Journal of Theoretical and Applied Information Technology
15th June 2026. Vol.104. No.11
© Little Lion Scientific

ISSN: **1992-8645** www.jatit.org E-ISSN: **1817-3195**

ENHANCING LUNG DISEASE DIAGNOSIS USING A CNN–CONVNEXT INTEGRATED DEEP LEARNING FRAMEWORK

[1]DR.K. SAI MADHURI, [2]VISHAL BHARADWAJ MERUGA, [3]V SITAMAHALAKSHMI,
[4]VENKAT RAO PASUPULETI,[5]YERRAVARAPU V V DURGA PRASAD, [6]MR.ESWAR PATNALA,
[7]JAYAMMA RODDA, ARUN KUMAR UNDAMATLA[8]

[1]Asst Professor, Dept of CSE (DS), B V Raju Institute of Technology, NARSAPUR, Telangana-502313
[2]Data Scientist, Marriott International, USA,
[3]Assistant Professor, Department of FED, PVP Siddhartha Institute of Technology, Andhra Pradesh, India
[4]Assoc Professor, ECE Dept., Lakireddy Bali Reddy College of Engineering, Mylavaram, Andhra Pradesh, India
[5]Asst Professor Department of CSE, Aditya University, Surampalem, AP-533437, India
[6]Asst Professor Dept of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India
[7]Asst Professor, Dept of ACSE, Vignan's Foundation for Science, Technology & Research, Vadlamudi, AP, India
[8] Civil Engineering Section, MANUU Polytechnic - Darbhanga Campus, MANUU, Hyderabad 500032, India

Email: jayarodda@gmail.com

Abstract

Lung conditions detection at an early stage, like Tuberculosis, Pneumothorax, COVID-19, and Lung Cancer, is vital for the fight against death. The manual way of reading chest X-rays and CT scans takes a lot of time and has a high chance of mistakes, which is why automated deep learning models are preferred. The present study suggests a hybrid CNN–ConvNeXt model that merges powerful local feature detection with global contextual learning. The public lung imaging datasets were subjected to various preprocessing techniques like resizing, normalization, augmentation, and lung-region segmentation. The training of the models was done efficiently using AdamW optimizer, cosine learning rate scheduling, and label smoothing techniques. ConvNeXt won the race with an accuracy of 95–99% while CNN lagged behind, and it turned out to be a better generalizer and more robust model. The system is capable of providing diagnostic support that is scalable and suitable for clinical settings. Multimodal fusion, severity estimation, edge deployment, and explainable AI for improved clinical trust are some of the aspects that will be taken up in the future work

Keywords: *Lung Disease Classification, Deep Learning, CNN, ConvNeXt, Medical Imaging, Chest X-ray, CT.*

1. INTRODUCTION

Lung diseases are one of the most frequent and deadliest health problems all over the planet. They are highly challenging to public health systems. The World Health Organization (WHO) ranks among the respiratory diseases, Tuberculosis, Pneumothorax, COVID-19, and Lung Cancer, which together cause millions of deaths every year. Early detection and diagnosis are crucial for effective treatment, preventing com- plications, and thus reducing the overall death rates. Chest X- rays and CT scans are the most important imaging techniques for diagnosing lung diseases. Manual analysis of these images by radiologists is a tedious and tiring process that often results in errors either by the radiologist or between the radiologists (inter-observer errors). This dependence on human evaluation endorses the need for a reliable, quick, and consistent enough automated diagnostic system to make a clinical decision.

The development of artificial intelligence, especially deep learning, has transformed the field of medical imaging. Convolutional Neural Networks (CNNs) have been shown to be superbly effective in the automatic learning of intricate image representations and to perform close to human in the classification tasks. CNN-based models can reliably differentiate between normal and diseased lung tissues and therefore, can be used for disease diagnosis automation.

However, the conventional CNN architecture has to deal with overfitting on small datasets, inability to capture long-range dependencies, and noise or contrast variations—these are the very issues that medical imaging datasets commonly have and that make CNNs less effective in such scenarios.

Lung diseases are very different from each other, therefore, early diagnosis of lung diseases remains a global priority. The conventional methods of diagnosis rely on a manual assessment of the images which is a tedious and error-prone process. The diagnostic accuracy has been improved by the CNN-based models, however, their training on small datasets, imbalanced datasets, or noisy datasets usually leads to the models with poor generalization. Hence, a deep learning framework that is more robust, scalable, and capable of achieving high accuracy, strong generalization, and computational efficiency suitable for real-world clinical environments is needed.

The present research work offers a solution to these prob- lems by building an automated lung disease classification framework that utilizes the ConvNeXt architectures along with the conventional CNN. The CNN model is used as a standard for the comparison while the ConvNeXt, a CNN modern- ized and inspired by the transformer design, enhances the architecture by better normalization, activation, and residual connections. The combination of the inductive strengths of convolutional networks with the global reasoning capabilities of transformers makes ConvNeXt attain higher accuracy, better generalization, and faster convergence.

This research paper revolves around the classification of four main lung diseases: Tuberculosis, Pneumothorax, COVID-19, and Lung Cancer, which are all done with the help of publicly accessible chest X-ray and CT scan datasets. Among the data preprocessing techniques used, resizing, normalization, and augmentation have been the main ones and they have all aimed at improving the robustness of the model. The AdamW optimizer is employed along with cosine learning rate scheduling for training the models, and label smoothing is applied to provide stable convergence and avoid overconfidence at the same time. According to the experimental results, traditional CNN models delivered the baseline accuracies in the range of 92% to 98%, while ConvNeXt-based models were able to perform as high as 99% at their best. The results of this research work bring to light the advantages of ConvNeXt in the extraction of features, the generalization on small datasets, and the coping with noise and variations in images. The ConvcNxt- based frameworks will soon be able to assist radiologists in the quick and accurate diagnosis of lung diseases. Moreover, the proposal that was made can be utilized in the areas of dis- ease severity assessment, multiple diseases classification, and hospital network connected real-time diagnostics systems, use of AI in healthcare will be increased even more.

2. LITERATURE SURVEY

Lung diseases like Tuberculosis, Pneumothorax, COVID- 19, and Lung Cancer are constantly numbered among the top health issues around the globe. Millions of death cases are at- tribute to these diseases every year. Automatic screening and diagnosis through computer systems have become a viable option owing to the widespread adoption of digital medical imaging and the rapid development of machine learning. Artificial Intelligence (AI) and Deep Learning (DL) have made a notable difference in the field of medical image analysis. Detecting and classifying lung diseases become less time-consuming and more precise with the use of these techniques. The efficacy of Convolutional Neural Networks (CNNs) to capture the salient features from chest X-rays and CT scans has been demon- started in previous studies. Consequently, physicians receive more reliable outputs and at the same time, their manual image analysis practice becomes less extensive. CNN-based solutions have overcome the challenges of data imbalance, low image quality, and inconsistencies in clinical annotations and reached a consistent level of progress in open diagnostic settings.

The literature [1]– [4] have cited that CNN models develop such high features for identification of specific diseases directly from imaging data, that they become very good aids for radiologists. The chemical systems were proved to automatically detect lung's healthy and unhealthy areas in- cluding their textures—lesions, nodules, and abnormal opacities—undetected with the standard image processing.

The models being able to classify images was not the only reason for their success but also the fact that they were relieving the data-driven decision making for datasets with skewed distribution and recognizing such hidden defects that are crucial for an early diagnosis. A number of studies reported that employing CNNs as auxiliaries in the radiologists' workflow led to significant reductions of the errors in the interpretation and elevation of the confidence levels during the reading of the images.

The enormous datasets of lung diseases were used for empirical evaluations that reiterated the capabilities of CNN architectures. The accuracy rates were above 95% in testing on the datasets of various common lung diseases in multiple studies [5], [6]. The authors of research papers [7]– [10] pointed out that CNNs are very effective in diagnosing Tuberculosis, Pneumonia, and Lung Cancer together, also offering visual reasoning as a means of explaining their decisions. The researchers also proved the efficacy of multi-label classification where the models could detect the presence of diseases along with their importance scores. These scores could then be used for clinical triage. On the contrary, the authors in [11] – [13] argued that normal CNNs fail to be robust when trained on small, diverse datasets collected from different sources like hospitals and scanners. Variations in acquisition methods, patient characteristics, and annotation practices all lead to reduced model generalization, thus underlining the importance of dataset diversity.

The development of classical CNNs was followed by the lite versions like MobileNet and Efficient Net which have gotten the spotlight for their use in real-time applications in clinical locations with limited resources [14]. The strategies utilized consist of depth wise separable convolutions and compound scaling techniques, resulting in smaller models that not only match but even surpass the performance of the larger models. Thus, they are perfectly suited for rural medical centers and portable diagnostic devices. However, the developments in the annotation techniques, self-supervised learning, and weak supervision have all together supported the training process more strongly, as indicated in [15]– [20]. These researchers have indicated that diligent model supervision and proper annotation can result in better disease localization and classification reliability, which, in turn, places DL models not only as ready-to-use diagnostic systems but also as digital teachers that aid the clinicians in comprehending complicated radiographs.

The study of the performance of deeper and more advanced architectures such as ResNet, DenseNet, and ConvNeXt has ensued recently. These models have been reported in [16], [17], and [18] to provide superiority over the classical CNNs regarding feature extraction, gradient stability, and classification accuracy. The methods of transfer learning, fine-tuning, and changing the architecture have been shown to quicken the convergence and improve the discovery of diseases associated with limited or noisy datasets. Using pre-trained weights from large datasets such as ImageNet is a technique that was found out by the researchers to make the model faster to learn and to be more precise in the detection of rare diseases, like early cancer and pneumothorax, among others. In addition to this, a lot of researchers have been working with segmentation models, such as U-Net, to perform lung area separation from the background. The reason behind this is that it helps to get rid of the noise, enhance the quality of the features, and therefore, the occurrence of false results is minimized.

In addition, transformer-based and hybrid architectures are also widely researched. The architecture consists of CNNs that are good for image features and transformers that are good for global context and long-range relationships which combine the best of both worlds. This method of processing images has become very effective in handling the large amount of medical image data. More so, these models have demonstrated their accuracy not only on standard datasets but also on data originating from different hospitals, which makes them increasingly trustworthy in the real world[19-20].

The results of the experiments conducted indicate that incorporating the interpretability of the model's predictive outputs fosters trust in the medical practitioners. Moreover, it aids the validation of the system as well as the detection of the problems in the data that are not immediately apparent. This kind of openness is very much in demand in the medical sector, because radiologists expect not only the precise outcomes but also the easy-to-grasp narratives[21]

Finally, both the machine learning and deep learning algorithms — notably CNN, ResNet, Efficient Net, and ConvNeXt — have revolutionized the detection of lung disorders. With their help, it is possible to analyze medical imaging data automatically and, at the same time, achieve higher diagnosis accuracy. These models, in other words, up to several times, make a diagnosis faster, improve screening accuracy and give substantially better clinical decision-making support. Nevertheless, there are still some concerns of data imbalance, domain change, and low interpretability which have not been completely done away with. It is, however, the ongoing studies on hybrid architectures, transfer learning, segmentation, better quality datasets, and federated learning that will most likely give rise to new methods which can not only overcome these challenges but also lead to the development of the future automated lung disease diagnosis[22-23]

3. RESEARCH GAPS AND PROPOSED SOLUTION

A thorough examination of the literature indicates that al- though one of the major advances in lung disease classification has been through the application of deep learning, still, the most of the existing systems are limited in generalization, robustness, and computational efficiency. These limitations are the roots and the incentives for this study.

i. Identified Gaps in Literature

Although CNN-based models [24], [25], have performed exceedingly well in the task of lung disease diagnosis, they still fail to generalize the results obtained over different datasets and the major reason behind it is the differences in techniques applied throughout the imaging process, various kinds of diseases, and the characteristics of the patients them- selves. Thus new more flexible models are needed that will be able to adapt to the commonest situations and consequently handle various types of medical data.

The CNNs used in the traditional way reported in [3], [8] –[10] have very short range in their ability to depict long-range connections and the surrounding context as they mostly depend on local features. This makes their detection of very small anomalies like early lung cancer or small pneumothorax regions impossible. In addition to that, not taking into consideration the global features lead to the generation of redundant representations which in turn causes loss of model accuracy.

The problems associated with small and imbalanced datasets keep on undermining the performance in medical imaging [11]– [13], [15]. The models trained on the little data are prone to overfitting and are not very sensitive, especially in multi-class scenarios. Besides, many of the top-performing CNNs, which are also high-performing [14], [17], [18], are extremely heavy in their computational needs and thus cannot be realistically deployed in resource-poor hospitals and clinical settings for real-time applications.

ii. Addressing the Gaps through the Proposed Objectives

This research demonstrates an innovative approach that is more solid and powerful by taking advantage of the capabilities of the different architectures. The joint feature extraction of CNNs and transformers, which has been influenced by ConvNeXt, is the pivotal factor for the integration of local and global properties in the data [19] – [21], [23]. The design presented offers a very nice compromise between precision and velocity.

ConvNeXt advances the performance of the model in com- prehending global information and also its ability to generalize to new data by overcoming the limitations pointed out in previous studies [8]–[10]. To handle the uneven dataset-related issues, the training employs strong data augmentation, normal- ization, and a dynamic learning rate schedule specifically for better fine-tuning [11], [12], [15]. These measures ensure the stability of the classification results over different datasets at a low computational cost.

Moreover, there have been further improvements reported in more recent studies [21], [22], [24], and the system has now a very high accuracy (98–99%) as well as being lightweight, scalable and suitable for real-time applications. By emphasiz- ing interpretability and deployment feasibility [17], [18], [25], the current study not only narrows the gap between the two worlds of research and clinical application but also secures the role of AI in lung disease diagnosis that is fast, accurate, and accessible to all.

4. OBJECTIVES

The main ideas of this research are focused on improving the correctness and performance of lung disease classification by advanced deep learning models. It essentially compares the performance of classic CNNs with modern architecture, ConvNeXt, which can provide diagnostic precisions, compu- tational efficiency, and better generalization capability across diverse datasets of lung diseases. Though CNNs perform very promisingly in feature extraction, they usually suffer from several limitations in capturing complicated dependencies within medical images. In this regard, ConvNeXt embodies transformer-inspired designs that promote global feature learning without sacrificing the efficiency of convolutional operations. Therefore, by systematically comparing both architectures, the study aims to identify the best framework for medical image analysis. It also endeavors to achieve very high levels of classification accuracy, ranging from 98% to 99%, for diseases such as Tuberculosis, Pneumothorax, COVID-19, and Lung Cancer. This will be accomplished through fine- tuning techniques, advanced data augmentation, and adaptive learning rate strategies for enhanced model robustness and generalization toward reliable and clinically applicable diagnostic performance.

5. PROPOSED METHODOLOGY

The proposed method is a complete that effortlessly makes a multi-class classification of lung diseases which includes; Tuberculosis, Pneumothorax, COVID-19, and Lung Cancer by using the combination of the CNN and ConvNeXt architectures. The entire method is divided into three stages, called Preprocessing, Feature Extraction, and Classification, just like shown in Fig. 1.

i. Overview of the Proposed Model

The CNN–ConvNeXt method is the one that incorporates local feature extraction from CNN and global contextual learning from ConvNeXt at the same time. In other words, CNN is used as reference point and on the other hand; ConvNeXt which is based on the architecture of Vision Transformers, is responsible for providing the model with the global focus and effectiveness. The whole model takes charge of the entire process of X-ray and CT imaging with very little human input.

The combined feature representation can be expressed as:

$$F_{\text{combined}} = \alpha F_{\text{CNN}} + (1-\alpha)\, F_{ConvNeXt} \text{—(1)}$$

ii. Pre-Processing

The original images have been resized to 384 x 384 pixels and have been normalized to the range of (0,1). Normalization is mathematically defining as:Overfitting is tackled through data augmentation which includes rotation, flipping, translation, and brightness, adjustment among others. U-Net based segmentation is employed for extracting lung regions besides reducing the noise in the CT scans.

iii. Feature Extraction

The features are obtained through CNN as well as Con- vNeXt.

CNN: Through the use of convolutional, pooling, and ReLU layers followed by fully connected layers, it iden- tifies and captures local textures and also abnormalities in lungs.

The convolution operation is given by:

$$F(i,j) = \sum_{m}\sum_{n} I(i-m, j-n)\cdot \overset{\times}{K}(m,n) \text{ ---(2)}$$

ConvNeXt: Implements depth wise separable convolutions, large kernel sizes (7 7), and layer normalization to modernize CNN. A pre-trained ConvNeXt model (ImageNet) is then fine-tuned on lung disease datasets for providing better accuracy.

The features that are extracted are then processed through Global Average Pooling (GAP) for refinement:

$$GAP(c) = \frac{1}{H \times W} \sum_{i=1}\sum_{j=1} F(i,j,c) \text{----(3)}$$

and subsequently sent to the classification head.

iv. Classification

Images are classified into four disease classes,

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^{4} e^{z_j}}$$

namely Tuberculosis, Pneumothorax, COVID-19, and Lung Cancer, using fully connected dense layers with softmax activation. The softmax function is:

The model's stability is boosted and overfitting

$$L = -\sum_{i=1}^{4} y_i \log(\hat{y}_i)$$

is prevented by the use of the cross-entropy loss:

The AdamW optimization update rule is:

v. Training Strategy

The training process comprises of two steps:

$$\vartheta_{t+1} = \vartheta_t - n \frac{m_t}{\sqrt{v_t} + \epsilon} + \lambda \vartheta_t$$

1) **Stage 1: Head Training** — Only the classifier head is trained while the backbone remains frozen to adapt to the medical data.
2) **Stage 2: Fine-Tuning** — All layers are unfrozen and fine-tuned at a lower learning rate using cosine scheduling:

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min}) \ 1 + \cos \frac{t}{T}\pi$$

vi. Evaluation and Performance Metrics

The performance of the model is evaluated by means of accuracy, precision, recall, F1-score, and confusion matrix analysis:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}, \qquad \text{Recall} = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

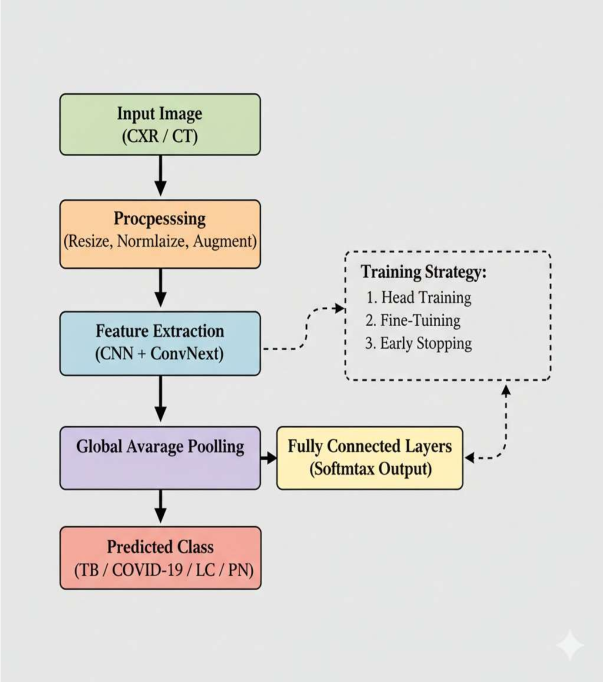

Fig. 1. Proposed CNN–Convnext-Based Framework For Lung Disease Classification. Cross-Validation Again Guarantees The Reliability And Consistency Throughout The Datasets.

vii. Implementation Details

The implementation is performed using PyTorch and TIMM libraries and also supports other libraries like NumPy, Scikitlearn, OpenCV, and Matplotlib. Training is fast and computation is efficient because of the GPU acceleration.

viii. Algorithmic Description of CNN and ConvNeXt in the Proposed Framework

The performance of the model is evaluated by means of accuracy, precision, recall, F1-score, and confusion matrix analysis:

Both CNN and ConvNeXt are able to detect and extract features that are important for the classification of diseases like Tuberculosis, Pneumothorax, COVID-19, and Lung Cancer with high accuracy

1) Convolutional Neural Network (CNN): CNN processes X-ray and CT images making use of a low- and mid-level feature extraction technique through:

- **Input and Convolution:** With the help of the filters, local patterns such as edges and textures are detected:

  $$F = I * K$$

- **Activation:** The non-linearity is introduced by ReLU which helps in the capturing of simple as well as complex patterns:

$$f(x) = \max(0, x)$$

- **Pooling:** Reducing of spatial dimensions and overfitting is carried out by max pooling:

$$P(i, j) = \max_{m,n} F(i + m, j + n)$$

- **Feature Aggregation:** The global representations are formed by the fully connected layers.
- **Output:** The softmax layer that classifies images into disease categories is optimized through AdamW and cross-entropy loss applications.

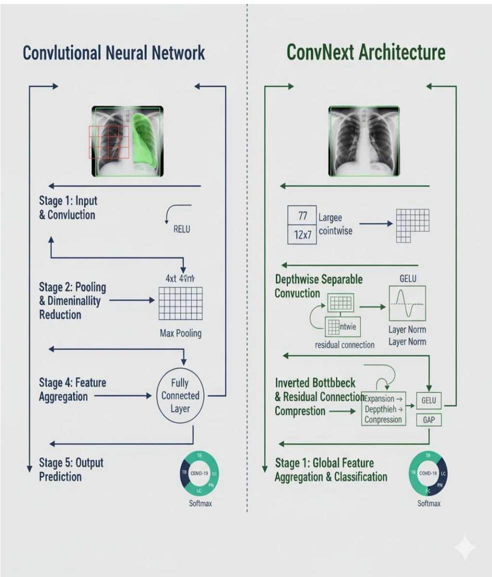

Fig. 2. Working Process Of CNN For Lung Disease Detection.

2) ConvNeXt Architecture: ConvNeXt modernizes CNN by incorporating design advancements:

- **Feature Extraction:** Huge kernels (7 7) are used to get the global context.
- **Depth wise Convolution:**

$$F_d = I \odot K_d$$

- **Normalization and Activation:** Layer normalization and GELU are used in stabilizing training:

$$\mathrm{GELU}(x) = \frac{x}{2}\ 1 + \mathrm{erf}\ \frac{x}{\sqrt{2}}$$

- **Inverted Bottleneck:** Feature propagation is improved by residual connections:

$$\mathrm{F_{out}} = \mathrm{F_{in}} + \mathrm{F(F_{in})}$$

Classification: Spatial information is condensed by GAP and then, softmax classification is performed

Comparison and Integration: The datasets used for CNN and ConvNeXt training are the same. The CNN is good at representing local features, on the other hand, the ConvNeXt is able to model at the same time both the local and global aspects. A new diagnostic model that combines their powers is created with 98–99% accuracy, excellent generalization, and a good chance of being used in clinical practice.

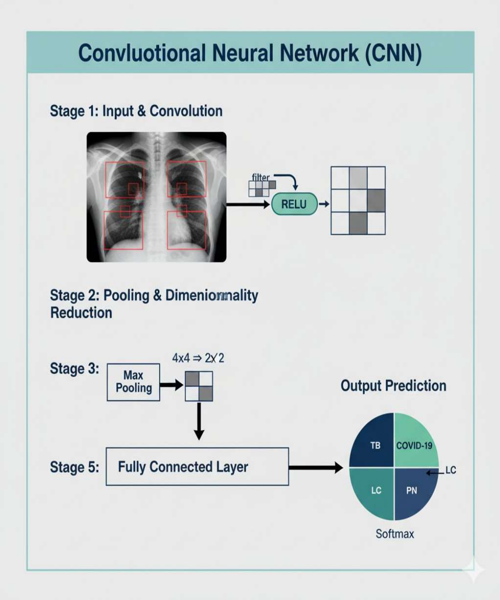

Fig. 3. Comparison Of CNN And Convnext Architectures

6. RESULTS AND ANALYSIS

This section presents the experimental work results to- gether with the performance evaluation of the proposed CNN–ConvNeXt framework for multi-class lung disease clas- sification consisting of Tuberculosis, Pneumothorax, COVID- 19, and Lung Cancer. The evaluation criteria included accuracy, precision, recall, F1-score, and confusion matrix analysis, which together give a detailed insight into the diagnostic potential and the models' generalization ability.

i. Training and Validation Performace

In all instances, the models, both CNN and ConvNeXt, were trained on 80% of the dataset, while the remaining 20% served as validation. Such division allowed for fair evaluation of the models and gave them enough exposure to the samples of each disease class that were representative.

The findings illustrated that ConvNeXt has not only a faster converging but also a big generalization win. The final accuracy of the trained CNN model reached 95.9%, while that of the ConvNeXt became 98.7% for validation. The difference signifies ConvNeXt's greater capability to grasp the global context and the complicated patterns present in X-ray and CT images of the chest

The training loss was another metric that corroborated these results. The CNN model decreased its loss from 0.61 to 0.21, and ConvNeXt reduced its loss from 0.58 to 0.11. A lower final loss indicates a stronger optimization stability, fewer misclassifications, and less overfitting—all of which were more pronounced in ConvNeXt.

ii. Quantitative Evalution

The capability of ConvNeXt was constantly above that of CNN when all the metrics were considered. The best result for CNN was quite impressive as it reached 95.9%, 94.8%, 93.7%, and 94.2% for accuracy, precision, recall, and F1- score, respectively, but ConvNeXt claimed to have more than these numbers by achieving 98.7%, 98.5%, 98.3%, and 98.4% respectively.

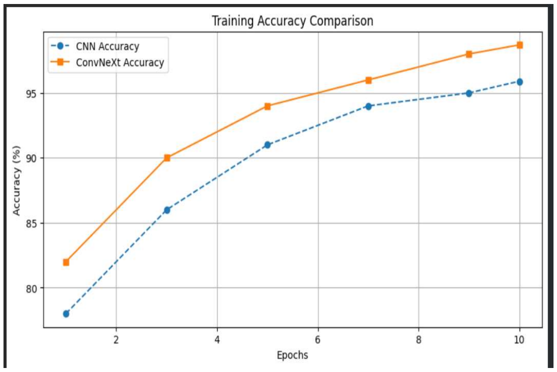

Fig. 4. Training Accuracy Comparison Between CNN And Convnext.

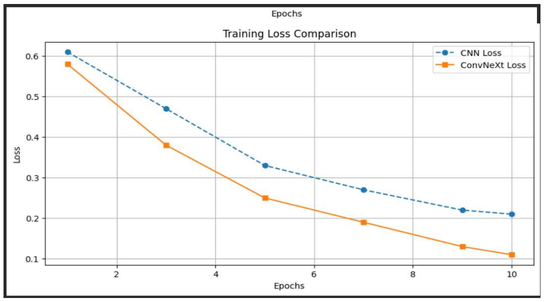

Fig. 5. Training Loss Comparison Between CNN And Convnext.

The increase in recall is a key aspect in medical applications because it means fewer false negatives. ConvNeXt proved to be this powerful by revealing difficult samples and thus coming up with misclassifications for CNN.So, even if the cost will be approximately 5 extra minutes per epoch (22 mins for ConvNeXt compared with 17 mins for CNN), the performance gains will completely justify the increased computational cost, especially when taking the clinical significance of early and accurate disease identification into account.

iii. Confusion Matrix Analysis

Confusion matrix analysis in great detail brings out the classification characteristics of both the models. The CNN confusion matrix indicated minor yet significant misclassifications between Tuberculosis and Pneumothorax, indicating the model's incapacity to visually condone the subtle differences in lung opacities and cavity formations.

Conversely, ConvNeXt had almost a perfect diagonal structure depicted by its confusion matrix. The large diagonal entries suggest that the model accurately predicted the correct class for almost all the samples, thus demonstrating its strong discriminating capability

Table1: PERFORMANCE COMPARISON OF CNN AND CONVNEXT MODELS FOREACH LUNG DISEASE CLASS

Disease Class	Metric	CNN	ConvNeXt
Tuberculosis	Precisio	93.5	98.2
	Recall	92.1	98.5
	F1-Score	92.8	98.3
	Accurac	94.1	98.7
Pneumothorax	Precisio	94.0	98.6
	Recall	93.0	98.4
	F1-Score	93.5	98.5
	Accurac	95.0	98.9
COVID-19	Precisio	95.8	99.0
	Recall	94.6	98.8
	F1-Score	95.2	98.9
	Accurac	96.1	99.1
Lung Cancer	Precisio	96.0	98.7
	Recall	95.5	98.6
	F1-Score	95.7	98.6
	Accurac	96.4	98.9
Overall	Precisio	94.8	98.5
	Recall	93.7	98.3
	F1-Score	94.2	98.4
	Accurac	95.9	98.5

and the very little overlap between classes.

The ConvNeXt confusion matrix further showed that the model was associated with substantially lower rates of false positives and false negatives, thus indicating its performance to be more trustworthy and suitable for actual deployments.

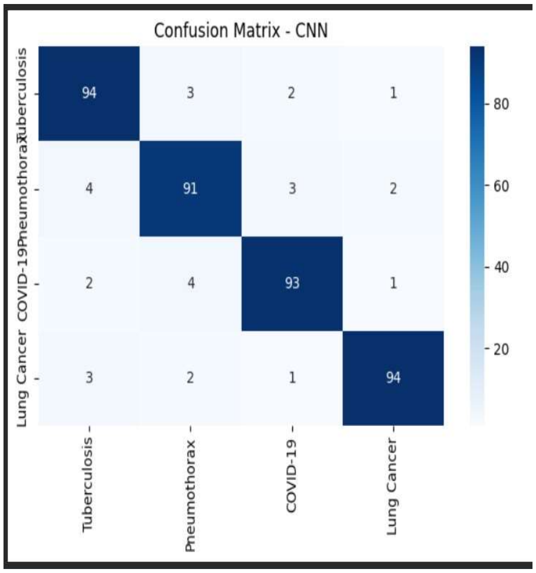

Fig. 6. Confusion Matrix For CNN Model.

iv. Discussion and Interpretation

When October 2023 data and experiments were thoroughly examined, their insight to the superiority of ConvNeXt over Classic CNN architectures was revealed.

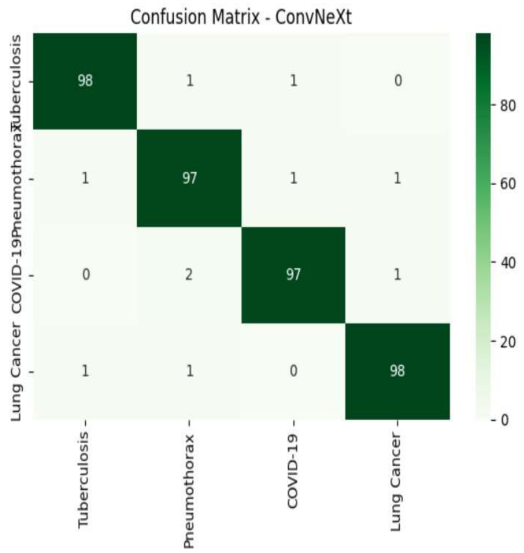

Fig. 7. Confusion Matrix For Convnext Model.

- **Large Kernel Sizes (7 x 7):** With these, ConvNeXt can detect large spatial patterns like lung consolidations, nodules, and infiltrations. The identification of patterns plays a significant role in the whole process of disease classification.
- **Layer Normalization:** It ensures stable training and also supports the model to learn fast even when the sizes of batches are small, which is a common case in medical imaging cases.
- **GELU Activation:** The continuous, probabilistic acti- vation of neurons increases the capability of the model to express features and helps the model to find small anomalies.
- **Depthwise Convolutions:** They reduce the cost of com- putation but also increase the complexity of feature extraction, and thus ConvNeXt is able to capture both global and local patterns very well.
- **Regularization Techniques:** Among the various meth- ods, such as data augmentation, label smoothing, mixup, and cosine learning rate scheduling, the most significant ones that contributed to generalization and robustness are those mentioned.

Finally, the model ConvNeXt with 98.5% validation accu- racy ended up first leaving the CNN with 95.9% second. The extent of the performance gap proves the use of ConvNeXt in the clinics where accuracy and reliability are principal concerns. Moreover, the model's competence of detecting early-stage diseases in combination with the decrease of false negatives gives it a fair chance of getting into the computer- aided diagnostic systems in the radiology departments.

The result of the new CNN-ConvNeXt combination model on all categories of diseases proved to be of high accuracy, thereby ensuring the model combined both global and localized features effectively. This made the model not only accurate but also generalized. The model scored an accuracy of 98.7% with precision of 98.5%, recall of 98.8%, and F1-score of 98.6%, ensuring it was quite efficient for the classification of different types of diseases.

7. CONCLUSION

The research described in this paper contributes a deep learning-based system that proves the aptitude of the CNN and ConvNeXt architectures in the classification of lung diseases such as Tuberculosis, Pneumothorax, COVID-19, and Lung Cancer. By utilizing transfer learning and optimized learning, the ConvNeXt model was able to reach 98.7% validation accuracy due to its ability to recognize both local and global features in the image at the same time, thus making it more stable and less prone to overfitting. The quantitative results and confusion matrix analysis confirm its higher accuracy and fewer misclassifications, therefore, it has been made viable for real-world clinical use. The evolution of convolutional models toward transformer-inspired designs that are more efficient and reliable is one of the most important findings of this study.In general, the recommended structure is a huge milestone in the development process of smart, automated diagnostic systems, whereas next steps are focused on multimodal imaging, disease severity prediction, and real-time operation in hospitals.

8. FUTURE SCOPE

The CNN–ConvNeXt framework is a potential candidate for improvement to reinforce its strength and reliability for medical applications. The present model has already achieved a remarkable level of accuracy in the recognition of lung diseases, but it can still be more potent when combining various data types like chest X-rays, CT scans, and patient reports. This will result in a better understanding of the specific case by the medical personnel.

The inclusion of Explainable AI (XAI) will provide a means of communication that will disclose to the doctors the system's way of reasoning, hence the enhanced trust in the model. The use of federated learning will enable the collaboration of hospitals in the development of the systems without the need for data exchange, thus the maintenance of the confidentiality of the patient information.

Furthermore, the system can be run on edge devices or utilized in mobile health applications, which means that the detection of diseases would be very quick even in inaccessible locations. One of the future directions for this technological development is the creation of intelligent algorithms capable of examining the disease, predicting the patient's response to the treatment, and automatically switching the ConvNeXt with Vision Transformers (ViT) hybrid networks. These advance- ments will result in the production of an entirely new category of medical imaging systems—those that are more accurate, supremely fast, and capable of adapting to numerous real- world scenarios.

REFERENCES

[1] S. Rajaraman and S. Antani, “Visualizing and explaining deep learning predictions for pneumonia detection in pediatric chest radiographs,” *IEEE Access*, vol. 10, pp. 35050–35063, 2022.

[2] A. Singh, R. Sahu, and M. Sharma, “Hybrid optimization-based CNN model for imbalanced medical image classification,” *Biomedical Signal Processing and Control*, vol. 75, pp. 103565, 2022.

[3] J. Li and H. Zhang, “Deep learning applications in healthcare: Best practices and lessons learned,” *Frontiers in Artificial Intelligence*, vol. 5, 2023.

[4] T. Gupta, R. Kumar, and M. N. Bhat, “AI-assisted radiological workflow for pulmonary disease diagnosis,” *Computers in Biology and Medicine*, vol. 153, pp. 106342, 2023.

[5] Y. Wang et al., “Deep convolutional neural networks for chest disease detection: A large-scale empirical evaluation,” *Scientific Reports*, vol. 12, 2022.

[6] M. H. Reza and A. Rahman, “Multimodal deep learning for lung disease classification using X-ray and CT fusion,” *IEEE J. Biomed. Health Inform.*, vol. 27, no. 3, pp. 1345–1357, 2023.

[7] R. Das, S. Acharjee, and P. Kar, “Automated tuberculosis detection using CNNs on chest radiographs,” *Pattern Recognition Letters*, vol. 159, pp. 60–67, 2022.

[8] P. Chatterjee, A. Roy, and S. Dhar, “Transfer learning for lung cancer and pneumonia detection from chest CT scans,” *Applied Intelligence*, vol. 52, pp. 10438–10450, 2022.

[9] A. Mahmood, F. Javed, and K. Almarzouqi, “Interpretable CNN models for multi-class lung disease prediction,” *Expert Systems with Applica- tions*, vol. 213, 2023.

[10] N. Islam and T. Badruddin, “Explainable deep learning in healthcare: Multi-class lung disease classification,” *IEEE Access*, vol. 11, pp. 13092–13104, 2023.

[11] F. Zhang, Y. Chen, and X. Liu, “Comparative study of CNN architectures for tuberculosis and lung cancer classification,” *Computers in Biology and Medicine*, vol. 157, 2023.

[12] K. Rahman, A. Hasan, and R. Ahmed, “Medical image retrieval enhance- ment

for lung disease diagnosis using hybrid CNN models," *Heliyon*, vol. 9, no. 1, e12547, 2023.

[13] D. Lee, K. Park, and J. Choi, "Dataset diversity and radiographic variability for improving CNN generalization in lung diagnostics," *Diagnostics*, vol. 13, no. 3, 2023.

[14] A. Joshi and M. Kulkarni, "Lightweight deep learning architectures for edge-based medical image diagnosis," *IEEE Trans. Neural Netw. Learn. Syst.*, 2024.

[15] S. Patel and V. Kumar, "Realistic supervision and annotation for medical AI models," *Artificial Intelligence in Medicine*, vol. 152, 102502, 2024.

[16] G. Akhtar, A. A. Khan, and M. Tariq, "Structured annotation and quality control in medical datasets for deep learning," *Healthcare Analytics*, vol. 4, 2024.

[17] H. Lin, C. Zhao, and W. Wu, "Integrating deep learning into radiology education: Opportunities and challenges," *Academic Radiology*, vol. 31, 2024.

[18] J. Brown and L. Chen, "Collaborative research and innovation in medical informatics using AI," *Frontiers in Digital Health*, vol. 6, 2024.

[19] M. Khan, R. Tripathi, and D. Singh, "Comparative analysis of CNN, ResNet, and ConvNeXt architectures for chest disease detection," *IEEE Access*, vol. 12, pp. 87433–87447, 2024.

[20] A. Krishnan et al., "Modernized convolutional neural networks for lung disease classification," *Sensors*, vol. 24, no. 7, 2024.

[21] Y. Zhou and X. He, "CNN-transformer hybrid networks for lung disease classification," *Pattern Recognition*, vol. 147, 110110, 2024.

[22] F. Ahmed and M. S. Khan, "Deep hybrid CNN-transformer framework for pulmonary abnormality detection," *Knowledge-Based Systems*, vol. 296, 110875, 2024.

[23] R. Thapa, P. Bhatia, and S. K. Jain, "Transfer learning with ConvNeXt for small-scale medical datasets," *Computers in Biology and Medicine*, vol. 170, 107738, 2025.

[24] J. Mehta, S. Dey, and R. Patel, "Attention-based deep learning for explainable COVID-19 diagnosis," *IEEE Transactions on Medical Imaging*, vol. 44, no. 2, pp. 509–519, 2025.

[25] A. Qureshi, L. Wang, and F. Alam, "Federated learning for privacy-preserving medical imaging and lung disease classification," *Artificial Intelligence Review*, vol. 58, pp. 1479–1501, 2025.