Electronic Letters on Computer Vision and Image Analysis 25(2):39-54, 2026

Lung Disease Classification Using Ensemble of CNN Models with Voting-Based Decision Fusion

Champa Tanga [1], Amarjit Roy [2], Jagdeep Rahul[1], Mohiul Islam [3], Chiranjit Sain [2], and Taha Selim Ustun [4]

[1] *Department of Electronics and Communication Engineering, Rajiv Gandhi University, Doimukh, Arunachal Pradesh, India*
[2] *Department of Electrical Engineering, Ghani Khan Choudhury Institute of Engineering and Technology, Malda, West Bengal, India*
[3] *School of Electronics Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India*
[4] *Fukushima Renewable Energy Institute, Koriyama, Japan*

Received 9 September 2025; revised 21 December 2025; accepted 22 April 2026

Abstract

Lung diseases such as pneumonia, tuberculosis, and COVID-19 continue to pose major threat to public health worldwide. For an effective clinical decision-making, it is still very important to make accurate and timely diagnosis using medical imaging. This paper suggests a voting-based automated ensemble framework that uses pretrained convolutional neural network (CNN) models to classify lung diseases from chest X-ray images. To make the diagnosis more accurate and reliable, three well-known CNN architectures such as ResNet50, AlexNet, and VGG19 are combined. This combination predicts using a majority voting strategy. A multi-stage classification framework based on transfer learning is employed to categorize lung images into four classes: normal, pneumonia, tuberculosis, and COVID-19. A dataset comprising over 900 chest X-ray images has been considered for performance evaluation. Experimental results show that the proposed ensemble model works better than individual CNN classifiers, with an overall classification accuracy of 92.30% and a sensitivity of 100%. Performance has been analyzed using different standard evaluation metrics such as accuracy, precision, specificity, sensitivity, and area under the curve (AUC). Comparative analysis with recent state-of-the-art techniques further validates the effectiveness of the proposed method. Comparative analysis shows that it can be used as a reliable computer-aided diagnostic framework for classifying lung diseases.

Key Words: Lung disease classification, Convolutional Neural Network (CNN), Ensemble Learning, Voting based decision fusion, transfer learning, Medical Image Analysis.

1 Introduction

Lung-related illnesses continue to represent a major global health issue, especially in less developed countries, where diseases such as lung cancer, tuberculosis, and pneumonia account for a substantial proportion of mortality. Medical imaging plays a vital role in life science research and serves as an essential accessory tool for clinical diagnosis. Common medical imaging modalities include magnetic resonance imaging (MRI), X-ray imaging, ultrasound, positron emission tomography (PET) and computed tomography (CT). In medical

Correspondence to: champa.tanga@rgu.ac.in

Recommended for acceptance by Angel D. Sappa
ELCVIA ISSN:1577-5097
Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

sciences, automatic identification of diseases is still a major challenge because of the large volume, complexity and higher dimensionality associated with the images involved in clinical diagnosis [1]. Consequently, current research trends in medical imaging show an extensive use of deep learning algorithms for developing computer-aided diagnosis (CAD) systems enabling advanced diagnosis, enhanced clinical decision-making, and prognostic assessment based on available data. Classification of medical images plays a critical role in clinical diagnosis and decision-making [3]. Prior to the adoption of deep learning, the classification of medical images predominantly depended on handcrafted features such as texture, color and shape descriptors [4]. Different conventional machine learning algorithms, such as support vector machines and logistic regression, were utilized to classify handcrafted features [5]. However, manual feature extraction and selection is a time consuming process and requires domain expertise. Moreover, the performance ceiling of handcrafted feature-based approaches has largely been reached [6]. CNNs have transformed medical image analysis by automating feature extraction and hierarchical representation learning from complex image data. CNNs excel at image categorization, detection, and segmentation. In several cases, CNNs have achieved or surpassed human-level accuracy[7, 8]. Single CNN models can learn a lot, but they may overfit, dataset bias, and be too sensitive to noise. This limitations can restrict CNNs practical applicability in real-world clinical settings. To address these limitations, ensemble learning techniques have emerged as a powerful alternative. Ensemble approaches reduce prediction variance and increase generalization by mixing basic learners. Voting-based ensemble approaches combine CNN model decisions in a simple yet effective way. They take use of the differences in network designs and training strategies. Due to these benefits, this study presents a voting-based CNN classification framework for lung image analysis. This improves diagnostic performance by aggregating the outputs of multiple CNN models. The proposed approach is validated using publicly available datasets. Compared to individual CNN designs, this work is more durable, sensitive, and accurate. The key contributions of this work are as follows:

- A voting-based ensemble classification framework is proposed that integrates multiple pretrained CNN models for accurate classification of lung related diseases in chest X-ray images.
- A multi-stage transfer learning strategy is employed using ResNet50, AlexNet, and VGG19 to classify lung images into four groups which include normal, COVID-19, tuberculosis and pneumonia.
- A majority voting-based decision fusion mechanism is introduced to enhance classification robustness by aggregating predictions from individual CNN classifiers.
- The effectiveness of the proposed framework is validated using different performance measures such as accuracy, specificity, precision, sensitivity, and AUC against various state-of-the-art methods.

Section 2 discusses about existing literature on deep learning based lung disease classification. Section 3 presents the proposed voting-based CNN framework and its underlying methodology. Results and discussions have been presented in Section 4. At last, Section 5 concludes the research work.

2 Related Work

Timely clinical treatment and diagnosis of lung-related diseases are necessary to reduce morbidity worldwide. Lung disorders such as COVID-19, lung cancer, tuberculosis (TB) and pneumonia are still a major problems for public health. These diseases are often associated with abnormal pathological changes in pulmonary tissues, including malignant cell proliferation. This can lead to fatal outcomes if undetected at an initial stage. A delayed diagnosis may further result in disease progression and its spread to other organs. As a result, it is now necessary to create precise and accurate computer-aided diagnostic (CAD) frameworks in order to properly categorize lung infections.

Medical imaging methods such as CT scans, chest X-rays and histopathological images are commonly used techniques for the automated lung disease analysis. Chest X-ray imaging is one of these methods that is utilized a lot because it is cheap and easy to get. In recent years, the automatic classification of lung disorders from medical images has significantly advanced with the adoption of different deep learning algorithms especially CNNs. CNN-based models have been used to detect lung illnesses in many studies. Some studies used single-network topologies, while others used ensemble-based techniques to increase diagnostic accuracy and generalization [9, 10, 11].

Early automated lung disease diagnosis research focused on TB detection utilizing chest X-rays. Lakhani and Sundaram[12] utilized pretrained CNN architectures, such as AlexNet and GoogLeNet, achieving diagnostic accuracy comparable to that of expert radiologists. Despite these excellent outcomes, these methods were tuberculosis-specific. TB was also difficult to distinguish from other lung illnesses with comparable radiological features. Pasa et al. [13] modified CNN topologies to improve tuberculosis detection in low-quality radiographs, making them more noise and image resistant. CNNs may be effective for identifying TB, but there are no clinical interpretability processes. Also, this method needs huge, well-labelled datasets to work well.

A lot of research has been done on CNN-based classification algorithms for identifying pneumonia. Kermany et al. [14] achieved classification accuracies surpassing 90% by utilizing transfer learning with the InceptionV3 architecture on pediatric chest X-ray images. But these kinds of models still depend a lot on how effectively the weights were put up before training. Rajpurkar et al. [15] proposed a 121-layer DenseNet model called as CheXNet. This model has the capability to detect pneumonia on par with radiologists. Subsequent experiments enhanced model comprehensibility and facilitated lesion identification by including attention mechanisms and multi-view learning techniques[16].

Lung cancer detection from CT and X-ray images is another critical area of investigation. Setio et al.[17] developed a multi-view CNN architecture that outperformed CAD methods for CT-based lung nodule diagnosis. However, the sensitivity of the method can be further improved by incorporating complete three-dimensional contextual information. To overcome this problem, Shen et al.[18] proposed a 3D CNN architecture for volumetric lung nodule categorization. This technique increased performance, but handcrafted image features may improve its clinical efficacy. Additionally, Causey et al.[19] demonstrated the diagnostic capability of CNNs in oncology by developing a deep learning model. This model is able to differentiate between benign and malignant lung nodules with high sensitivity. However, limited dataset size and the lack of cross-dataset validation restricted the generalizability of their findings.

Several deep learning algorithms have been proposed for early and accurate COVID-19 diagnosis. Apostolopoulos and Mpesiana [20] claimed a classification accuracy of 96.78% using transfer learning with the VGG19 architecture on chest X-rays. Afshar et al.[21] created a capsule network-based architecture with ensemble learning to improve generalization. Furthermore, Islam et al. [22] proposed a hybrid CNN–LSTM ensemble model to capture both spatial and sequential characteristics of COVID-19 progression. Despite showing that deep learning can enable real-time clinical decision support, several algorithms were limited by dataset imbalance and sample sizes, which hindered generalization and exacerbated overfitting.

Ensemble learning methods including stacking, weighted averaging, and majority voting have gained popularity for boosting classification reliability and resilience in CNN models. Zhou et al.[23] found that merging ResNet, DenseNet, and InceptionV3 architectures by majority voting improved COVID-19 CT image detection. Mahdy et al.[24] improved prediction stability by using a hybrid voting-based ensemble architecture for early COVID-19 identification. These methods demonstrate that ensemble-based CNN models can improve complex medical imaging diagnostic performance over single-network approaches.

Table 1 shows a brief overview of different methods for classifying lung diseases for better understanding.

Voting-based ensemble methods are particularly effective when the constituent models exhibit complementary behavior, as they mitigate the effects of overfitting and class imbalance. Ensemble and voting-based deep learning frameworks improve medical diagnosis classification accuracy and reliability [25, 26, 27, 28]. In applications like glaucoma detection, melanoma classification, retinopathy of prematurity screening, and lung lesion analysis, weighted or majority voting to integrate multiple CNN models addresses data imbalance and over-

Table 1: Comparative Analysis of Deep Learning Techniques for classifying Lung Diseases

Ref.	Authors & Year	Disease	Technique
[14]	Kermany, D. S. et al. (2018)	General medical diagnosis (including pneumonia)	Transfer learning using CNNs (Inception-V3)
[15]	Rajpurkar, P. et al. (2017)	Pneumonia detection	DenseNet-121 CNN (CheXNet)
[16]	Guan, Q. et al. (2018)	Thorax disease classification	Attention-guided CNN
[17]	Setio, A. A. A. et al. (2016)	Pulmonary nodule detection (CT)	Multi-view CNN
[18]	Shen, W. et al. (2017)	Lung nodule malignancy classification	Multi-crop CNN (MC-CNN)
[19]	Causey, J. L. et al. (2018)	Lung nodule malignancy prediction	Deep CNN + clinical features
[20]	Apostolopoulos, I. D. & Mpesiana, T. A. (2020)	COVID-19 detection	Transfer learning with pretrained CNNs (VGG19, MobileNetV2)
[21]	Afshar, P. et al. (2020)	COVID-19 detection	Capsule Network (COVID-CAPS)
[22]	Islam, M. Z. et al. (2020)	COVID-19 detection	Combined CNN-LSTM hybrid model

lapping disease characteristics. Waisy et al.[29] used a hybrid ensemble architecture to diagnose lung illness with CNNs and recurrent neural networks (RNNs) and achieved cutting-edge results across datasets. In real-world clinical environments, ensemble-based methods enhance the robustness, reliability, and interpretability of AI-driven diagnostic systems.

Despite these developments, voting-based ensemble frameworks for lung disease classification are still underdeveloped and underutilized. Addressing this research gap, the present study introduces a CNN-based voting ensemble method. The method integrates predictions from multiple CNN architectures trained on carefully curated datasets incorporating pneumonia, TB and COVID-19. The goal is to create a standard, high-performance, and clinically relevant categorization system that can accurately diagnose several lung diseases.

3 Proposed Voting based CNN model

This study presents a CNN-based voting ensemble lung disease classification system. The overall workflow of the proposed methodology is shown in Figure 4. After data collection and pre-processing, pretrained CNN architectures are chosen for transfer learning. Multi-stage categorization and voting-based decision fusion are used in the framework. Finally, the proposed approach is rigorously evaluated using standard image classification metrics. The transfer learning method uses multiple pretrained CNN models to address a multiclass classification issue.

a) **Data collection and pre-processing:** Chest X-ray images of pneumonia, TB, and COVID-19 are collected from various public sources. As CNN architectures require fixed input dimensions, the resizing of images are done to match the selected networks' inputs. To enhance dataset diversity and mitigate overfitting, data augmentation techniques are applied to increase training dataset size and variability.

b) **Pretrained CNN selection:** This study uses three popular pretrained CNN architectures: ResNet50, AlexNet, and VGG19. These networks are utilized to extract high-level discriminative features from

lung images during the training process. Thus, the model benefits from knowledge learned on large-scale benchmark datasets.

c) **Multi-stage classification:** Transfer learning adapts selected pretrained CNN models to the target problem and applies them in a multi-stage classification framework. Each model independently classifies input images as COVID-19, pneumonia, TB, or normal.

d) **Voting-based fusion:** A voting-based decision fusion strategy is adopted to combine the predictions obtained from the individual CNN classifiers. The final diagnostic decision is determined using a majority voting scheme. This enhances classification robustness by reducing the influence of erroneous predictions from any single model.

e) **Performance evaluation:** The presented algorithm is assessed using performance metrics like Specificity, sensitivity, accuracy and AUC. The comparative analysis of the obtained results are performed with those of state-of-the-art lung disease classification techniques. The analysis demonstrates the superiority and reliability of the presented approach.

3.1 Basic CNN Model

Figure1 demonstrates CNN's basic architecture. CNN, a fundamental form of deep learning, is derived from the idea of simple and complicated cells in the brain's visual cortex method demonstrated in [8]. CNN is associated with three layers such as pooling, convolutional and fully connected (FC). The fully connected layers are trained for the final classification after numerous alternating convolutional and pooling layers. CNN has shown to be very successful when employed in picture classification tasks.

The fundamental operation of CNN is expressed using the following mathematical formulation:

$$f(x) = \text{sign}\left[2\sigma_{\text{soft}}\left\{\omega_3\,\sigma_{\text{relu}}\left(\omega_2\,\sigma_{\text{relu}}(\omega_1 b_1) + b_2\right) + b_3\right\} - 1\right] \tag{1}$$

As indicated in Eq.1, the CNN's first convolutional layer and rectified linear unit (ReLU) operate on the input x. Here, the symbols represent the activation functions and b1, b2, and b3 denote corresponding bias functions.

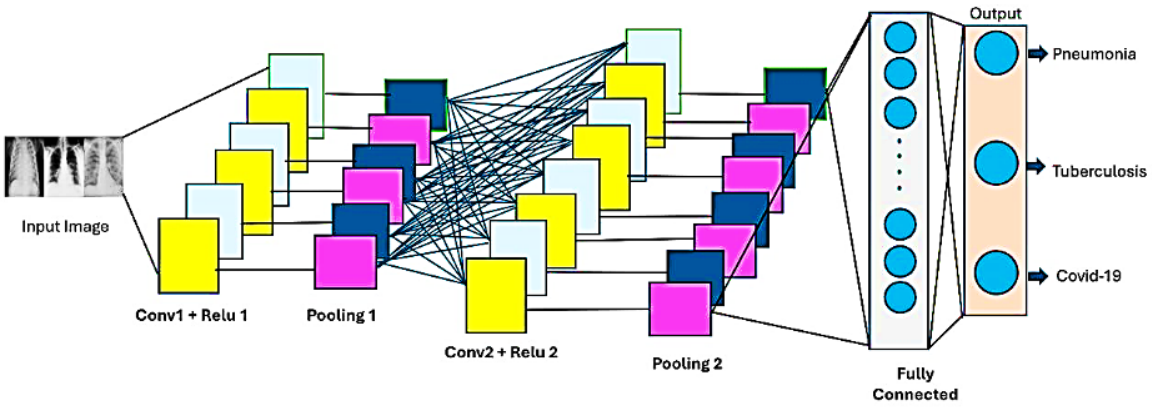


Figure 1: Basic architecture of CNN

3.1.1 Convolutional Layer

It is associated with three operations which are convolution, batch normalization and pooling. The convolutional filter output is followed by:

$$o_i = w_i\, x_i + b_i, \quad \text{for } i = 1, 2, 3, \ldots, N \tag{2}$$

In Eq.2, o_i denotes the output of the ith layer, while w_i, x_i, and b_i represent the weight, input, and bias, respectively, in the ith layer.

3.1.2 Batch normalization and Relu

It is a method used during the training of deep neural networks to standardize the inputs within each mini-batch. Normalization stabilizes and accelerates network training by decreasing internal covariate shift, enabling faster convergence and greater learning rates. Thus, batch normalization speeds up deep CNN model training and decreases epochs. The batch normalization operation applied during training is expressed as follows:

$$\hat{o}_i = \sigma\, \hat{x}_i + b_i, \quad \text{for } i = 1, 2, 3, \ldots, N \tag{3}$$

In Eq.3, σ and b_i denote the scale and shift parameters, respectively, and $\hat{x}_i$ represents the normalized input.

An activation function in a deep neural network is defined by the ReLU layer, through which non-linearity is introduced into the training model. In this function, positive input values are retained, while zero is assigned to negative input values. This process helps mitigate the vanishing gradient problem. A common generalized form of the ReLU function is given as follows:

$$f(x_i) = \max(o, x_i) \tag{4}$$

In Eq.4, x_i is depicted as normalized input subjected to relu layer.

3.1.3 Pooling Layer

After the convolutional layer, the pooling layer reduces input feature map spatial dimensions while keeping the most important information. This operation is performed independently on each depth channel of the input volume. During pooling, a two-dimensional kernel is systematically slid over each feature map to aggregate information within the receptive field, thereby summarizing local features and reducing computational complexity[30]. Max, average, and global pooling are used for feature abstraction and dimensionality reduction.

Let us consider the dimension of the feature vector as $N_h \times N_w \times N_c$. After passing through the pooling layer, the output dimension is given in Eq.5.

$$\left(\frac{N_h - F + 1}{S}\right) \times \left(\frac{N_w - F + 1}{S}\right) \times N_c \tag{5}$$

Here, N_h and N_w denote the height and width of the feature map, respectively, and N_c represents the number of channels. Moreover, F denotes the filter size, while S represents the stride length.

3.1.4 Softmax Layer

Following the final convolution layer comes the softmax layer, that generates the output classification's probability distribution function. The representation of softmax layer is represented in Eq.6.

$$\sigma(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{k} e^{x_j}}, \quad \text{for } j = 1, 2, \ldots, k \text{ and } i = 1, 2, \ldots, N \tag{6}$$

Here, k denotes the number of output classes, $\mathbf{x}$ represents the softmax input vector, and x_i denotes its elements. Each x_i of the input vector $\mathbf{x}$ is normalized by dividing by the sum of exponentials after applying the exponential function. The sum of the outputs of the softmax layer is given in Eq.7.

$$\sum_{i=1}^{k} \sigma(x_i) = 1 \tag{7}$$

3.1.5 Classification Layer

Each input image is classified in the classification layer based on the network's probability distribution. The size of the final layer before the classifier determines the output classes. Thus, the network architecture has a fully connected layer with four output nodes and a softmax activation function for the four target classes. To facilitate effective classification using deep neural networks, an average pooling layer is employed in place of the conventional final FC layer of the CNN, reducing model complexity and mitigating overfitting. This is succeeded by a completely connected layer combined with a softmax classifier. The proposed framework supports both binary classification (normal versus infected) using two output nodes. It also supports multiclass classification using four output nodes corresponding to normal, tuberculosis, pneumonia, and COVID-19 cases. By combining predictions from several CNN architectures with voting-based decision fusion improves classification robustness. The CNN models utilized in this study include ResNet50, VGG19, and AlexNet. Figure 2 shows the block diagram of the ResNet50 architecture utilized in the proposed framework.

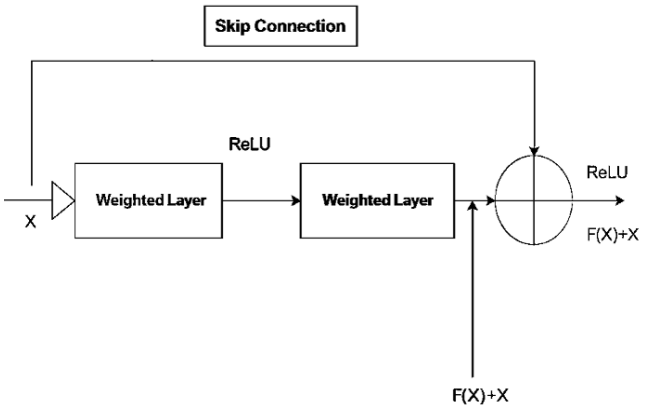


Figure 2: ResNet50 Block diagram

ResNet50 comprises of 48 convolutional layers, followed by single max-pooling layer and single average-pooling layer. The network accepts input images of size 224 × 224 pixels. The ImageNet dataset, which comprises more than a million tagged images covering 1,000 object categories, is used to pretrained ResNet50. Owing to its deep architecture and residual connections, ResNet50 is capable of learning rich hierarchical feature representations while mitigating the vanishing gradient problem.

The VGG19 network is also pretrained on the ImageNet dataset. The 19-layer architecture includes 16 convolutional layers, three FC layers, and a softmax classifier for prediction. VGG19 requires input images of size 224 × 224 pixels. It is noted for its consistent architecture and ability to capture the fine-grained spatial features through small convolutional filters [2].

AlexNet, an eight-layer deep CNN was also trained using ImageNet dataset. The architecture has three FC layers after five convolutional layers with max-pooling operations. Rectified Linear Unit (ReLU) activation algorithms introduce nonlinearity across the network, and dropout layers in fully connected layers prevent overfitting and increase generalization.

3.2 Incorporation of Classifier Fusion based network

In the proposed framework, three pretrained CNN architectures—ResNet50, VGG19, and AlexNet are initially employed as individual classifiers to detect and classify the different lung diseases. Subsequently, the decisions obtained from these individual classifiers are integrated using a classifier fusion (CF) strategy to generate the final diagnostic outcome. Classifier fusion combines classifiers outputs to improve overall prediction reliability and robustness. CF techniques vary in fusion structures and decision-combination mechanisms. In this study, a majority voting-based fusion technique is adopted due to its simplicity and effectiveness. This method aggregates CNN classifier outputs and assigns a class label based on the majority of votes. The output of the classifier fusion has been shown in Figure 3. Here, the output is fused in such a way that O1, O2, and O3 are outputs respectively from ResNet50, VGG19 and AlexNet. The flowchart of the proposed CNN voting based techniques for classification of lung images has been presented in Figure 4.

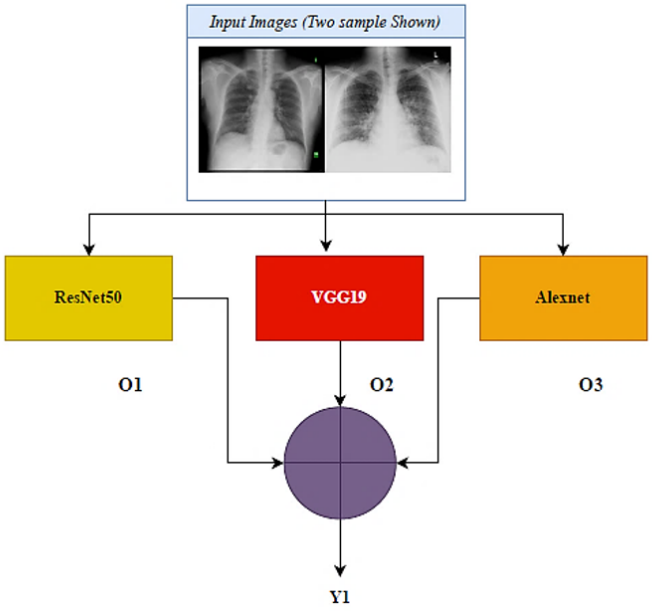


Figure 3: Classifier Fusion Technique

4 Results and Discussions

The proposed CNN-based voting ensemble classifier is evaluated on more than 900 color lung images, including pneumonia, tuberculosis (TB), and COVID-19[35, 36, 37]. All experiments are conducted in MATLAB R2024b on a Windows 10 Pro system equipped with an Intel® Xeon® W-2125 processor and 16 GB RAM.

The classification performance of the proposed framework is assessed using a confusion matrix and standard evaluation metrics, including accuracy, precision, sensitivity, specificity, and area under the curve (AUC). After applying data augmentation techniques to lung images, the results section quantitatively evaluates the classification performance of the model.

Figure 5 illustrates various types of data augmentation techniques applied to COVID-19 and TB infected chest X-ray images. Figure 5 shows data augmentation methods for COVID-19 and TB chest X-rays. To diversify lung images, data augmentation generates images with diverse orientations and transformations. This technique boosts dataset diversity, minimizes class imbalance, and improves trained model generalization by exposing them to more image variations.

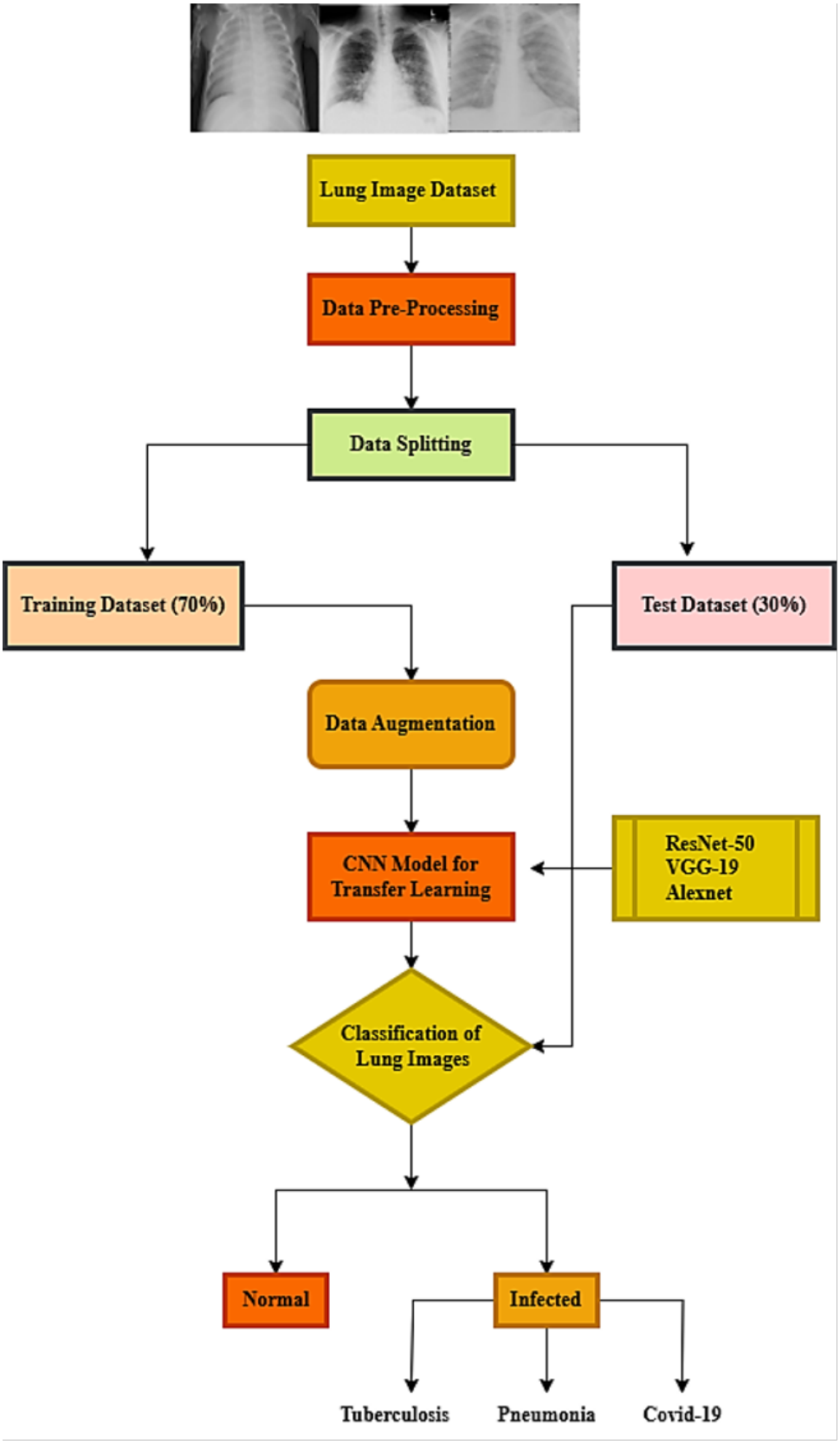


Figure 4: Flow chart of Proposed Technique

This section also provides an overview of the training parameters used during the model training process. For the four-class classification task, a batch size of 200 is employed. The adaptive moment estimation (Adam) optimizer is utilized to update the network weights. Multiple experiments are conducted using different numbers of training epochs to analyze their impact on classification performance. The classification accuracy achieved at various epoch values for each pretrained CNN model is summarized in Table 2. The results indicate that the optimal performance is achieved at 50 training epochs.

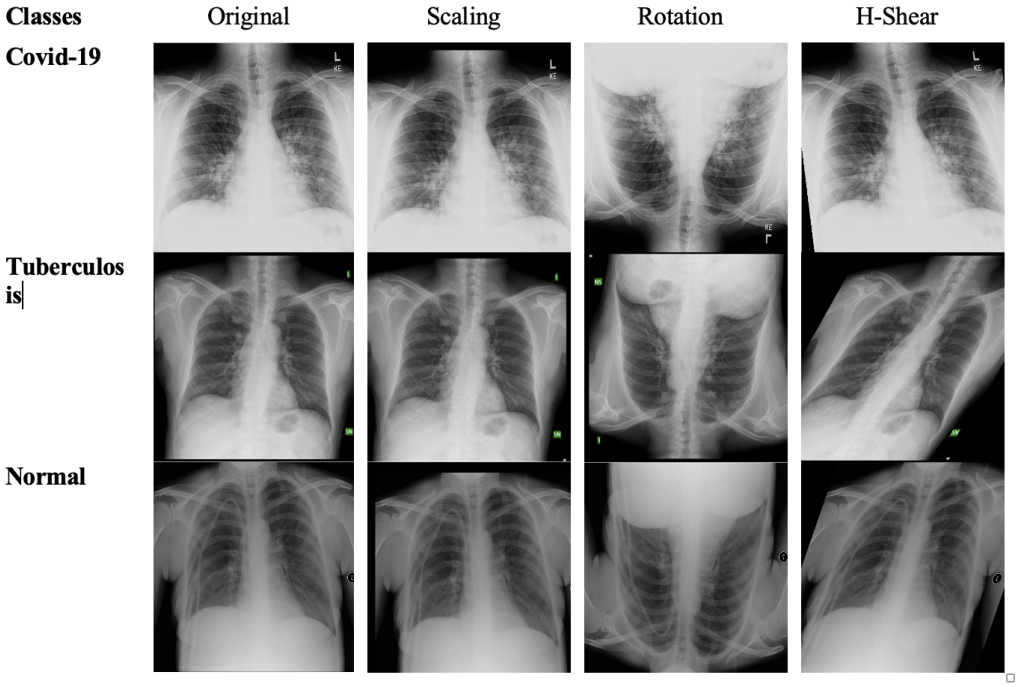


Figure 5: Different types of augmented images for Covid-19 and Tuberculosis Chest X-ray images.

Table 2: Accuracy for Different Epoch Values in terms of different pretrained models

CNN Model	Accuracy for Different Epoch Values				
	10	20	30	40	50
ResNet 50	87.55	87.82	88.10	88.54	89.85
AlexNet	87.20	87.53	87.71	88.08	89.30
VGG19	85.81	86.17	86.43	86.69	87.13

The performance comparison has been carried out using the following parameters, as shown in Eq. 8 - 11:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{9}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{10}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{11}$$

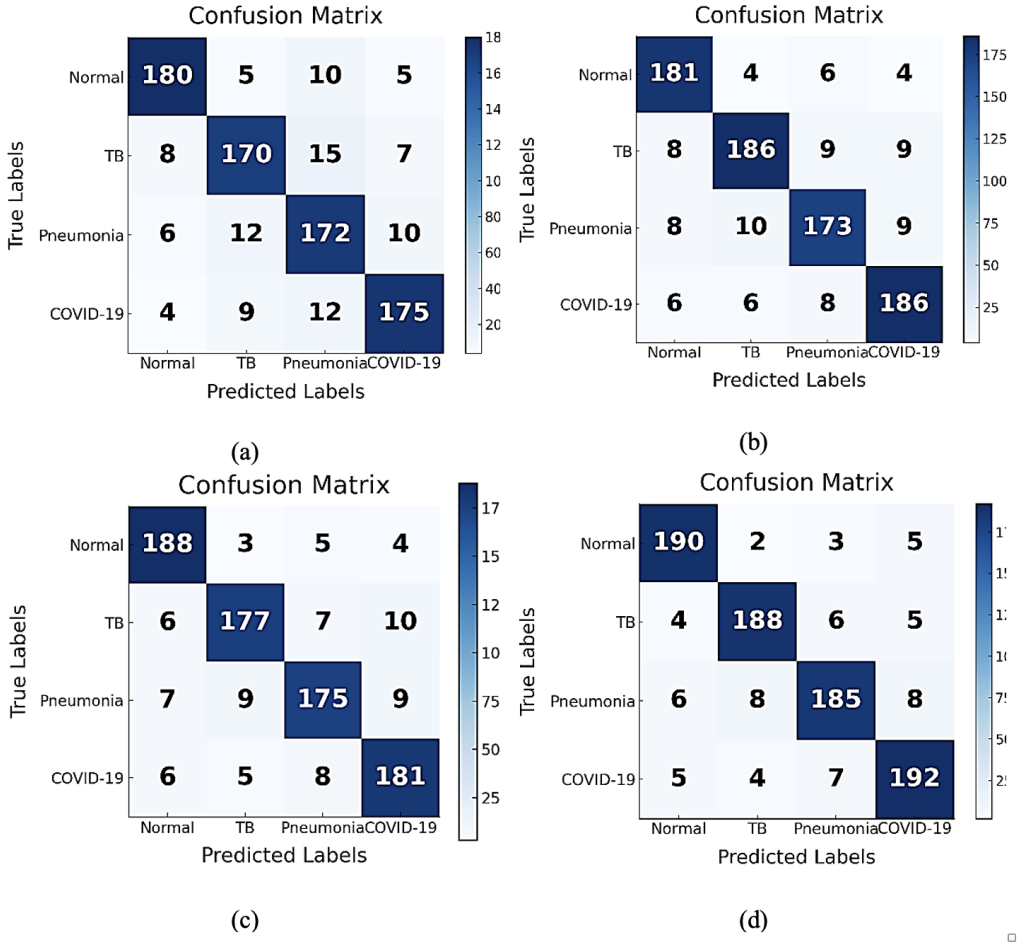


Figure 6: Confusion Matrix for Pneumonia, Tuberculosis, COVID-19, and Normal lung images using (a) VGG-19, (b) AlexNet, (b) ResNet-50 and (d) Classifier fusion

True positive (TP) denotes correctly classified infected lung images, while true negative (TN) shows correctly classified normal lung images. False positive (FP) corresponds to normal lung images that are incorrectly categorised as infected, whereas false negative (FN) refers to infected lung images that are incorrectly classified as normal.

Figure 6 presents the confusion matrices that classifies normal, pneumonia, tuberculosis and COVID-19 lung images. To facilitate comparative performance analysis, four separate confusion matrices are illustrated, corresponding to the classification results obtained using (a) VGG19, (b) AlexNet, (c) ResNet50, and (d) the proposed classifier fusion approach.

The confusion matrices presented in Figure 6 depicts the performance of the proposed classifier fusion method in comparison with individual transfer learning models, as summarized in Table 3. The proposed fusion-based classification framework achieves an improvement of approximately 5.62% in classification accuracy in contrast to VGG19 transfer learning model. Similar performance gains are also observed in terms of precision, sensitivity, and specificity. This demonstrates the efficiency of the classifier fusion strategy in enhancing overall diagnostic performance.

The AUC is an important metric for assessing the discriminative capability of a classification model across multiple image classes which includes pneumonia, tuberculosis, COVID-19, and normal cases. Figure 7 illustrates the AUC curve obtained using the ResNet50 model. An AUC value closer to unity indicates superior class separability and overall classification performance. The results demonstrate that ResNet50 achieves a high

Table 3: Performance metrics for different pretrained models

Model	Performance Metrics (In terms of Percentage)				
	Accuracy	Precision	Specificity	Sensitivity	AUC
VGG19	87.12	98.67	95.71	98.62	99.08
AlexNet	89.30	98.53	96.43	98.54	99.03
ResNet50	89.75	98.90	96.58	98.88	99.25
Proposed Classifier Fusion	92.30	100	97.43	100	99.11

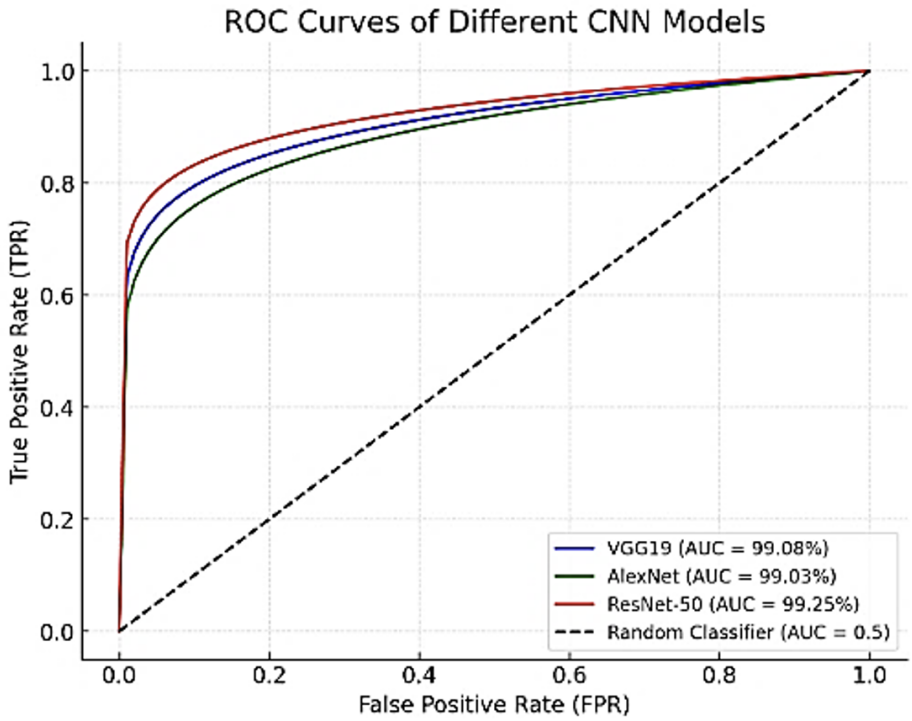


Figure 7: AUC curve for ResNet-50

AUC value, reflecting its strong capability to effectively distinguish among the different lung disease categories.

The proposed classifier fusion framework demonstrates superior performance over several state-of-the-art lung disease classification algorithms, as summarized in Table 4. By integrating three pretrained CNN architectures i.e., VGG19, AlexNet, and ResNet50 into a voting-based ensemble, the framework effectively exploits complementary feature representations learned at different network depths. This fusion strategy enhances robustness and reduces the limitations associated with single-model approaches. As a result, the classification accuracy of the presented method achieves a classification accuracy of 92.30% and a sensitivity of 100% for a challenging four-class problem involving normal, COVID-19, TB, and pneumonia cases.

The experimental results demonstrate the superiority of the proposed approach in terms of performance in comparison to the existing literature. Rahimzadeh et al.[32] used ResNet50V2 to classify three classes (Covid-19, pneumonia, and normal) with 91.40% accuracy and 80.53% sensitivity. Similarly, Hussain et al.[34] used XGBoost, KNN, Naive Bayes, and CART for a four-class issue and achieved 89.78% accuracy and 89.79% sensitivity. The suggested fusion model not only does better than these results in terms of accuracy and sensitivity, but it also solves a more difficult multi-class problem. Furthermore, Das et al.[33] also used Inception, DenseNet201, and ResNet50V2 for a binary classification test (Normal vs. COVID-19). It reported an accuracy of 91.62%. Goyal et al. [31] proposed an F-RRN-LSTM model with 87.20% accuracy and sensitivity on a

three-class dataset. This method outperforms conventional methods in a more difficult four-class classification. This outcome illustrates the efficacy of the ensemble fusion approach.

The improvement in performance can be attributed to the ensemble's ability to combine diverse feature representations, reduce model variation, and make similar sickness patterns more distinct. The Voting-based fusion guarantees crucial case detection by balancing sensitivity and specificity. So, the suggested classifier fusion system offers a strong and clinically useful way to automatically diagnose lung illness.

Table 4: Comparison with state-of-the-art techniques

Authors	Methodologies	Classification Dataset	Accuracy %	Sensitivity %
Rahimzadeh et al. [2020] [32]	ResNet50V2	3 Class (Normal, Pneumonia, Covid-19)	91.40	80.53
Hussain et al. [2020] [34]	XGBoost Linear, XGBoost, k-NN, CART, Tree Naive Bayesnet	4 Class (Covid-19, VP, BP, Normal)	89.78	89.79
Das et al. [2021] [33]	Inception + DenseNet201 + ResNet50V2	2 Class (Normal, Covid-19)	91.62	91.62
Goyal et al. [2023] [31]	F-RRN-LSTM	3 Class (Normal, Pneumonia, Covid-19)	87.20	87.20
Proposed Classifier Fusion	Pretrained (VGG19 + AlexNet + ResNet50)	4 Class (Normal, Covid-19, Tuberculosis, Pneumonia)	92.30	100

5 Conclusions and future scopes

This study developed a CNN-based voting ensemble framework utilizing chest X-ray images for the multiclass classification of lung diseases. These include pneumonia, tuberculosis (TB), COVID-19, and normal cases. Through a majority voting mechanism, the proposed method integrates three pretrained deep learning architectures such as ResNet50, AlexNet, and VGG19 to improve diagnostic accuracy and stability. Experimental results showed that the suggested fusion model significantly outperformed individual CNN classifiers and other state-of-the-art methods in accuracy, sensitivity, and specificity. These results demonstrate ensemble learning's effectiveness in improving medical image analysis diagnostics, generalization, and resilience. The proposed method demonstrates consistent and reliable performance improvements across most evaluation metrics.

The evaluation of the model using large and more varied datasets may be the main goal of future study. It may also be extended it to encompass additional respiratory disorders, and optimizing it for real-time implementation in clinical environments. Future work may be extended to bigger and more varied datasets for improved generalization, incorporating additional lung diseases such as COPD and lung cancer, and enhancing interpretability through explainable AI techniques. Moreover, optimization for real-time clinical deployment and integration into healthcare systems could further strengthen its practical applicability in medical diagnostics.

DECLARATIONS

Ethics approval and consent to participate

Not applicable (public datasets are used in this study).

Consent for publication

Not applicable (public datasets are used in this study).

Availability of data and materials

Data used for the findings are publicly available at:

"Chest X-ray images (pneumonia)," https://www.kaggle.com/paultimothymooney/chest-xray-pneu accessed: 2024-10-03.

"Tuberculosis (tb) chest x-ray dataset," https://www.kaggle.com/tawsifurrahman/tuberculosis-tb- accessed: 2024- 10-03.

"Covid-19 chest x-ray dataset," https://github.com/ieee8023/covid-chestxray-dataset, accessed: 2024-10-03.

Competing interests

Authors declare no competing interests.

Funding

This project received no funding.

Authors' contributions

Conceptualization, Investigation, Writing – initial draft, Writing – review and editing; C.T., A.R., J.R., M.I., C.S., H.A.S., T.S.U. Acknowledgements Not applicable

References

[1] G. Litjens et al., "A survey on deep learning in medical image analysis", *Medical Image Analysis*, 42:60–88, 2017.

[2] M. K. Chaitanya et al., "Artificial intelligence based approach for categorization of COVID-19 ECG images", *Biomedical Physics & Engineering Express*, 9(3):035012, 2023.

[3] M. M. Adnan et al., "Automated image annotation using deep ResNet50-SLT", *IEEE Access*, 2023.

[4] N. Dong et al., "Feature fusion based deep learning for white blood cell classification", *Journal of Ambient Intelligence and Humanized Computing*, 2022.

[5] S. S. Devi et al., "Malaria infected erythrocyte classification using hybrid classifier", *Multimedia Tools and Applications*, 77(1):631–660, 2018.

[6] S. S. Yadav and S. M. Jadhav, "Deep CNN based medical image classification", *Journal of Big Data*, 6(1), 2019.

[7] T. Mahmood et al., "Advancements in deep learning for medical image analysis", *IEEE Access*, 2023.

[8] Y. Guo et al., "Deep learning for visual understanding: a review", *Neurocomputing*, 187:27–48, 2016.

[9] Y. Gordienko et al., "Deep learning for chest X-ray analysis", *Advances in Computer Science for Engineering and Education*, pp. 638–647, Springer, 2019.

[10] W. Sun and B. Zheng, "Computer aided lung cancer diagnosis", *Proc. SPIE Medical Imaging*, pp. 241–248, 2016.

[11] D. Narin and T. Ö. Onur, "Hyperparameter effects in lung cancer classification", *Erzincan Univ. J. Sci. Technol.*, 15:258–268, 2022.

[12] P. Lakhani and B. Sundaram, "Deep learning for tuberculosis classification", *Radiology*, 284(2):574–582, 2017.

[13] F. Pasa et al., "Deep network architectures for tuberculosis screening", *Scientific Reports*, 9(1):6268, 2019.

[14] D. S. Kermany et al., "Medical diagnosis using deep learning", *arXiv preprint*, 2018.

[15] P. Rajpurkar et al., "CheXNet pneumonia detection", *arXiv:1711.05225*, 2017.

[16] Q. Guan et al., "Attention guided CNN for thorax classification", *arXiv:1801.09927*, 2018.

[17] A. A. A. Setio et al., "Pulmonary nodule detection using CNN", *IEEE Transactions on Medical Imaging*, 35(5):1160–1169, 2016.

[18] W. Shen et al., "Multi-crop CNN for lung nodule classification", *Pattern Recognition*, 61:663–673, 2017.

[19] J. L. Causey et al., "Prediction of lung nodule malignancy", *Scientific Reports*, 8(1):9286, 2018.

[20] I. D. Apostolopoulos and T. A. Mpesiana, "COVID-19 detection using CNN", *Physical and Engineering Sciences in Medicine*, 43:635–640, 2020.

[21] P. Afshar et al., "COVID-CAPS framework", *Pattern Recognition Letters*, 138:638–643, 2020.

[22] M. Z. Islam et al., "CNN-LSTM for COVID-19 detection", *Informatics in Medicine Unlocked*, 20:100412, 2020.

[23] T. Zhou et al., "Ensemble deep learning for COVID-19 CT images", *Applied Soft Computing*, 98:106885, 2021.

[24] L. N. Mahdy et al., "COVID-19 classification using SVM", *MedRxiv*, 2020.

[25] H. Wang et al., "Voting ensemble for pulmonary lesion classification", *Computers in Biology and Medicine*, 197:111064, 2025.

[26] A. Ennaji et al., "Weighted voting classifier for melanoma diagnosis", *Multimedia Tools and Applications*, 84(9):6073–6089, 2025.

[27] S. S. Krishna and S. S. Reka, "Voting classifier for retinopathy detection", *International Journal of Computational Intelligence Systems*, 18(1):238, 2025.

[28] S. J. Edan and M. R. Feizi-Derakhshi, "Glaucoma classification using CNN fusion", *Baghdad Science Journal*, 22(8):2813–2823, 2025.

[29] A. S. Al-Waisy et al., "COVID-DeepNet hybrid model", *Computers, Materials and Continua*, 67(2):2409–2429, 2021.

[30] V. K. Velpula et al., "Glaucoma detection using explainable AI", *IET Image Processing*, 18(13):3827–3853, 2024.

[31] S. Goyal and R. Singh, "Detection of lung diseases using ML and DL", *Journal of Ambient Intelligence and Humanized Computing*, 14(4):3239–3259, 2023.

[32] M. Rahimzadeh and A. Attar, "Modified CNN for COVID-19 detection", *Informatics in Medicine Unlocked*, 19:100360, 2020.

[33] A. K. Das et al., "COVID-19 detection using ensemble CNN", *Pattern Analysis and Applications*, 24:1111–1124, 2021.

[34] L. Hussain et al., "Chest X-ray classification using ML", *BioMedical Engineering OnLine*, 19:1–18, 2020.

[35] "Chest X-ray images (pneumonia)", `https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia`, accessed: 2024-10-03.

[36] "Tuberculosis chest X-ray dataset", https://www.kaggle.com/tawsifurrahman/tuberculosis-tb-chest-xray-dataset, accessed: 2024-10-03.

[37] "COVID-19 chest X-ray dataset", https://github.com/ieee8023/covid-chestxray-dataset, accessed: 2024-10-03.