**AN EXAMINATION OF ARTIFICIAL INTELLIGENCE METHODS FOR DISEASE DIAGNOSIS**

**J. Fasila Afreen1**

1Student, Department of CSE, Kings College of Engineering, Pudukkottai, Tamil Nadu, India

**ABSTRACT**

 Emerging computer-based technologies are developing too quickly. Many potentials exist in digital healthcare to lower human error, enhance therapeutic outcomes, and monitor data over time. Artificial Intelligence (AI) technologies, including Machine Learning (ML) and Deep Learning (DL) algorithms, are widely employed in the prediction and diagnosis of numerous diseases, especially those whose diagnosis is based on imaging or signaling analysis. AI can also aid in determining whether groups of people or geographical regions are more likely to have a disease or engage in high-risk activities. Due to the sophisticated algorithms that allow for the automatic extraction of better features, ML approaches have seen substantial success in the analysis of medical images. There are three types of machine learning (ML) that are based on learning techniques: supervised (classification, regression, and composition), unsupervised (association, clustering, and dimensionality), and reinforced learning.

A successful method for classifying and analyzing images is the convolution neural network (CNN), which is a supervised deep learning (DL) model. Fully connected layers with standard weights make up CNN, which reduces the number of parameters needed to train features through the backpropagation method. They are made to take spatial data out of the input images. CNN attempts to automatically classify picture data, learn hierarchical features, and extract their characteristics. This algorithm's key benefit is learning extremely abstract characteristics with few parameters and straightforward preprocessing

1. **INTRODUCTION**

 Emerging computer-based technologies are developing too quickly. Many potentials exist in digital healthcare to lower human error, enhance therapeutic outcomes, and monitor data over time. Artificial Intelligence (AI) technologies, including Machine Learning (ML) and Deep Learning (DL) algorithms, are widely employed in the prediction and diagnosis of numerous diseases, especially those whose diagnosis is based on imaging or signaling analysis. AI can also aid in determining whether groups of people or geographical regions are more likely to have a disease or engage in high-risk activities. Due to the sophisticated algorithms that allow for the automatic extraction of better features, ML approaches have seen substantial success in the analysis of medical images. There are three types of machine learning (ML) that are based on learning techniques: supervised (classification, regression, and composition), unsupervised (association, clustering, and dimensionality), and reinforced learning.

A successful method for classifying and analyzing images is the convolution neural network (CNN), which is a supervised deep learning (DL) model. Fully connected layers with standard weights make up CNN, which reduces the number of parameters needed to train features through the backpropagation method. They are made to take spatial data out of the input images. CNN attempts to automatically classify picture data, learn hierarchical features, and extract their characteristics. This algorithm's key benefit is learning extremely abstract characteristics with few parameters and straightforward preprocessing.

1. **METHODOLOGY**

Support Vector Machines (SVM), Decision Trees, and Random Forests are examples of supervised learning algorithms that could be used to build a machine learning-based system for multiple illness prediction. A dataset containing the medical records of patients with different conditions would be used to train the system. The model would receive as inputs characteristics like the patient's age, gender, medical history, way of life, and environmental factors. With the data at hand and the model that has been trained, it is possible to determine which disease a patient is most likely to have. In order to aid in the diagnosis and treatment of a certain disease, the model can also be used to pinpoint risk factors related to that condition. The system might potentially make use of natural language processing to examine and extract data from text sources.

**Decision Tree**

This particular supervised machine learning technique focuses primarily on categorization issues. Creating a training model that may be used to forecast the class or values of the desired value is the primary goal of employing decision trees. To do this, basic decision procedure inference is learned from previously collected data (training data. Using the decision tree approach, we begin by predicting the class at the tree's root. We combine the root characteristic's values with the trait of the data. We move to the next node with the branch that is parallel to that value on the basis of differentiation. The decision Tree divides the symptoms in this system according to their classification, which reduces the complexity of the dataset.

**Random Forest**

A random forest can be constructed by merging N decision trees, and then it can be used to make predictions for each tree that was produced in the first step. Supervised machine learning algorithms include Random Forest. Although it is utilized for both classification and regression, classification problems are its main focus. The use of Random Forest is fairly simple, as is the implementation. If we need to create a model quickly, Random Forest is the ideal replacement. A large number of decision trees are built during the training phase of Random Forest, an ensemble learning technique. By using voting, it determines which answer is best. Several decision trees make up Simple Random Forest. A forest of trees is produced. The accuracy rate is precisely related to the number of trees in the forest, which also solves the overfitting issue. Because Random Forest is not dependent on overfitting and is insensitive to dataset noise, it performs well when applied to real-world issues. It is very effective and offers superior performance to other tree-based algorithms. Bootstrap aggregation or bagging is commonly used for tree learning.

**KNN**

KNN is a supervised learning algorithm that classifies fresh data points based on metrics of similarity. It is a non-parametric method, which means that it makes no assumptions about the underlying data. KNN is a lazy learner approach as well because it doesn't do any training and just keeps the training dataset without trying to derive any discriminative functions from it. KNN is based on the concept of the feature similarity approach, which assumes that homogenous objects exist nearby. Also, it is an approach for instance-based learning in which the function is roughly local. This approach is utilized when there are non-linear decision-making divisions between classes and it can handle vast amounts of data.

**XGBoost**

XGBoost is an open-source machine-learning library for gradient boosting. It is an optimized distributed gradient boosting library designed to be highly efficient, flexible, and portable. It implements machine learning algorithms under the Gradient Boosting framework. XGBoost is used for supervised learning tasks such as classification and regression. XGBoost provides a parallel tree boosting (also known as GBDT, GBM) that solves many data science problems in a fast and accurate way. XGBoost is an implementation of gradient-boosted decision trees designed for speed and performance. It is a scalable, portable, and accurate library for large-scale machine learning. XGBoost uses a more regularized model formalization to control over-fitting, which gives it better performance. It also has tools for parallel processing, allowing it to exploit the power of multiple computers to perform computations in a fraction of the time. It has a wide range of hyperparameters that can be tuned to improve the model's performance.

1. **MODELING AND ANALYSIS**

By employing the system, the patient can predict the disease. The output is displayed based on the trained model of the user input after the user adds the input for the particular ailment.



1. **RESULTS AND DISCUSSION**

The application of DL and ML approaches holds considerable promise for revolutionizing the disciplines of disease detection and prediction. When it comes to illness, the accuracy and correctness of the diagnosis are the most crucial components of the treatment process. AI has proven to be extremely accurate in detecting diseases from photos and forecasting treatment outcomes in terms of survival rate and response to therapy. The system's diabetes illness prediction model used the knn algorithm, the xgboost method, and the random forest algorithm since they each offered the best levels of accuracy. The disease-specific parameter that the patient adds will reveal whether or not the patient has the condition in issue.

1. **CONCLUSION**

The fields of disease diagnosis and prediction have a great deal of promise to be revolutionized by DL and ML techniques. The accuracy and correctness of the diagnosis are the most important aspects of the treatment procedure when it comes to sickness. When it comes to the detection of diseases based on images as well as the forecasting of treatment results in terms of survival rate and treatment response, AI has demonstrated a high degree of accuracy. The vast amount of image data necessitates its application into processing stages through quick, accurate, and reliable computing power offered by AI techniques. Accurate disease identification, efficient treatment, and patient safety are crucial factors in disease diagnosis. AI consists of huge and varied data, algorithms, deep learning techniques, different types of neural networks, and cutting-edge technologies that are continually developing to fulfill human requirements. The results of this study show that SVM performs the best for predicting cardiac problems. Due to their high accuracy and quick image recognition, supervised DL networks, such as CNN-based models, are frequently utilized, particularly for the diagnosis of respiratory, lung, skin, and brain illnesses, which has produced notable results. KNN typically performs best when combined with other networks, such as SVM, for the diagnosis of breast cancer. Thus, DL and ML have a significant impact on the prognosis of many diseases mentioned in this article because of their excellent experimental outcomes in the detection and classification of medical pictures. Hence, by maximizing the utilization of various resources, AI-based technologies help medical systems diagnose and anticipate illnesses. Also, with the quick advancement of AI technology, it will no longer be difficult to objectively diagnose a variety of ailments.

1. **REFERENCES**
2. Xi X, Meng X, Yang L, Nie X, Yang G, Chen H, Fan X, Yin Y, Chen X. Automated segmentation of choroidal neovascularization in optical coherence tomography images using multi-scale convolutional neural networks with structure prior. Multimed Syst. 2019;25(2):95.
3. Rizk Y, Hajj N, Mitri N, Awad M. Deep belief networks and cortical algorithms: a comparative study for supervised classification. App Comput Inform. 2019;15(2):81.
4. Bhatt C, Kumar I, Vijayakumar V, Singh KU, Kumar A. The state of the art of deep learning models in medical science and their challenges. Multimedia Syst. 2021;27(4):599.
5. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. Imagenet large-scale visual recognition challenge. Int J Comput Vision. 2015;115(3):211.
6. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Commun ACM. 2017;60(6):84.
7. Owais M, Arsalan M, Choi J, Mahmood T, Park KR. Artificial Intelligence-based classification of multiple gastrointestinal diseases using endoscopy videos for clinical diagnosis. J Clin Med. 2019;8(7):986.
8. Nithya A, Appathurai A, Venkatadri N, Ramji DR, Palagan CA. Kidney disease detection and segmentation using artificial neural networks and multi-kernel k-means clustering for ultrasound images. Measurement. 2020;149: 106952.
9. Depeursinge A, Chin AS, Leung AN, Terrone D, Bristow M, Rosen G, et al. Automated classification of usual interstitial pneumonia using regional volumetric texture analysis in high-resolution computed tomography. Invest Radiol. 2015;50(4):261.
10. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans Med Imag. 2016;35(5):1207.
11. Abedi V, Khan A, Chaudhary D, Misra D, Avula V, Mathrawala D, et al. Using artificial intelligence for improving stroke diagnosis in emergency departments: a practical framework. Therap Adv Neurol Disorders. 2020;13:1756286420938962.
12. Gorris M, Hoogenboom SA, Wallace MB, van Hooft JE. Artificial intelligence for the management of pancreatic diseases. Digestive Endoscopy. 2021;33(2):231.
13. Sinagra E, Badalamenti M, Maida M, Spadaccini M, Maselli R, Rossi F, et al. Use of artificial intelligence in improving adenoma detection rate during colonoscopy: might both endoscopists and pathologists be further helped. World J Gastroenterol. 2020;26(39):5911.
14. Xu L, Gao J, Wang Q, Yin J, Yu P, Bai B, et al. Computer-aided diagnosis systems in diagnosing malignant thyroid nodules on ultrasonography: a systematic review and meta-analysis. European Thyroid J. 2020;9(4):186.
15. Xia C, Li X, Wang X, Kong B, Chen Y, Yin Y, et al. A multi-modality network for cardiomyopathy death risk prediction with CMR images and clinical information in lecture notes in computer science. Cham: Springer; 2019.
16. Wu X, Liu X, Zhou Y. Proceedings of 2021 Chinese intelligent systems conference: a review of unsupervised learning techniques in lecture notes in electrical engineering. Singapore: Springer; 2022.
17. Dhal P, Azad C. A comprehensive survey on feature selection in the various fields of machine learning. Appl Intell. 2022;52(4):4543.
18. Dabowsa NI, Amaitik NM, Maatuk AM, Aljawarneh SA. A hybrid intelligent system for skin disease diagnosis in the international conference on engineering and technology (ICET), Antalya Turkey, 2017.
19. Jader R, Aminifar S. Fast and accurate artificial neural network model for diabetes recognition. NeuroQuantology. 2022;20(10):2187.
20. Momin MA, Bhagwat NS, Dhiwar AV, Chavhate SB, Devekar NS. Smart body monitoring system using IoT and machine learning. Int J Adv
21. Bharti R, Khamparia A, Shabaz M, Dhiman G, Pande S, Singh P. Prediction of heart disease using a combination of machine learning and deep learning. Comput Intell Neurosci. 2021;2021:8387680.
22. Suresha PB, Wang Y, Xiao C, Glass L, Yuan Y, Cliford GD. A deep learning approach for classifying nonalcoholic steatohepatitis patients from nonalcoholic fatty liver disease patients using electronic medical records. In: Explainable AI in Healthcare and Medicine. Switzerland: Springer; 2021.
23. Goenka N, Tiwari S. Deep learning for Alzheimer prediction using brain biomarkers. Artif Intell Rev. 2021;54(7):4827.
24. Ramani V, Shendure J. Smash and DASH with Cas9. Genome Biol. 2016;17:42.
25. Rajaraman S, Candemir S, Kim I, Thoma G, Antani S. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. Appl Sci. 2018;8(10):1715.