<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.

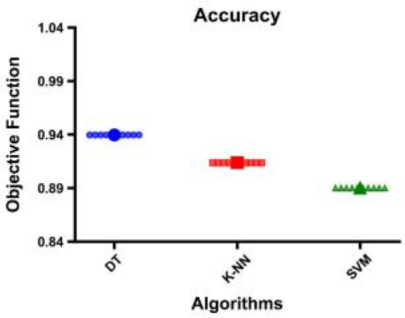

Fig. 4. Accuracy Comparison Between Models

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.