International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

PULMONARY IMAGE CLASSIFICATION USING NEURAL NETWORKS AND ENSEMBLE LEARNING TECHNIQUESMODEL

[1]V. ARUNA KUMARI, [2]P.ANUSHA, [3]. GOLLA SWATHI, [4]RAMACHANDRA EDIGA, [5]MOHAMMAD MUNWAR

[1,2,3,4,5]Assistant professor, Department of MCA, Ashoka women's engineering college (Autonomous), Kurnool, Andhra Pradesh, India

E-Mail: [1]arunakumari@ashokacollege.in, [2]p.anusha@ashokacollege.in, [3]gswathi@ashokacollege.in, [4] ramchandra535@gmail.com, [5] mohdmunwar@gmail.com

Abstract: *Pulmonary diseases, such as pneumonia, tuberculosis, and lung cancer, pose significant health risks and demand accurate and timely diagnosis. This study proposes an advanced pulmonary image classification framework that integrates neural networks with ensemble learning techniques to enhance diagnostic performance. The system utilizes convolutional neural networks (CNNs) for deep feature extraction from chest X-ray and CT images, capturing complex patterns and spatial hierarchies within the medical data. To improve robustness and classification accuracy, ensemble strategies such as bagging, boosting, and stacking are employed, combining the strengths of multiple models. The proposed method is evaluated on publicly available pulmonary image datasets, demonstrating superior performance in terms of accuracy, sensitivity, specificity, and AUC compared to individual classifiers. The synergy between deep learning and ensemble methods significantly reduces misclassification rates and improves generalizability. This approach holds promise for assisting radiologists and clinicians in making more accurate and efficient diagnoses of pulmonary conditions.*

Keywords: Pulmonary Image Classification, Neural Networks (NN), Ensemble Learning, Medical Diagnosis, Pulmonary Prediction, Machine Learning.

INTRODUCTION

Now a days machine learning is widely used for various diseases prediction accurately with provided and trained datasets.This paper provides is a study of Predictive Analysis of Pulmonary nodules Disease Based onstudy of appropriate neural network selection and by using ensemble learning. As our proposed pulmonary image classification based neural network selection using VGG_16 Model, Inception_v3, ResNet50 and VGG19.Nodule detection is an acute pulmonary infection caused by bacteria, viruses, or fungi that infects the lungs, producing inflammation of the air sacs and pleural effusion (fluid in the lung). It is the cause of over 15% of all deaths in children under the age of five.[21]. Lung infections are more common in undeveloped and underdeveloped countries, where overcrowding, pollution, and unsanitary environmental circumstances worsen the problem, and medical resources are limited. As a result, early detection and treatment can help prevent the disease from progressing to the point of death. The use of computed tomography (CT), magnetic resonance imaging (MRI), or radiography (X-rays) to examine the lungs is commonly employed for diagnosis[23,24,25]. X-ray imaging is a non-invasive and painless method of obtaining information. Figure 1 displays an example of a lung X-ray with a damaged and a healthy lung. Infiltrates, or white patches on the Lung X-ray Chest X-ray exams for infection identification, on the other hand, are vulnerable to subjective variability. As a result, an automated technique for detecting Nodules Infection is necessary. We created a method based on deep learning methods for dealing with such automation difficulties in this research. The most extensively used and common type of clinical assessment for pulmonary nodules is a chest X-ray film.

However, due to the fast increase in the number of infection diseases, which is also a major potential source of diagnostic error, This explosion is definitely outpacing the amount of radiologists available. As of January 2020, infectious diseases had killed more than half a million Americans, out of a total of two and a half million deaths worldwide, according to the Centers for Disease Control and Prevention.. (CDC) [15].

What was initially assumed to be a respiratory virus began to express itself in other parts of the body, with a long list of symptoms ranging from arrhythmia, heart attacks, blood clots, liver and kidney damage, rashes, and more. Despite this, respiratory problems are still the most common symptom of infectious diseases. In terms of diagnosis, thoracic radiography's specificity for infectious diseases is debatable, and its usefulness for frontline prescreening is also debatable. Several radiology organizations, such as the American College of Radiography[6,7,8], advise against utilizing clinical radiography to diagnose pulmonary causes. Nonetheless, a few researchers believe that a lung scan examination might be utilized as a primary tool for screening various locations, and that it could provide essential information for diagnosis and, in particular, the management of respiratory tract infections. We contributed a better investigation and established a novel Protocols to measure ML models when using heterogeneous data sources, particularly with a large number of patient cases, because to the limits of Strategies for ensuring that the ML models' visual features are particularly documenting the locations of lung anomalies rather than bright objects like medical equipment or hard tissue; Algorithms for tracking the position of a feature in a CXR image processing task and evaluating [22] the relationship with essential factors linked to a variety of viral diseases in the lungs.

The following is the outline for this paper. We begin by reviewing current studies in order to identify potential flaws in employing neural networks to process radiography pictures[9,14]. Then, using the open-access benchmark dataset Infectious Diseases as a case study, we present protocols and strategies for evaluating deep learning models for segmentation and classification.

Deep learning is a powerful artificial intelligence technology that can help solve a variety of difficult computer vision problems [4, 5, 6]. For diverse picture categorization issues, deep learning models, notably convolutional neural networks (CNNs), are widely used. However, such models work best when they are given a huge amount of data to work with. Such a large volume of labelled data is challenging to obtain for biomedical image classification challenges

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

because it requires professional doctors to classify each image, this is a time-consuming and costly task. A workaround for overcoming this barrier is transfer learning. In this strategy, a model trained on a big dataset is re-used and the network weights determined in this model are employed to solve a problem with a small dataset. For biological image classification tasks, CNN models trained on a large dataset like ImageNet [7], which contains over 14 million images, are widely utilized.

Ensemble learning is a popular method for combining the decisions of multiple classifiers to produce a final prediction for a test sample. It is done so that discriminative information from all of the base classifiers may be captured, resulting in more accurate predictions. The most often used ensemble approaches in the literature were average probability, weighted average probability, and majority voting. In the average probability-based ensemble, each member base learner gets equal priority. However, for a certain situation, one base classifier may be superior than others at capturing information. Assigning weights to all of the base classifiers is thus a more effective technique[11]. The most extensively used and common type of clinical assessment for pulmonary nodules is a chest X-ray film. However, because to the fast increase in the number of infectious diseases, which is also a major potential source of diagnostic error, the number of radiologists clearly cannot keep up with this eruption. The deep learning method is the most suited method for dealing with such automation issues. Previously, experts offered a framework for conducting ML research in medicine[23,29]. In order to evaluate the existing ML algorithm in Lung Infectious Disease [8] prediction, systematic review and meta-analysis, which are the foundations of modern evidence-based medicine, must be undertaken.

LITERATURE SURVEY

Marcin Wozniak , DawidPołapproposed a method to perform computer aided diagnosis the goal of this study is to investigate the possibilities of using deep learning algorithms to diagnosis respiratory diseases images by using firefly algorithm, artificial bee colony algorithm ,artificial ant colony, cuckoo algorithm, practical swarm algorithm and extraction is carried out by bim tissue keypoints and aggregated key points ,In the images of lung illnesses like pneumonia, lungs sarcoidosis and cancer medical experts search for tissues that have changed structure. These types of changes are visible in x-ray images with a solid structure similar to bone tissues, which are not permeable to x-ray radiation and therefore visible in images. Schematic Tissue Key-Area’s position detection in x-ray image is performed by the proposed BIM approach over the input image [1]. S.Mukherjee et al. [2] proposed a method for autonomously detecting lung nodules based on geometric parameters. The x-ray pictures are used to classify benign and malignant pulmonary nodules based on shape factors such as roundness, eccentricity, diameter, and aspect ratio.Noise Removal using Bilateral Filtering then Image Binarization and Segmentation and classification is carried out by using Bayesian classifier.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Woniak et al. . proposed a probabilistic neural network-based lung cancer classification system. This method is basic, yet it has a decent classification effect and can detect nodules with low contrast. The following probabilistic neural network was used to extract features from a lung image:. As a result, a vector is generated, the elements of which show how close the input is to single classes in Mahalanobisdistance.. By using this vector, the pattern layer computes a probability vector whose components define the belonging to the different classes. Finally, the output layer selects the largest value of the probability vector to predict the target class, determining whether an input vector belongs to that class.

Here, to create and apply feature extraction methods and algorithms, most of these methods require required professional knowledge or a significant amount of time and effort. With the progressive advancement of deep learning research, the technology that can be employed with photos has also made a qualitative leap. [4] by using a variety of datasets to train a certain domain and a variety of model architectures.

Long and Wang proposed a method To address the problem of domain adaptation in transfer learning [5] they introduced a unique Deep Adaptation Network (DAN), which extends Deep Convolutional Neural Network [30] to domain Adaptation. This architecture optimizes the transferability of features from the task-specific layers of the neural network. In a replicating kernel Hilbert space, mean-embedding matching of multi-layer representations across domains can considerably increase feature transferability. While an unbiased estimate of the mean embedding naturally leads to a linear time approach, which is particularly desirable for deep learning from large-scale datasets, an efficient multi-kernel selection strategy boosts embedding matching effectiveness even more.

The usage of multiple classifier systems (or ensemble systems) and then merging the results of their outputs is one of the suitable ways for improving classification accuracy. The "creation of ensemble" and "combination of class label" are the two main components of a multiple classifier system.[6]. To create and apply feature extraction methods and algorithms, most of these methods require required professional knowledge or a significant amount of time and effort. The technology that can be applied to photos has evolved qualitatively as a result of the progressive advancement of deep learning research, giving rise to the notion of medical picture categorization based on deep learning. Deep learning does not demand any medical or engineering technology qualities, nor does it necessitate any medical-related specialist knowledge. To categories pulmonary pictures, the existing system uses the inception-v3 transfer learning model. On the JSRT database, the neural network model based on transfer learning outperforms the original DCNN model in pulmonary image categorization. Then, automatically extract features from pulmonary images using the fine-tuned Inception-v3 model based on transfer learning.

METHODOLOGY

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

We created an ensemble framework of several classifiers in this study. Using a weighted average ensemble technique, in which the weights assigned to the classifiers are produced using a novel scheme, as detailed in the sections below. Using Ensemble learning in the Proposing system, the categorization of Pulmonary Image and performance may be improved. The following are the study's primary contributions. An ensemble architecture was designed to improve the performance of the base CNN learners in the categorization of lung nodules. A weighted average ensemble approach was used for this purpose.

The evaluation measures were combined to calculate the weights assigned to the classifiers: accuracy, recall, f1-score, and AUC. We employed a hyperbolic tangent function to set the weights instead of relying exclusively on the accuracy of classifiers or the results of tests.

The proposed model was tested on the JSRT dataset [4] and two publicly available chest X-ray datasets. The results outperform those of state-of-the-art methods, demonstrating that the method is viable for usage in the real world. Ensemble learning is the process of creating and combining many models, such as classifiers or experts, to solve a given computational intelligence problem. Ensemble learning is frequently used to improve the performance of a model (classification, prediction, function approximation, etc.) or to lessen the risk of an unintentional poor model selection.

Dataset: A laser digitizer with a 2048x2048 matrix size (0.175-mm pixels) and a 12-bit grey scale digitized 154 conventional chest radiographs with a lung nodule (100 malignant and 54 benign nodules) and 93 radiographs without a lesion for the database (no header, big-endian raw data). Additional information in the database includes the patient's age, gender, and diagnosis (malignant or benign).

The proposed system is depicted in Fig. 2 as a systematic overview. In brief, the system accepts lung CT scans as input and processes them using two key techniques: image processing and classification. In the first module, noise is removed from photos, segmentation is performed, backgrounds are removed, and the interested items and their features are extracted from raw images. The remaining potential objects are categorized in the second module based on their attributes extracted during the feature extraction phase, allowing lung cancers to be diagnosed. Among suspicious items, the system would be able to differentiate between nodule and non-nodule. This classification is based on a committee of three different classifiers comprising VGG-16, INCEPTION-V3,RESNET-50,and VGG-19 . Following, the steps of the proposed system have been described, respectively..

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

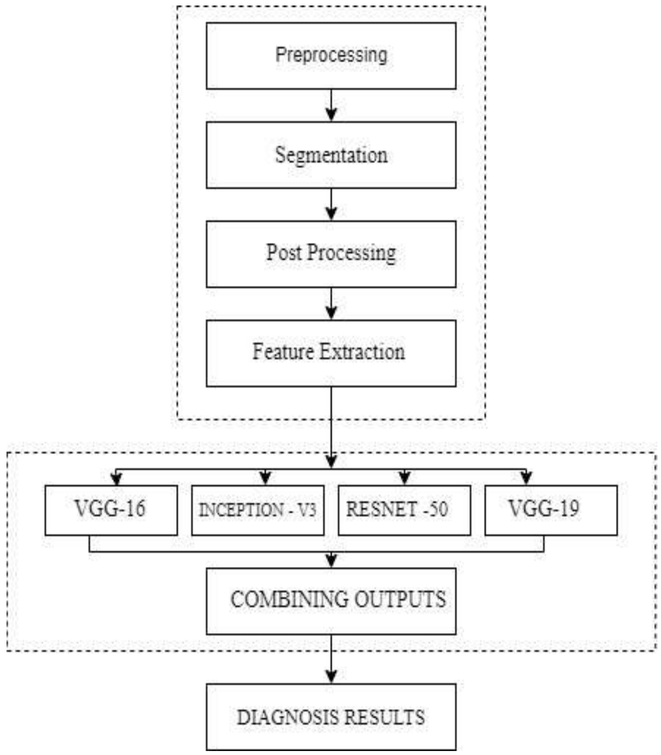

Figure 1: Classification Process

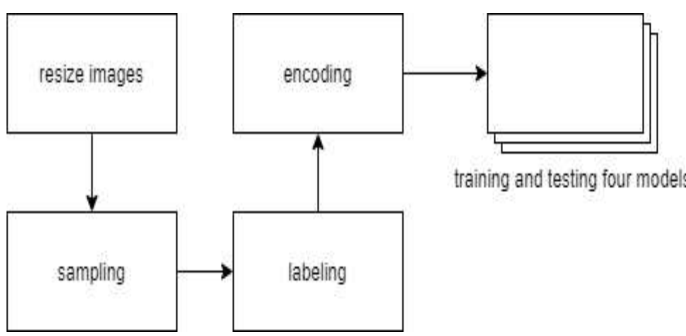

Figure 2: Preprocessing Process

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

RESULT ANALYSYS

In the second part of each ensemble system, the decisions of all these classifiers are combined to create an ensemble output. The most common combination method is Majority voting, though many more powerful techniques such as: Naïve Bayes, Decision, Templates, Dempster, Shafer minimum, Sum, Maximum, Mean rule, and product rule have also been proposed and in many instances may provide even higher classification performance. In this study, each base classifiers has its own opinion and identifies suspended objects as nodule or non-nodule. Then, majority voting techniques is used to combine the results of base classifiers. According, lung nodule among all suspected objects is detected through proposed system.

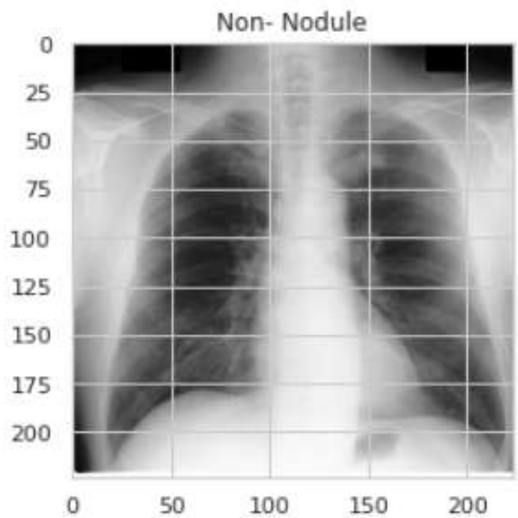

Figure 3: X-ray Showing Non Nodule Image with pulmonary readings

This study was performed by using the 247 images including both men and women collected by jsrt(Japanese society of Radiological Technology).The proposed system detects suspected objects in each image as nodule or non-nodule. In fact, the collected dataset is divided into two categories (positive and negative). There are a variety of metrics that may be used to assess the success of categorization methods that are regularly employed in automatic medical diagnosis systems. TP is the number of accurate predictions for a positive instance; FN denotes the number of erroneous predictions for a negative instance; For a positive instance, FP signifies the number of incorrect predictions; for a negative instance, TN denotes the number of accurate predictions. We could generate the measures below to evaluate the system's performance based on these indicators.

A confusion matrix is a table that lists the actual and predicted categories in a classification system. The data in the matrix is commonly used to assess a system's performance. The confusion matrices for base classifiers and ensemble systems are shown here.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Figure 4: Heat map demonstrating Non Nodule and Lung Nodule Comparisons

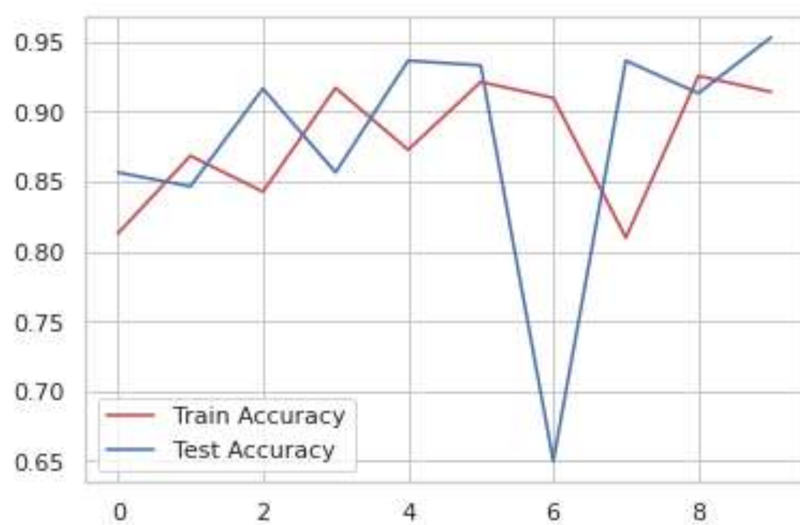

Figure 5: displaying Train and Test Accuracy Comparisons

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

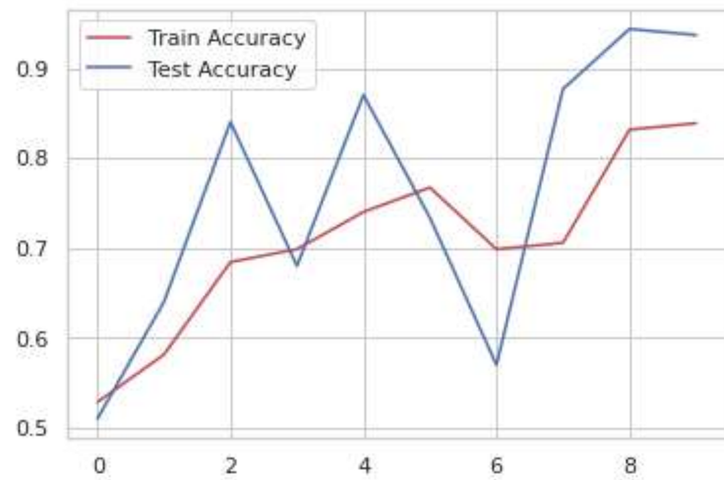

Figure 6: demonstrating Train and Test Accuracy Comparisons

CONCLUSION

This study demonstrates the effectiveness of combining neural networks with ensemble learning techniques for accurate classification of pulmonary images. By leveraging the deep feature extraction capabilities of convolutional neural networks (CNNs) and enhancing model performance through ensemble methods such as bagging, boosting, and stacking, the proposed system achieves improved diagnostic accuracy and robustness. The hybrid approach significantly reduces the risk of misclassification and enhances generalization across diverse pulmonary conditions. Experimental results on benchmark datasets confirm that this integrated method outperforms traditional single-model approaches. The findings highlight the potential of advanced AI-driven solutions to support radiologists and clinicians in early and reliable detection of lung diseases, ultimately contributing to better patient outcomes and more efficient healthcare delivery. Future research may explore real-time deployment, multimodal data integration, and interpretability to further strengthen clinical adoption.

REFERENCES

1) Thatha, V. N., Chalichalamala, S., Pamula, U., Krishna, D. P., Chinthakunta, M., Mantena, S. V., Vahiduddin, S., & Vatambeti, R. (2025b). Optimized machine learning mechanism for big data healthcare system to predict disease risk factor. Scientific Reports, 15(1). https://doi.org/10.1038/s41598-025-98721-6.
2) Shariff, V., Paritala, C., & Ankala, K. M. (2025). Optimizing non small cell lung cancer detection with convolutional neural networks and differential augmentation. Scientific Reports, 15(1). https://doi.org/10.1038/s41598-025-98731-4.

3) S Phani Praveen et al., "AI- Powered Diagnosis: Revolutionizing Healthcare With Neural Networks", Journal of Theoretical and Applied Information Technology, vol. 101, no. 3, February 2025
4) C. S. Kodete, D. Vijaya Saradhi, V. Krishna Suri, P. Bharat Siva Varma, N. S Koti Mani Kumar Tirumanadham and V. Shariff, "Boosting Lung Cancer Prediction Accuracy Through Advanced Data Processing and Machine Learning Models," 2024 4th International Conference on Sustainable Expert Systems (ICSES), Kaski, Nepal, 2024, pp. 1107-1114, doi: 10.1109/ICSES63445.2024.10763338.
5) K. V. Rajkumar, K. Sri Nithya, C. T. Sai Narasimha, V. Shariff, V. J. Manasa and N. S. Koti Mani Kumar Tirumanadham, "Scalable Web Data Extraction for Xtree Analysis: Algorithms and Performance Evaluation," 2024 Second International Conference on Inventive Computing and Informatics (ICICI), Bangalore, India, 2024, pp. 447-455, doi: 10.1109/ICICI62254.2024.00079.
6) V. Pasupuleti, B. Thuraka, C. S. Kodete, V. Priyadarshini, K. M. Kumar Tirumanadham and V. Shariff, "Enhancing Predictive Accuracy in Cardiovascular Disease Diagnosis: A Hybrid Approach Using RFAP Feature Selection and Random Forest Modeling," 2024 4th International Conference on Soft Computing for Security Applications (ICSCSA), Salem, India, 2024, pp. 42-49, doi: 10.1109/ICSCSA64454.2024.00014.
7) S. Vahiduddin, P. Chiranjeevi and A. Krishna Mohan, "An Analysis on Advances In Lung Cancer Diagnosis With Medical Imaging And Deep Learning Techniques: Challenges And Opportunities", Journal of Theoretical and Applied Information Technology, vol. 101, no. 17, Sep. 2023.
8) C. S. Kodete, D. Vijaya Saradhi, V. Krishna Suri, P. Bharat Siva Varma, N. S Koti Mani Kumar Tirumanadham and V. Shariff, "Boosting Lung Cancer Prediction Accuracy Through Advanced Data Processing and Machine Learning Models," 2024 4th International Conference on Sustainable Expert Systems (ICSES), Kaski, Nepal, 2024, pp. 1107-1114, doi: 10.1109/ICSES63445.2024.10763338.