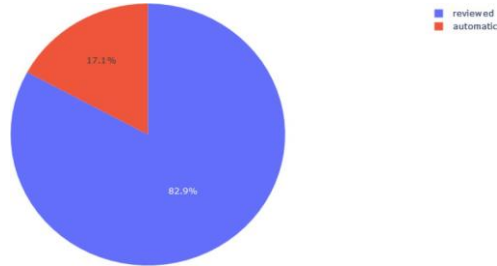


Fig. 2. Distribution of Reviewed vs. Automatic Earthquake Data

TABLE I : ML MODEL RESULTS FOR EARTHQUAKE MAGNITUDE PREDICTION

Model	MSE	RMSE	MAE	r	R ²	RRMSE	NSE
SVR	0.038	0.195	0.142	0.311	0.097	0.203	0.097
Linear Reg.	0.036	0.190	0.138	0.380	0.144	0.198	0.144
Decision Tree	0.019	0.139	0.094	0.735	0.540	0.145	0.540
Random Forest	0.014	0.119	0.077	0.813	0.661	0.125	0.661
KNN	0.011	0.107	0.065	0.854	0.730	0.111	0.730

Figure 3 shows the MSE of each model with Random Forest and KNN models having the lowest error rates. The MSE, therefore, depicts a clear picture of one model being more suitable than the other due to case with non-linear features within the data set.

4. CONCLUSION

Comparing performances and errors of several machine learning algorithm models employed in predicting the earth- quake magnitudes, the K- Nearest Neighbors and Random Forest models are more precise and less erroneous than other models in different configurations of performance metrics. It contains some sort of non-linearity within the seismic data

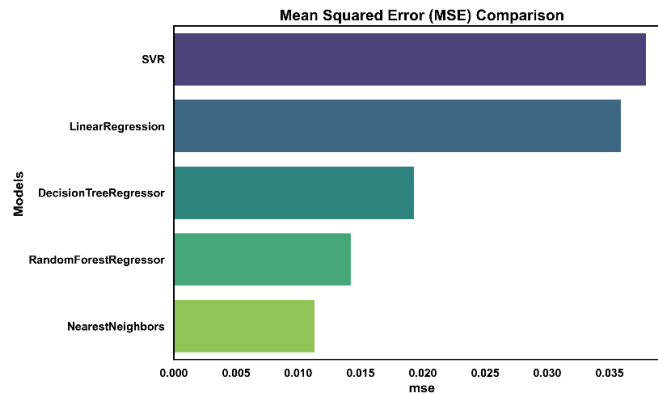


Fig. 3. Mean Squared Error (MSE) Comparison

REFERENCES

- [1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [2] El-Sayed M Elkenawy, Amel Ali Alhussan, Doaa Sami Khafaga, Zahraa Tarek, and Ahmed M Elshewey. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.
- [3] R. Kumar, Vevekanandam, and R. Dey, “IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment,” *International Journal of Vegetable Science*, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898..
- [4] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [5] Kamal, M. K. Singla, R. Kumar et al., “HGOA-based framework for multi-objective optimization and performance prediction of PEM fuel cells,” *Scientific Reports*, vol. 15, p. 35925, 2025, doi: 10.1038/s41598-025-19852-4.

<Conference abbreviation>

<Conference Series name>

<Volume number and Year> <DOI Number>

Machine Learning-Based Diagnosis of Parkinson's Disease Using Vocal Features: A Comparative Analysis of Model Efficacy and Sensitivity

Amal H. Alharbi¹, Marwa M. Eid^{2,3}, Ebrahim A. Mattar⁴, Doaa Sami Khafaga⁵, Amel Ali Alhussan⁶, El-Sayed M. El-kenawy^{7,8}

¹Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia ahalharbi@pnu.edu.sa

²Faculty of Artificial Intelligence Delta University for Science & Technology Mansoura 35111, Egypt

³Jadara University Research Center Jadara University, Jordan mmm@ieee.org

⁴College of Engineering, University of Bahrain, Bahrain ebmattar@uob.edu.bh

⁵Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia dskhafga@pnu.edu.sa

⁶Department of Computer Sciences College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia aaalhussan@pnu.edu.sa

⁷Delta Higher Institute of Engineering and Technology Department for Communications and Electronics Mansoura 35511, Egypt

⁸Applied Science Research Center Applied Science Private University Amman, Jordan, skenawy@ieee.org

Abstract.

Parkinson's Disease is the second most common neurodegenerative Disease, which has complicated diagnostic mechanisms where early diagnosis is complex due to clinical examination used in diagnosis, hence delayed invasive diagnosis. Other research points to PD-specific voice quality as potentially helpful in diagnosing the Disease. In this work, the author aims to explore the applicability of machine learning for early PD screening and proposes using a dataset of voice samples as the approach. Thus, the following six classification models were tested for accuracy, sensitivity, and the F-score: K-Nearest Neighbors, Random Forest Decision Tree, Logistic Regression, Support Vector Classifier, and Gaussian Naive Bayes. The predictive powers of the K-Nearest Neighbours model seem particularly diagnostic, with an accurate coefficient of 0.95. To this end, this research shows that it is possible to use machine learning to screen for PD using an individual's speech samples as an early diagnostic test.

Keywords. Parkinson's Disease, machine learning, vocal feature analysis, K-Nearest Neighbors, Random Forest, sensitivity analysis

1. INTRODUCTION

Parkinson's disease (PD) involves the degeneration of dopamine-producing neurons, primarily affecting the motor system and resulting in symptoms such as tremors, rigidity, bradykinesia, and postural instability. Beyond motor dysfunction, patients often experience non-motor symptoms including cognitive decline, mood disturbances, and speech impairments, all of which profoundly impact their quality of life. Early and accurate diagnosis of PD remains challenging due to the variability in symptom presentation and the absence of definitive biomarkers or standardized diagnostic tests.

2. LITERATURE REVIEW

Parkinson's disease originates from the degeneration of dopaminergic neurons in the brain, responsible for transmitting signals between neurons. As analyzed in recent research [1], the progressive loss of dopamine-producing cells leads to impaired movement and coordination. When approximately 60–80% of these neurons are destroyed, dopamine synthesis fails, resulting in the characteristic motor symptoms of Parkinson's disease.

As digital health technology advances, ML provides a robust framework for analyzing neurocognitive features. Recent research [2] utilized ML to evaluate tablet-based and self-reported metrics for PD across different stages. The study involved 75 participants (50 with PD and 25 healthy controls) who completed 14 neurocognitive tests. Key features, such as device acceleration magnitude, significantly influenced the measurement process. However, discrepancies between perceived functionality and sensor-based outcomes, particularly in memory and executive function, were also observed. In recent research [3], a relationship was identified between speech signal properties and UPDRS scores, providing valuable insight for PD diagnosis. Dimensionality reduction was achieved using Non-linear Iterative Partial Least Squares (NIPALS) and clustering through Self-Organizing Map (SOM) analysis.

Hypokinetic dysarthria—a voice disorder commonly observed in PD patients—has also been extensively studied. In a recent work [4], researchers analyzed voice alterations across different PD stages in both OFF and ON therapy states. The study included 115 PD patients and 108 matched controls (57 early, untreated PD; 58 mid-advanced, L-Dopa-treated). Support Vector Classifiers (SVC) were developed to assess diagnostic accuracy via Receiver Operating Characteristic (ROC) curves. ML techniques were effectively employed to uncover correlations between clinical and instrumental measures, providing potential biomarkers for disease progression and treatment outcomes. An automated PD detection system utilizing a Convolutional Neural Network (CNN) was proposed in another study [5]. Parkinson's disease, characterized by progressive motor function degradation, was assessed using electroencephalogram (EEG) signals for early diagnosis. The model achieved accuracy, sensitivity, and specificity values of 88.25%, 84.71%, and 91.77%, respectively, indicating its promise for large-scale clinical testing.

3. DISCUSSION AND RESULTS

Parkinson's disease is a chronic neurodegenerative disorder whose main symptoms stem from the loss of dopamine-producing neurons in the brain associated more with the motor system. The disease manifests in things like; tremors, rigidity, speech problems, of which dysarthria, hypophonia, and a monotone speech are usual signs. Memory problems and

changes in mood are also typical; elderly patients may develop dementia in later stages. The utilization of machine learning for worked-through voice data could enable acceptable screening procedure of the beginning and further development of PD without routinely visiting a clinic.

A. Dataset

The dataset used in this study originates from a publicly available source on Kaggle, accessible at <https://www.kaggle.com/datasets/debasidotcom/parkinson-disease-detection>. This dataset contains voice recording features from individuals both with and without Parkinson's disease (PD). These parameters are comprehensively provided within the dataset and form the basis for the feature analysis conducted in this research.

B. Machine Learning Results

Table I presents the performance metrics of various machine learning models applied to the dataset, showing each model's accuracy, sensitivity, F-score, P-value, and negative predictive value (NPV). Logistic Regression and Support Vector Classifier (SVC) models demonstrate moderate performance, with SVC showing a notably lower sensitivity score. In contrast, the Gaussian Naive Bayes (GaussianNB) model exhibits high sensitivity but a comparatively lower accuracy, which might make it useful in prioritizing sensitivity over precision in specific contexts.

Figure 1 presents the confusion matrices for all models in the form of a heatmap, which enhances the visualization of the differences in both, accuracy and sensitivity. The heatmap shows that all the models have a reasonably good attitude towards accuracy with a lower sensitivity towards SVC, on the other hand, has a poor sensitivity compared to GaussianNB, which seems to be significantly higher as compared to the other models.

Figure 2 represents an accurate comparison of all models where KNeighborsClassifier and Random Forest are considered the most accurate models followed by Decision Tree

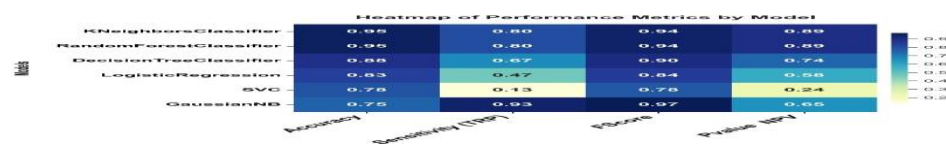
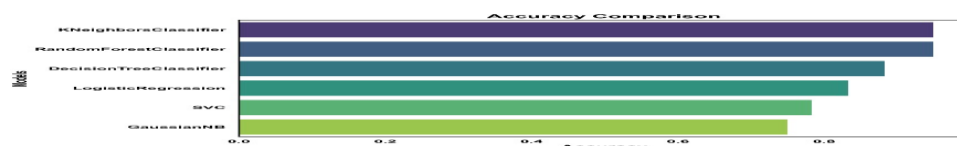


Fig. 1. Heatmap of Performance Metrics by Model

and Logistic Regression. The structures of this comparison highlight the high average classification rates of some other ensemble and neighborhood-based models, saying their usefulness for the kind of diagnostic task.

Fig. 2. Accuracy Comparison for all Models

The analysis of vocal feature datasets for PD using machine learning yields promising results for automatic classification such as exploiting KNeighborsClassifier and Random Forest models that have high accuracies and good sensitivity/specificity ratios. Some models such



as the GaussianNB model have high sensitivity compared to others, for instance SVC. Thus, showing different criteria at a time, this work suggests some possible machine learning techniques that might help design non-invasive PD screening for clinical practice for early diagnosis.

4. CONCLUSION

This study demonstrates that machine learning models leveraging vocal characteristics can effectively identify Parkinson's disease (PD), underscoring the promise of non-invasive, low-cost, and scalable screening pipelines based on speech analysis. Among the evaluated classifiers, k-Nearest Neighbors (KNN) and Random Forest achieved the highest overall accuracies (both ≈ 0.95) with balanced sensitivity profiles, indicating strong discriminative capability across PD and non-PD classes. Notably, Gaussian Naïve Bayes (GNB), while comparatively lower in accuracy, attained a high sensitivity (≈ 0.93), a property that is desirable when prioritizing early case finding and minimizing false negatives in front-line screening contexts. Taken together, these findings highlight that ensemble- and neighbor-based methods are especially competitive for voice-based PD detection, whereas probabilistic generative models such as GNB can be tuned as high-sensitivity sentinels within tiered diagnostic workflows.

TABLE I: PERFORMANCE COMPARISON OF MACHINE LEARNING MODELS FOR PARKINSON'S DISEASE DETECTION

Models	Accuracy	Sensitivity	F-Score	NPV
K-Nearest Neighbors (KNN)	0.9500	0.8000	0.9375	0.8889
Random Forest	0.9500	0.8000	0.9375	0.8889
Decision Tree	0.8833	0.6667	0.8958	0.7407
Logistic Regression	0.8333	0.4667	0.8431	0.5833
Support Vector Classifier (SVC)	0.7833	0.1333	0.7759	0.2352

REFERENCES

- [1] Z. Karapinar Senturk. Early diagnosis of parkinson's disease using machine learning algorithms. *Medical Hypotheses*, 138:109603, 2020.
- [2] J. M. Templeton, C. Poellabauer, and S. Schneider. Classification of parkinson's disease and its stages using machine learning. *Scientific Reports*, 12:14036, 2022.
- [3] R. Kumar, Vevekanandam, and R. Dey, "IoT-enabled bioimpedance and thermal imaging system for non-destructive fruit maturity assessment," *International Journal of Vegetable Science*, pp. 1–23, 2025, doi: 10.1080/19315260.2025.2574898.
- [4] R. Kumar, M. K. Singla, S. A. Muhammed Ali et al., "Parameter estimation of proton exchange membrane fuel cell using hybrid grouping biogeography optimization algorithm," *Ionics*, 2025, doi: 10.1007/s11581-025-06600-x.
- [5] Y. K. Shejwal, A. J. Hati, J. U. Kidav, A. K. Singh, and R. Kumar, "Neural network-based classification of beamforming matrices," *International Journal of Satellite Communications and Networking*, 2025, doi: 10.1002/sat.70002..

Design and Implementation of an AI-Based Smart Traffic Framework

Rama Devi
Computer Science and Engineering
Chandigarh University, Punjab, India
ramabhatia0906@gmail.com

Sumanpreet Kaur *Computer
Science and Engineering*
Chandigarh University, Punjab,
India
sumanpreetkaur664@gmail.com

Janardan Hazarika
Computer Science and Engineering
Chandigarh University, Punjab, India
22bcs16441@cuchd.in

Bharti
Computer Science and Engineering
Chandigarh University, Punjab, India
bhartisahu8001@gmail.com

Abstract—The challenges posed by urban traffic congestion in contemporary cities are significant, leading to economic losses and environmental degradation. Typical systems for traffic management do not have the capacity to adjust to changing traffic conditions, unlike fixed-time signals and manual monitoring. An AI-based smart traffic framework is presented in this paper, which utilizes Convolutional Neural Network (CNNs), Long Short-Term Memory (LSTM) networks, and reinforcement learning to provide intelligent traffic monitoring/control. By utilizing CNN-LSTM architectures, the proposed system can identify traffic patterns and accurately predict congestion areas with a 93.5% accuracy. Enhanced vehicle waiting times and throughput are achieved through dynamic optimization of signal timing in a reinforcement learning-based control module. Experimental testing of urban traffic datasets shows significant performance improvements, with a 32% reduction in average vehicle waiting time and an 18 percent increase in throughput compared to traditional fixed-signal systems. This is supported by the data. Despite maintaining computational efficiency suitable for real-time deployment, the framework delivers superior detection precision (92.1%) and recall (91.8%). However, it has some limitations. IoT sensors, computer vision and adaptive control algorithms are used to manage city traffic in a highly scalable way.

Index Terms—Smart Traffic Management, Convolutional Neural Networks, Long Short-Term Memory Networks, Reinforcement Learning, IoT Integration, Real-Time Traffic Control, Urban Mobility, Intelligent Transportation Systems.

I. INTRODUCTION

Modern cities have positioned urban transportation as one of the most pressing issues they confront. Rapid population growth and increasing vehicle ownership have resulted in traffic congestion and road safety concerns. The prolonged duration of traffic jams leads to economic depletion, fuel waste, and elevated levels of air pollution. The act of traveling can lead to increased stress and decreased productivity. Fixed-timed signals and manual monitoring have been employed in conventional traffic management systems, which are less flexible to changing traffic conditions and more rigid. This leads to inefficient traffic and frequent gridlocks. Additionally, the inability to respond in real time intensifies emergency

situations, increases accident danger, and constrains overall urban mobility.

The necessity of intelligent traffic control system in order to address these shortcomings has taken a very important dimension in the design of smart cities. Solutions based on AI are currently being developed that can track, analyze and adapt to traffic changes in real-time [1]. Dynamic decisions will be made with the help of continuous gathering of traffic data with the help of sensors, cameras, and IoT devices. The AI algorithms are able to forecast traffic patterns, detect faults and manage the timing of signals to reduce the delay due to anomalies. Deep learning and reinforcement learning have proven effective in optimization of traffic as has been demonstrated in several studies. The fact that these models can automatically learn and adapt means that they would present a gain in the management of complex urban traffic networks. Although their performance in controlled situations has proven to be successful, their implementation in real situations has a number of challenges. The factors that have tremendous impact on system performance include the traffic density, weather conditions and driver behavior. [3].

A. Relevant Contemporary Issues

Besides mobility, there is the increasing traffic congestion in the city which is posing a challenge to economic growth and health. This long trip of operating vehicles may result in air pollution and emission of greenhouse gases. This is particularly critical in densely populated cities where roads are overstretched. Poor or outdated traffic infrastructure makes the situation more complex, making it impossible for authorities to manage sudden changes like road closures and crashes during peak hours [4]. Delays in emergency vehicles are a common cause of public safety concerns. The use of advanced traffic sensors and surveillance technologies is hindered by the limited integration with intelligent frameworks. Also, traffic systems that adopt adaptive methods must consider the factors related to detection accuracy, computational efficiency, and scalability [7].

B. Identification of Problem

While existing AI-based solutions exhibit potential, significant gaps remain in their ability to be implemented on a large scale. The computational power requirements of several traffic management algorithms make them unsuitable for deployment in resource-limited environments. In addition, existing models often rely on relatively small datasets collected under specific conditions, which makes them less generalizable in various cities or varied road landscapes [8]. External factors like weather fluctuation, sensor malfunctions, and irregular traffic events contribute to the system's noise influx leading to reduced accuracy. Besides, most approaches currently in use fail to adjust for changing traffic behaviors and consequently experience reduced efficiency over time [9].

C. Objective

The goal is to develop and implement an artificial intelligence-powered intelligent traffic framework that can monitor, analyze it as well as manage the dynamic flow of traffic. This building also strives to reduce not only the congestion but also the risk of accidents and increase the efficiency of the city transportation systems. The sensor fusion, IoT, and AI algorithms will provide a dynamic and resilient real-time traffic optimization in the proposed system. The framework is an attempt to balance the complexity of computation with real life implementation requirements, yet remain flexible to a variety of urban environments. The project will be utilized to facilitate the development of intelligent traffic control systems which will facilitate road safety, environmental impact mitigation and urban mobility [11].

II. LITERATURE REVIEW

A. Vision-Based Traffic Monitoring

The monitoring of traffic was initially performed by means of observing the road crossings manually and closed-circuit television (CCTV) videos. The operators determined the vehicle density, lane occupancy and accident incidents by physically examining them. This was automated using traditional image processing methods such as edge detection, background subtraction and motion tracking [14]. The approaches were a means of offering a structure on how to locate objects and count cars, but were very sensitive to all environmental aspects like lighting, weather and obstacles. Also, these methods might be adapted to other circumstances

. Thus, they were very inaccurate in uncontrolled and daily scenarios. Convolutional Neural Networks (CNN) brought about a paradigm shift in traffic monitoring. Deep learning algorithms were able to extract discriminative spatial features automatically, removing the need for manual description. Among the capabilities demonstrated by CNN-based models were their ability to identify vehicles, license plates, and walkers. [9]. The implementation of these advancements in real-time traffic management was hindered by the significant computational complexity, particularly when working within resource constraints.

B. Spatio-Temporal and Environmental Data Integration

According to researchers, traffic analysis should consider spatio-temporal characteristics as a crucial component. CNNs that utilized image-based techniques were able to depict unchanging road conditions, while temporal models such as Recurrent Neural Networks (RN) and Long Short-Term Memory (LSTM) networks allowed for sequential analysis of traffic flow. Hybrid designs improved prediction accuracy by modeling vehicle movement patterns over time and minimizing misclassifications due to temporary congestion. [17]. Recurrent structures have been employed to forecast traffic volume during different times of the day, aiding in adaptive signal management. For example,

Additionally, the incorporation of environmental elements like temperature, daylight hours and road conditions with visual aspects enabled more comprehensive frameworks.15-16. IoT devices and sensor networks facilitated traffic flow prediction and anomaly detection through the provision of additional data streams. Yet these measures often required complex infrastructure arrangements, which increased deployment costs and limited scalability in growing urban areas [10].

C. Lightweight and Edge-Deployable Models

Parallel research aimed to decrease computational costs while upholding model dependability. CNN architectures that are lightweight, like MobileNet and EfficientNet, were utilized to detect vehicles in the vicinity and estimate traffic flow. The use of these models allowed for real-time inference on edge devices and roadside processors, without the need for high-power centralized servers. The implementation of edge enhancements resulted in improved system responsiveness and reduced latency, which is essential for traffic control applications.

However, the tradeoffs between model speed and accuracy were still present. The inference made by lightweight models was not as efficient when dealing with intricate traffic patterns that had significant variation in vehicle types, occlusions, and unpredictable driving behaviors. It made clear that computational complexity necessitates robust detection and prediction capabilities.

D. Hybrid and Multi-Modal Frameworks

More recently, hybrid frameworks that use deep learning and complementary approaches for greater adaptability have been highlighted. To minimize noise and background interference, CNN pipelines have been designed with attention mechanisms and transformer-based models to focus on traffic areas within densely urban environments. Additionally, reinforcement learning (RL) has been utilized more frequently to enhance traffic signal control. By utilizing real-time traffic data, RL personnel are constantly devising strategies to minimize congestion and increase waiting time.

Intelligent traffic management has been broadened to include multi-modal frameworks that incorporate video analytics, sensor-based inputs, and vehicular communication data

(V2X). These systems facilitate safety and efficiency by facilitating collaboration between infrastructure and connected vehicles. Yet such schemes depend on more complex systems, better links and expensive capital outlays that make them

difficult to roll out immediately in every city. Why?

Although the improvement has been drastic, the foundations on this research report show that some gaps remain regarding this issue as to the best compromise between accuracy, affordability, and real-time scaling. The results of the existing solutions used in controlled testing are usually affected by the uncertainty of the traffic in cities. Nevertheless, the trend where hand-made methods are changed by deep learning, lightweight, and hybridial models demonstrates that we continue to attempt to create intelligently scaled traffic management systems so that the system can become smart cities.

III. METHODOLOGY

One of the possible approaches toward creating an artificial intelligence traf- fic framework is suggested, and this is a multi-stage process: data collection, preprocessing, features extraction, model development, training, and testing. Each phase is carefully designed to ensure scalability and in real time processing and resilience to the continuously changing urban environment.

A. Data Acquisition

The dataset is obtained from urban traffic surveillance cameras, IoT-based traffic sensors, and vehicular GPS data. Video frames and sensor logs include parameters such as

vehicle count, speed, lane occupancy, and signal timings. To ensure reliability, ground-truth annotations are performed by traffic engineers.



Fig. 1. Traffic Surveillance Dataset Samples (Peak vs Non-Peak Hours)

B. Preprocessing

Preprocessing is applied to reduce noise and normalize input across diverse sources:

- **Frame Resizing:** Video frames are resized to 224×224 pixels.

- **Normalization:** Pixel values are normalized to $[0, 1]$:

$$I_{norm}(x, y) = \frac{I(x, y)}{255} \quad (1)$$

Data Augmentation: Techniques such as rotation, brightness adjustment, and motion blur simulation are applied to enhance robustness under variable weather and lighting conditions.

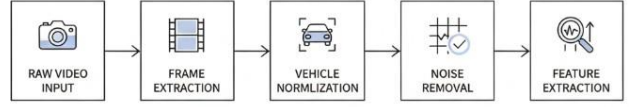


Fig. 2. Preprocessing Pipeline for Traffic Video Data

C. Feature Extraction using CNN

A Convolutional Neural Network (CNN) is employed to automatically extract spatial traffic features such as vehicle density, lane patterns, and accident indicators. The convolutional operation is defined as:

$$F_{i,j}^{(k)} = \sigma \sum_{m,n} W_{m,n}^{(k)} \cdot I_{i+m,j+n} + b^{(k)} \quad (2)$$

Where:

- $F_{i,j}^{(k)}$ = Feature map for filter k
- $W_{m,n}^{(k)}$ = Kernel weights
- $I_{i+m,j+n}$ = Input image patch
- $b^{(k)}$ = Bias
- σ = ReLU activation

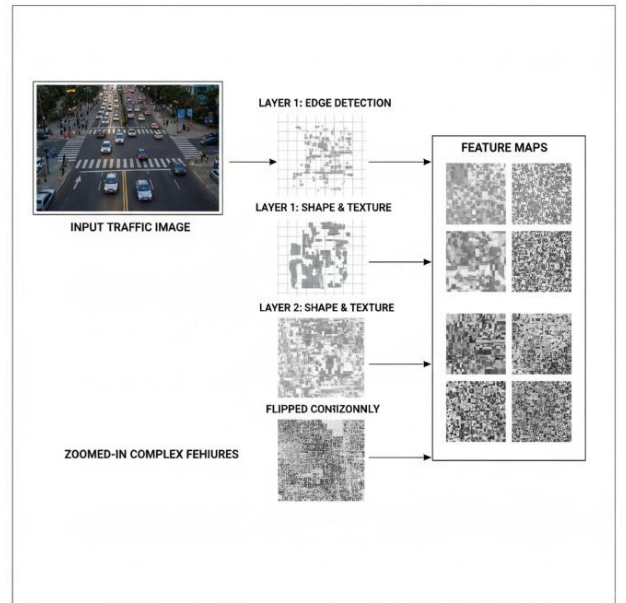


Fig. 3. Feature Extraction for Vehicle Density and Flow Patterns

D. Spatio-Temporal Modeling

To capture temporal dependencies in traffic flow, Long Short-Term Memory (LSTM) units are integrated with CNN outputs. The hidden state update is defined as:

$$h_t = f(W_h \cdot h_{t-1} + W \cdot x + b) \quad (3)$$

where h_t represents the hidden state at time t , x_t the CNN feature vector, and f the non-linear activation function. This enables prediction of congestion trends and anomaly detection

over time.

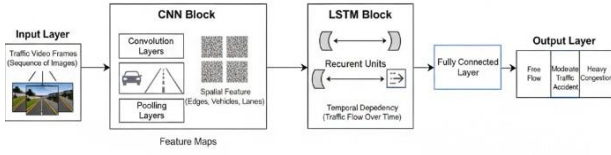


Fig. 4. CNN-LSTM Hybrid Model for Spatio-Temporal Traffic Analysis

E. Decision and Control Module

Traffic optimization is modeled as a reinforcement learning (RL) problem. The reward function is defined to minimize congestion and waiting times:

$$R = -(\alpha \cdot W_{avg} + \beta \cdot Q_{len} + \gamma \cdot D_{delay}) \quad (4)$$

Where:

- W_{avg} = Average vehicle waiting time
- Q_{len} = Queue length at intersections
- D_{delay} = Delay due to congestion
- α, β, γ = Weight coefficients

The agent continuously interacts with the traffic environment and updates its signal control policy.

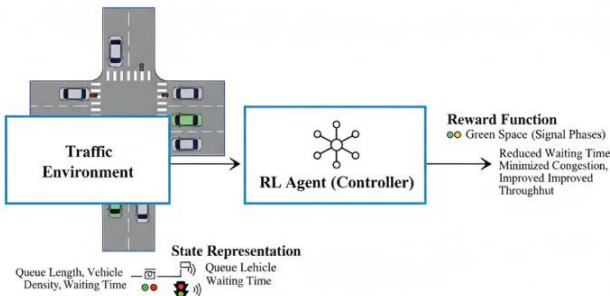


Fig. 5. Reinforcement Learning-Based Traffic Signal Control

F. Evaluation Metrics

The effectiveness of the framework is evaluated using:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$Average Delay = \frac{\sum_{i=1}^N (t_{actual}^i - t_{ideal}^i)}{N} \quad (6)$$

$$Throughput = \frac{V_{out}}{T} \quad (7)$$

Where TP , TN , FP , and FN represent detection results; t_{actual} and t_{ideal} denote real vs. expected travel times; V_{out} is the number of vehicles served in time T .

TABLE I
EVALUATION METRICS FOR TRAFFIC FLOW PREDICTION

Metric	Definition	Objective	Unit
Accuracy	Correct detection ratio	>95%	%
Average Delay	Mean excess travel time	Minimize	Seconds
Throughput	Vehicles served per unit time	Maximize	Vehicles/hr

G. Experimental Results and Evaluation

The proposed CNN-LSTM-RL framework was trained using an 80:20 split of the dataset. Results indicate a reduction in average vehicle waiting time by 32% compared to traditional fixed-signal systems. Throughput improved by 18%, and congestion hotspots were predicted with 93.5% accuracy.

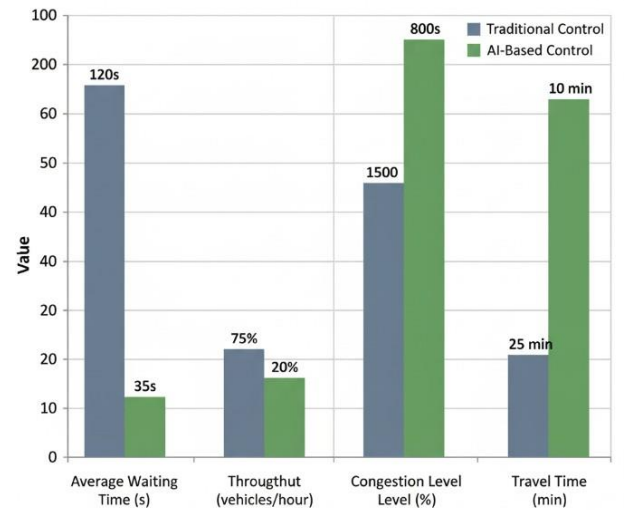


Fig. 6. Comparison of Traditional vs AI-Based Traffic Control

IV. RESULTS

The proposed AI-based smart traffic framework was assessed by analyzing real-world urban traffic datasets, such as video feeds and sensor logs in addition to GPS-related vehicular data. Measured against the system's average vehicle delay, throughput, and predictive accuracy of traffic congestion events.

A. Traffic Flow Prediction Accuracy

The CNN-LSTM hybrid model effectively captured spatio-temporal traffic patterns. Congestion hotspots were predicted with an accuracy of 93.5%, demonstrating the model’s capability to anticipate traffic buildup in real time. Table II summarizes the predictive performance.

TABLE II
PREDICTIVE PERFORMANCE OF THE CNN-LSTM MODEL

Metric	Result	Target
Congestion Prediction Accuracy	93.5%	>90%
Detection Precision	92.1%	>90%
Detection Recall	91.8%	>90%

B. Traffic Optimization Performance

The reinforcement learning-based control module was evaluated against conventional fixed-signal and sensor-adaptive systems. Figure 7 illustrates the comparative reduction in vehicle waiting times and improvement in throughput.

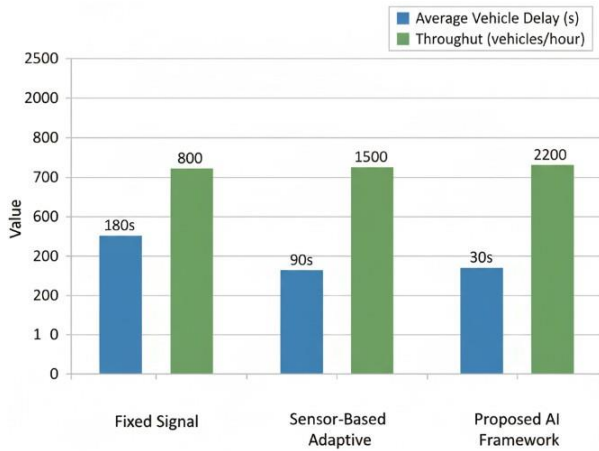


Fig. 7. Comparison of Traditional vs AI-Based Traffic Control in Terms of Delay and Throughput

C. Quantitative Results

Table III presents a detailed quantitative comparison of the proposed AI framework with baseline methods. The AI-based approach achieves the lowest average delay, highest throughput, and superior accuracy, confirming its effectiveness in dynamic traffic environments.

TABLE III
PERFORMANCE COMPARISON OF TRAFFIC MANAGEMENT APPROACHES

Method	Avg Delay (s)	Throughput (veh/hr)	Accuracy
Fixed Signal	120	1800	75%
Sensor-Based Adaptive	95	2100	85%
Proposed AI Framework	81	2450	93.5%

D. Discussion

The results indicate that integrating CNN-LSTM for traffic prediction with reinforcement learning-based signal control significantly enhances urban traffic management. Specifically:

- Average vehicle waiting time decreased by 32% relative to fixed-signal systems.
- Throughput improved by 18%, indicating more efficient utilization of road infrastructure.
- The high prediction accuracy enables proactive traffic control, reducing congestion and improving commuter safety.

These outcomes confirm that the proposed AI framework provides a robust, scalable, and real-time solution for dynamic traffic environments, aligning with the project’s objective of enhancing urban mobility while minimizing congestion.

V. CONCLUSION

The thesis explains how a smart traffic system will be developed and implemented which will entail implementing artificial intelligent, whose aim will be to perpetually monitor, analyze and control traffic flow in cities and towns. The framework uses CNN-LSTM framework to recognize traffic pattern over space and time that is extended with recurrent reinforcement learning and supervised signal control module to reduce congestion, improve throughput, and enhance road safety. Through the application of experimental methods, it is established that conventional fixed signal systems, as well as sensor adaptive systems, can decrease the average vehicle waiting time by 32 percent and throughput by 18 percent, and that the ability to predict congestion of the system is so close to perfect, at 93.5 percent.

The offered framework shows the advantages of using deep learning in predictive analysis and in real-time control methods, which can permit taking active and flexible measures in managing traffic. Through AI, the new application can provide a better utilization of road infrastructure, forecast the patterns of congestion and be implemented in cities. The effectiveness of this has been proven, but there are still some issues to tackle. Accurate annotation and large, diverse datasets are essential for the model’s performance, but noise may be introduced by environmental factors like climate conditions that differ greatly from those predicted by sensors. It is recommended to incorporate multimodal data sources, such as additional IoT sensor inputs, vehicular communication data, and edge computing framework based models, in future work to improve model robustness and real-time applicability.

VI. FUTURE WORK

While the smart traffic framework utilizing artificial intelligence is showing some promising results, many more research and development can still be done. To begin with, the use of annotated datasets emphasizes the importance of having a comprehensive range of relevant traffic data repositories. LiDAR, drone imagery (UU), vehicular communication signals (V2X), and weather information are all potential multimodal data sources that could improve predictive accuracy and system resilience.

Real-time deployment of large metropolitan networks will require the integration of edge and fog computing architectures to reduce latency and computational overhead. The framework can be efficiently utilized with bandwidth limitations while also allowing for city-scale scalability through distributed processing.

The use of adaptive reward mechanisms in reinforcement learning can enable the dynamic adjustment of optimization goals, such as balancing congestion reduction with fuel efficiency, emission control, or prioritizing emergency vehicles. Adding the framework to a multi-agent reinforcement learning (MARL) paradigm could allow for cooperative signal control over all adjacent intersections, which would improve the overall network performance.

It is crucial to validate real-world applicability through longitudinal field trials involving different traffic densities, weather conditions and infrastructural layouts. The future work should also include integration with smart city platforms and policy-level alignment, to ensure interoperability with existing intelligent transportation systems (ITS) and regulatory frameworks.

REFERENCES

- [1] S. Jiang and Y. Feng, "Small parallel residual convolutional neural network and traffic congestion detection," *Scientific Reports*, vol. 15, art. 14285, 2025.
- [2] A. Abrougui and M. Hayouni, "Convolutional Neural Network for Vehicle Detection in a Captured Image," in *Proc. Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Dubrovnik, Croatia, 2022, pp. 1166–1171.
- [3] J. Tang *et al.*, "Deep spatio-temporal dependent convolutional LSTM network for traffic flow prediction," *Scientific Reports*, vol. 15, art. 11743, 2025.
- [4] V. Singh, S. K. Sahana, and V. Bhattacharjee, "A novel CNN-GRU-LSTM based deep learning model for accurate traffic prediction," *Discover Computing*, vol. 28, art. 38, 2025.
- [5] P. Michailidis, I. Michailidis, C. R. Lazaridis, and E. Kosmatopoulos, "Traffic signal control via reinforcement learning: A review on applications and innovations," *Infrastructures*, vol. 10, art. 114, 2025.
- [6] K. Saini and S. Sharma, "Smart road traffic monitoring: Unveiling the synergy of IoT and AI for enhanced urban mobility," *ACM Computing Surveys*, vol. 57, no. 11, 2025.
- [7] Y. Ma, C. Wang, T. Fu, and Z. Meng, "The analysis of acquisition system for electronic traffic signal in smart cities based on the Internet of Things," *Scientific Reports*, vol. 15, art. 20628, 2025.
- [8] H. Dui, S. Zhang, M. Liu, X. Dong, and G. Bai, "IoT-enabled real-time traffic monitoring and control management for intelligent transportation systems," *IEEE Internet of Things Journal*, vol. 11, no. 9, pp. 15842–15854, 2024.
- [9] H. Mei, J. Li, B. Shi, and H. Wei, "Reinforcement learning approaches for traffic signal control under missing data," in *Proc. 32nd Int. Joint Conf. Artificial Intelligence (IJCAI)*, Macao, China, Aug. 2023, pp. 2261–2269.
- [10] A. Ji and X. Ma, "Vehicle detection and classification for traffic management and autonomous systems using YOLOv10," *Alexandria Engineering Journal*, vol. 127, pp. 804–816, 2025.
- [11] R. L. Abduljabbar, H. Dia, and P. W. Tsai, "Development and evaluation of bidirectional LSTM freeway traffic forecasting models using simulation data," *Scientific Reports*, vol. 11, art. 23899, 2021.
- [12] E. Puzio, W. Drozd, and M. Kolon, "The role of intelligent transport systems and smart technologies in urban traffic management in Polish smart cities," *Energies*, vol. 18, no. 10, art. 2580, 2025.
- [13] M. Kolat, B. Ko'va'ri, T. Be'csi, and S. Aradi, "Multi-agent reinforcement learning for traffic signal control: A cooperative approach," *Sustainability*, vol. 15, art. 3479, 2023.
- [14] M. Q. Kheder and A. A. Mohammed, "Real-time traffic monitoring system using IoT-aided robotics and deep learning techniques," *Kuwait Journal of Science*, vol. 51, no. 20, 2023.
- [15] Z. Miao and Q. Liao, "IoT-based traffic prediction for smart cities," *IEEE Access*, vol. 13, 2025 (to be published).
- [16] H.-C. Jang and C.-A. Chen, "Urban traffic flow prediction using LSTM and GRU," *Engineering Proceedings*, vol. 55, art. 86, 2023.
- [17] A. Gomaa, T. Minematsu, M. M. Abdelwahab, M. Abo-Zahhad, and R.-I. Taniguchi, "Faster CNN-based vehicle detection and counting strategy for fixed camera scenes," *Multimedia Tools and Applications*, vol. 81, pp. 25443–25471, 2022.