APPLICATION OF MACHINE LEARNING TECHNIQUES FOR BIOMEDICAL DIAGNOSIS SYSTEM

M.Sheerin Banu[1], R. Pitchai[2], N. Balamurugan[3], R. Krishna Kumari[4],
Kavitha V P[5], Ashok Kumar[6], Swagata Sarkar[7]*

[1]Department of Information Technology, R. M. K. Engineering College, Kavaraipettai, Tamil Nadu 601206, India. **hod.it@rmkec.ac.in**

[2]Department of Computer Science and Engineering, B V Raju Institute of Technology, Narsapur, Telangana 502313, India. **pitchrks1984@gmail.com**

[3]Department of Mathematics for Excellence, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu 602117, India. **knbmurugan@gmail.com**

[4]Department of Mathematics, Panimalar Engineering College, Chennai, Tamil Nadu 600123, India. **krishrengan@gmail.com**

[5]Department of Electronics and Communication Engineering, Velammal Engineering College, Chennai, Tamil Nadu 600066, India. **kavithashri410@gmail.com**

[6]Department of Computer Science, Banasthali Vidyapith, Rajasthan 304022, India. **kuashok@banasthali.in**

[7]Department of Artificial Intelligence and Data Science, Sri Sai Ram Engineering College, Chennai - 600044, Tamil Nadu, India. **swagata.b.sarkar@gmail.com**

Corresponding author: **swagata.b.sarkar@gmail.com**

Abstract:-

The majority of machine learning technique research to far has been application-driven and has remained focused on technical challenges. The effective uses of machine learning techniques discussed during the Workshop on Machine Learning in Medical Applications are outlined in this letter. The workshop's objectives were to promote fundamental and applied research in the use of machine learning techniques for solving medical problems and conducting medical research, to offer a venue for disclosing noteworthy findings, to ascertain whether machine learning techniques can serve as the foundation for research and development on intelligent systems for medical applications, and to pinpoint those areas where further study is most likely to result in advancements. There were many suggestions made for a research agenda, including both technical and human-centered concerns.

Keywords:- COVID 19, Biomedical, Machine Leaning and 5G Networks

1. INTRODUCTION

In the last century, the novel coronavirus (COVID-19) epidemic has caused a catastrophic crisis that has spread over the whole planet. Around the world, COVID-19 prevalence is increasing quickly every day. Although there are currently no known vaccinations for this pandemic, deep learning algorithms have shown to be a potent weapon in the toolbox of doctors for the quick identification of COVID-19. This work intends to provide an overview of newly created deep learning systems that use several medical imaging modalities, such as computer tomography (CT) and X-ray. This paper contains information on well-known data sets that were used to train these networks and especially addresses the deep learning systems created for COVID-19 diagnosis. Additionally, it illustrates the numerous performance metrics and data partitioning strategies created by academics in this area. To properly understand the most current works, a taxonomy is created [1]-[5]. Finally, we wrap off by discussing the difficulties in using deep learning techniques to COVID-19 identification and potential future directions in this field of study. This research aims to provide specialists (medical or otherwise) and technicians fresh insights into the ways deep learning methods are used in this context and how they could function even better in containing the COVID-19 epidemic. The creation of COVID-19 diagnostic systems based on deep learning using information gathered from medical imaging samples. The evaluated systems are categorised using a taxonomy that uses deep transfer learning and tailored deep learning techniques to pre-train models. We examine the most important methods created for COVID-19 diagnosis, stressing certain elements including the experimental data, the data splitting method, and the assessment metrics. The difficulties of current deep learning-based systems are openly discussed, and a forecast of next developments is also provided [6]-[10].

In a range of medical disciplines, ML offers approaches, strategies, and tools that may assist in resolving diagnostic and prognostic issues. ML is being used to extract medical information for outcomes research, plan and support therapies, and manage patients as a whole. It is also being used to analyse the significance of clinical characteristics and their combinations for prognosis, such as the prediction of disease development. The application of ML in data analysis includes the interpretation of continuous data used in intensive care units, the discovery of regularities in the data by effectively handling defective data, and intelligent alarms that leads to effective and efficient monitoring. It is believed that the effective use of ML techniques may aid in the integration of computer-based systems in the healthcare environment, allowing chances to facilitate and enhance the work of medical professionals and eventually improving the effectiveness and quality of medical treatment. We have included some of the most important ML applications in medicine [11]-[15].

2. BIOMEDICAL TECHNNOLOGIES

Computer-based systems have several key applications in the field of medical diagnostic reasoning. Expert systems and model-based approaches provide methods in this context for the creation of hypotheses from patient data. For instance, rules are taken from expert systems' knowledge by experts. Unfortunately, experts often aren't aware of or may not be able to articulate the information they really use to problem-solving. Expert systems can now acquire and maintain information thanks to symbolic learning approaches (such inductive learning via examples). Learning in intelligent systems may be accomplished using ML techniques that can provide a systematic description of those clinical attributes that specifically describe the clinical circumstances given a group of clinical cases that serve as examples. Simple rules or decision trees are common methods for expressing this information. KARDIO, a system designed to read ECGs, is a prime example of this kind. This strategy may be expanded to address situations in which there is no prior knowledge on how to read and comprehend medical data. For instance, an intelligent system for detecting changes in a patient's state is detailed in the work. This system uses real-time patient data collected during cardiac bypass surgery to develop models of normal and pathological cardiac physiology. These models may also be used in a research context to generate early hypotheses that guide more testing.

Due to its ability to assist in the early diagnosis of illnesses in a relatively quick and accurate way, machine learning has become more popular in the healthcare industry. Machine learning may be used in the medical industry to help in illness diagnosis. It is often used to help with breast cancer detection using ultrasound or X-ray imaging. In this case, supervised learning classification is utilised to determine if cancer is widespread or not (a discrete binary label). Additionally, the algorithm may provide helpful information to health practitioners by describing why it classified a picture in the manner that it did. The prediction of Alzheimer's disease is another well-known example of how machine learning is used in the medical field. Using a collection of audio recordings, the machine learning model in this example looks for patterns in the speech of people who have this condition. The analysis makes use of the sounds' volume and frequency as well as the breaks between phrases. The approach helps a specialist in geriatric medicine identify verbal early warning indicators of Alzheimer's disease. The algorithms used in machine learning can handle a lot of data. They do it as part of their investigation of crucial connections in the choice-making process. There is a tremendous quantity of knowledge in the subject of medicine that has to be described. People used to save and store this data, but technology has since developed. Based on all collected data, ML algorithms are configured to give multidimensional solutions [16]-[20]. Additionally, new explanations may be discovered using medical machine learning. The information acquired may be used more effectively by medical institutions. In this circumstance, algorithms are really helpful. Information management is vital to save time for critical physicians. It is easy for doctors to help people with their problems. Medical institutions may lower the expenses associated with different treatment methods in the meanwhile.

3. MACHINE LEARNING

The artificial neural network (ANN) is a modelling strategy inspired by the human nervous system that enables pattern learning from representative data that represents an actual event or a decision-making process. A distinctive quality of ANN is its capacity to construct real correlations between independent and dependent variables as well as to extract complex knowledge and delicate information from representative data sets. The relationships between independent and dependent variables may be established without making any assumptions about a formal representation of the events. Compared to regression-based models, ANN models provide a number of benefits, including the capacity to manage noisy data.

An ANN consists of a layer of input nodes and a layer of output nodes connected by one or more layers of hidden nodes. Input layer nodes communicate with hidden layer nodes by activating them, whereas hidden layer nodes respond to evidence by activating or remaining latent. The hidden levels assess the evidence, and when a node's value, or the value of a group of nodes, exceeds a predetermined threshold, a value is transmitted to one or more nodes in the output layer. To train ANNs, a large number of instances must be employed (data). ANNs cannot be utilised for rare or severe situations if there is insufficient data to train the model. Quantitative data cannot be substituted with human judgement in ANNs. A neural network with more than three layers, including the inputs and outputs, is referred to as a deep learning algorithm. A basic neural network is a two- or three-layer neural network.

An imaging test called an ultrasound uses sound waves to provide an image of the organs, tissues, and other internal components of the body (also known as a sonogram). Contrary to x-rays, ultrasounds do not use radiation. The advantages of ultrasonic imaging over other medical imaging modalities include its simplicity, non-invasiveness, and real-time capabilities. Patients are exposed to radiation during computed tomography (CT), while magnetic resonance imaging (MRI) is non-invasive yet pricey and time-consuming. US imaging is thus frequently used in a range of medical specialties for both screening and final diagnosis is shown in Figure 1.

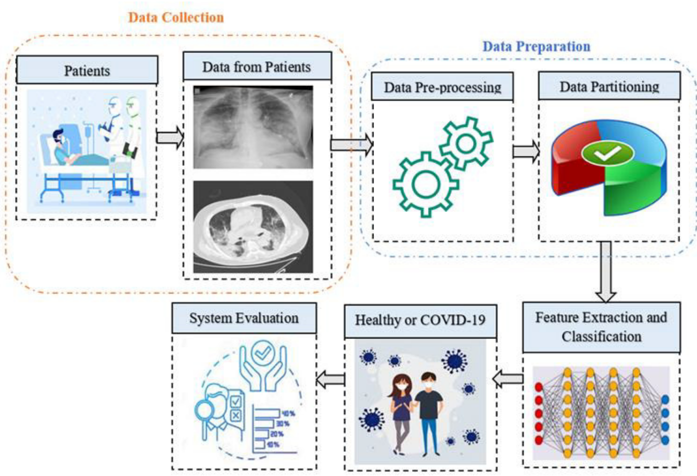

Figure.1. Biomedical Diagnosis Using Machine Learning

Sonologists, radiographers, and pathologists are among the medical professionals that heavily depend on pictures, and deep learning, a common AI technology, excels at identifying visual patterns. Although obstetric and gynaecological ultrasonography are two of the most used imaging techniques, artificial intelligence (AI) has so far had little of an influence in this area. However, AI holds great promise for assisting with repetitive ultrasound tasks, such as automatically choosing high-quality acquisitions and providing near-instant quality assurance. This potential requires interdisciplinary collaboration between ultrasound specialists and AI developers. The turning point that will enable us to close the healthcare gap in the cancer care continuum is new technology. Artificial intelligence (AI) has become a technology that is revolutionising industries. AI-driven clinical care has the potential to significantly reduce health disparities, particularly in low-resource environments. AI integration in cancer care may enhance diagnostic speed and accuracy, support clinical judgement, and provide better health outcomes.

The development of artificial intelligence (AI) has the potential to advance cancer monitoring, medication discovery, tumour genetic characterisation, and cancer screening. Cancer is a complex and multidimensional illness with several genetic and epigenetic alterations. AI-based algorithms offer a lot of promise for early detection of aberrant protein interactions and genetic anomalies. Bringing AI technology to clinics in a safe and ethical manner is another focus of contemporary biomedical research. Cancer remains one of the deadliest illnesses despite substantial advancements in treatment and diagnostic techniques. This may change in the near future thanks to AI. Future cancer cure announcements might result from its pattern-finding capabilities helping doctors plan treatments and lowering false positives and negatives.

The largest organ in the human body is the liver. It is in charge of all bodily metabolic processes, including the transformation of dietary materials into usable body chemicals and the storage and subsequent delivery of these molecules to the cells as required. It also oversees conversion. converting poisonous compounds into non-toxic materials One of the other crucial functions of the liver is bile production. proteins are synthesised, glucose is stored and released, haemoglobin is processed, blood is cleaned, and the immune system is protected. Factor synthesis, bilirubin elimination, and other processes. As a consequence, it is the most significant and fundamental organ in the body. To enhance general health, it must have a healthy state. However, the truth is that the majority of individuals Health is something that you should ideally disregard. Given the complexity of the early liver disease symptoms, it could be challenging to recognise them. Problems with liver disorders are sometimes not identified until it is too late since the liver continues to function even when partly damaged. The ability to spot problems early may save lives. Even the most skilled medical professional may not be able to see the early signs of a variety of illnesses. Early diagnosis increases a patient's life expectancy significantly. Multiple clinical parameters have been used in the development of machine learning algorithms to predict disease risk and outcomes, including evaluating liver fibrosis and steatosis, forecasting liver decompensation in primary sclerosing cholangitis, screening and selecting liver transplant recipients, and predicting post-transplant survival and complications.

Modern hospitals often use monitoring and data gathering equipment, and vast information systems collect and share data. The machine learning approach is currently ideally suited for analysing medical data, and much work is being done in the field of medical diagnosis, particularly in micro or minor diagnostic circumstances.

There are a few issues to think about when utilising machine learning to identify illnesses. First off, artificial intelligence cannot completely replace a human doctor. While machine learning may aid in sickness prognosis, it cannot replace the role of an expert in all situations. For instance, machine learning may help identify cancer very quickly and early, but the clinician still needs to choose the treatment. For machine learning to be effective, the data must be good. If there are no patterns in the data or the data is of low quality, it is meaningless. Similar to this, there is a possibility that previous patterns won't hold true for new data when a model is trained using a set of data. New models must be used by hospitals and doctors. The job description of a doctor does not usually include updating Machine Learning models. Finding time for a health expert to help with data and result verification might be challenging. It might be dangerous to use a trained model that no longer provides accurate projections, particularly when human health is at stake. It is also risky to go from a tried-and-true paradigm to one that may not work.

Understanding the meaning of this term and appreciating the development of machine learning the fundamental building block of AI that sparked its culmination in the biomedical sciences is crucial for understanding the current techniques used in AI and the computational algorithms that drive novel advancements within the field. While the Turing Test algorithm, which was developed during World War II to predict the location of U-Boats, used the same principles that underlie machine learning today and set the stage for its advancement for years to come, it is the more advanced form of machine learning, known as deep learning, that has enabled AI algorithms to function with the computational complexity and human-like intelligence that unlocks its great potential. Machine learning, or algorithms that use some form of memory-based and context-dependent learning, were the forerunners of artificial intelligence (AI). These algorithms employ both top-down and bottom-up analytical approaches, where memory from earlier training sessions is loaded into weights and biases to make predictions on the dataset, or feature extraction and subsequent analysis from pre-established metrics, respectively. The former is known as supervised machine learning, where training on a dataset is used to create "ground truth," or a point of reference for the algorithm to use when estimating and making predictions during the analysis of new (related) data, and the latter is known as unsupervised machine learning, where analysis and predictions on a dataset are made in accordance with predetermined metrics and formulas that can be applied to virtually any novel dataset in each new iteration. The most recent and widely used paradigm in AI, deep learning, is the result of combining various aspects of these two machine learning subsets. This algorithm uses previously stored memory of labelled datasets on which it was trained to make predictions and perform analysis on new datasets using this memory. We shall examine some of the algorithms that make up these different AI paradigms below.

Deep Learning is frequently used in the biomedical sciences through an algorithm known as the Convolutional Neural Network (CNN) to perform analysis and make predictions on a new dataset that is similar to that which the neural network was trained on. Deep Learning is modelled after the human brain, using facets of the cortex and multiple cortical layers that interact to allow intelligent decision-making and inference skills. The neural network, which is often used in image processing tasks, consists of numerous layers of "neurons," some of which are discrete and serve purposes that do not affect the final output, and others of which are engaged in representing outputs that are significant to the user. An approach for back propagation known as gradient descent loss is used to train a neural network on a particular dataset and perform future computation on a dataset of the same kind. Gradient descent algorithms, like convolutional neural networks, are frequently used in image analysis tasks where masks of the object or region of interest (ROI) are input into the algorithm along with the original image in question in order to train the network on the dimensions of the ROI when it encounters them in future, previously unseen datasets . Gradient descent loss iterations reduce loss to a local minimum on the gradient descent curve. This loss, also known as dice loss, is determined by contrasting the current iteration's image and dimensionality with the ground truth that was previously established. In this sense, loss is comparable to cost in that it asks "how different is this ROI from the ROI that I am trained to detect and analyse. Based on calculated loss, the neural network's hyper parameters, connection weights and the biases associated with those weights, which were used to arrive at its current prediction in the current iteration of training, are adjusted. The detection, segmentation, and analysis of a ROI in a given dataset may seem difficult, but after the algorithm has been trained on the available dataset, it performs in a highly robust and intelligent manner, frequently outperforming image specialists, radiologists, and scientists in their respective fields. Although applications and implications will be covered in a later section, it is crucial to note that these fundamental ideas that make up the CNN enable scientists and researchers to use a tool that provides high throughput, quick monitoring, segmentation, and analysis of their dataset, and can be adjusted or completely rebuilt to address any issue that the neural network can be trained to detect and analyse using a plentiful and labelled dataset. This very requirement of neural networks is one of their drawbacks; in order to train the network and establish the ground truth, a large amount of training data must be generated or, more frequently, readily available. Additionally, labelling the data and creating masks of the labelled regions can be a time-consuming task, which can be challenging. However, it is clear that the effort and engineering needed to create such an algorithm are worthwhile and may provide fresh perspectives on issues that were formerly thought to be insurmountable.

4. CONCLUSION

Despite the fact that these AI algorithms operate with a high level of computational and mathematical rigour, their applications in the field of biomedical imaging have greatly increased our understanding of previously challenging quantification tasks and measures that help in data analysis, especially that which is noisy and necessitates extensive time and scrutiny with the naked eye in order to yield results. The validity and accuracy of human selected ROIs and the quantification of biomedical images are also called into question by the lack of a standardised method or approach for assessing and analysing biomedical images, as each image is frequently subject to selection bias by the radiologist, image specialist, or rater in question. CNNs have been used in the past for a variety of segmentation and quantification tasks, including the segmentation of brain and abdominal tumour lesions from patient scans, the prediction of ventricular heart disease, and the histological analysis of cancerous tissue to detect markers of neoplasia to aid in the prediction of the ranostics. This has made it possible to conduct additional research and use AI to diagnose, provide a prognosis for, and predict treatment response for a variety of diseases in a range of imaging modalities, including MRI, CT, Positron Emission Tomography (PET) imaging, optical imaging, and microscopy. This creates a strong foundation for the advancement of radiomics and other fields within the developing field of precision medicine. Similar to this, large-scale gene expression data clustering and unsupervised machine learning techniques like the K means++ clustering method have been used to find molecular markers for cancer in populations of patient cohorts.

When compared to board-certified radiologists and other imaging specialists in the field, deep learning algorithms have performed with unprecedented accuracy, frequently analysing on par with or better than their human counterpart. This has been demonstrated by the Intra class Correlation Coefficient (ICC) from numerous studies used to evaluate these algorithms. This enables the segmentation of ROIs from various imaging modalities in a consistent, high-throughput, and quick manner, which may not only save time in the lab and clinic but also shine light on and shed insight on previously unrecognised issues and areas of investigation. As an illustration, artificial intelligence can be used in the field of stem cell tracking or genomic analysis to not only track the longitudinal development and differentiation of pluripotent stem cells or transplanted islets through various visible molecular patterns and biomarkers, but it can also recognise specific genetic patterns among large amounts of sequence data in a high-throughput manner. Furthermore, in the context of the various imaging modalities where deep learning and clustering algorithms perform segmentation tasks, the improvement in speed and accuracy in identifying and segmenting the ROI from background noise can allow for greater quantification measures in a pipeline that can permit prediction of various metrics that previously could not be calculated due to the limitations of image quantification such as selection bias and low contrast image scans that make it difficult to calculate metrics that are related to the various imaging modalities. What is crucial to understand is that AI in this context is not intended to replace radiologists and other experts in the field, but rather to act as a potent tool that will unlock significant potential in the field of biomedical research, support researchers and clinicians in their various processes, and standardise the current approach to biomedical image quantification to boost throughput, reliability, and accuracy in the industry.

REFERENCES

[1] Bansal, S., Kori, A., Zulfikar, W., Wexler, J., Markiewicz, C., Feingold, F., et al. (2021). High-sensitivity detection of facial features on MRI brain scans with a convolutional network. bioRxiv.

[2] Boonzaaier, D. (2017). Training a Neural Network for Checkers. Ph. D. thesis.

[3] Dey, D., Chaudhuri, S., and Munshi, S. (2018). Obstructive sleep apnoea detection using convolutional neural network based deep learning framework. Biomedical engineering letters, 8(1), 95-100.

[4] Julie, E. G., Robinson, Y. H., and Jaisakthi, S. (2021). Handbook of Deep Learning in Biomedical Engineering and Health Informatics: CRC Press.

[5] Kumar, S., Negi, A., and Singh, J. (2019). Semantic segmentation using deep learning for brain tumor MRI via fully convolution neural networks. In Information and Communication Technology for Intelligent Systems (pp. 11-19): Springer.

[6] Mansour, R. F. (2018). Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. Biomedical engineering letters, 8(1), 41-57.

[7] Park, C., Took, C. C., and Seong, J.-K. (2018). Machine learning in biomedical engineering: Springer.

[8] Raghunath, S., Cerna, A. E. U., Jing, L., Stough, J., Hartzel, D. N., Leader, J. B., et al. (2020). Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. Nature medicine, 26(6), 886-891.

[9] Wei, R., Zhang, X., Wang, J., and Dang, X. (2018). The research of sleep staging based on single-lead electrocardiogram and deep neural network. Biomedical engineering letters, 8(1), 87-93.

[10] Wood, D. A., Kafiabadi, S., Al Busaidi, A., Guilhem, E. L., Lynch, J., Townend, M. K., et al. (2021). Deep learning to automate the labelling of head MRI datasets for computer vision applications. European Radiology, 1-12.

[11] Weber, G., Mandl, K. & Kohane, I. Finding the missing link for big biomedical data. *JAMA* **311**, 2479–2480 (2014).

[12] Van Horn, J. & Toga, A. Human neuroimaging as a “big data” science. *Brain Imaging Behav.* **8**, 323–331 (2014).

[13] Zhou, L. & Verstreken, P. Reprogramming neurodegeneration in the big data era. *Curr. Opin. Neurobiol.* **48**, 167–173 (2018).

[14] Vallejos, C. A., Richardson, S. & Marioni, J. C. Beyond comparisons of means: understanding changes in gene expression at the single- cell level. *Genome Biol.* **17**, 1–14 (2016).

[15] Y. A. Qadri, A. Nauman, Y. Bin Zikria, A. V. Vasilakos and S. W. Kim, “The future of healthcare internet of things: A survey of emerging technologies,” *IEEE Communications, Surveys and Tutorials*, vol. 22, no. 2, pp. 1121–1167, 2020.

[16] J. Yanase and E. Triantaphyllou, “A systematic survey of computer-aided diagnosis in medicine: Past and present developments,” *Expert Systems with Applications*, vol. 138, pp. 112821, 2019.

[17] N. C. Caballe, J. L. Castillo-Sequera, J. A. Gómez-Pulido, J. M. Gómez-Pulido and M. L. Polo-Luque, “Machine learning applied to diagnosis of human diseases: A systematic review,” *Applied Sciences*, vol. 10, no. 15, pp. 1–27, 2020.

[18] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li *et al.*, “Arti□cial intelligence in healthcare: Past, present and future,” *Stroke and Vascular Neurology*, vol. 2, no. 4, pp. 230–243, 2017.

[19] J. Schaefer, M. Lehne, J. Schepers, F. Prasser and S. Thun, “The use of machine learning in rare diseases: A scoping review,” *Orphanet Journal of Rare Diseases*, vol. 15, no. 1, pp. 145, 2020.

[20] W. Rogers, B. Ryack and G. Moeller, “Computer-aided medical diagnosis: Literature review,” *International Journal of Bio-Medical Computing*, vol. 10, no. 4, pp. 267–289, 1979.