JRITM
JOURNAL OF RESEARCH AND INNOVATION IN TECHNOLOGY, COMMERCE AND MANAGEMENT

Journal of Research and Innovation in Technology, Commerce and Management
Vol. 3 Issue 9, September 2026, pp. 39040-39045
ISSN: 3049-3129 (Online)

An Architecture-Aware Deep Transfer Learning Framework for Robust Multi-Class Lung Disease Classification

S.Nandhinidevi
Department of Computer Science
Thiruvalluvar Government Arts College
Periyar University, Tamil Nadu
Email: nandhinidevi.ctug@kongu.edu

R. Dhanushmati
Department of Computer Technology-UG
Kongu Engineering College, Tamil Nadu
Email: dhanushmatir.24bcr@kongu.edu

G. Rajendran
Department of Computer Science
Government Arts and Science College, Tamil Nadu
Email: guru.ragendran@yahoo.com

Cite as: S.Nandhinidevi, R. Dhanushmati, & G. Rajendran. (2026). An Architecture-Aware Deep Transfer Learning Framework for Robust Multi-Class Lung Disease Classification. Journal of Research and Innovation in Technology, Commerce and Management, Vol. 3(Issue 9), 39040–39045. https://doi.org/10.5281/zenodo.22689422

DOI:https://doi.org/10.5281/zenodo.22689422

***Abstract*—** The accurate classification of lung diseases through medical imaging like X-Ray and CT-scan is vital for early diagnosis and clinical decisions. Deep learning and transfer learning have shown significant ability to automating this classification process, although the effectiveness of various pretrained models can differ based on the dataset's characteristics. This study employs three pretrained deep learning models viz ResNet50, and DenseNet121 and Vision Transformer (ViT-B/16) to perform the multi-class lung disease classification using X-Ray images. The performance of all the models is measured using various metrics such as accuracy, precision, recall, F1-score, and validation accuracy to identify the best model for the lung disease classification. From the results of these models, it is found that DenseNet121 outperforms than the others by achieving accuracy of 92.49%, precision of 92.43%, recall of 92.49%, and F1-score of 92.45%. ResNet50 also shows good results, with accuracy of 92.05% and F1-score of 91.97%, while ViT-B/16 reaches an accuracy of 88.18%. These findings suggest that DenseNet121 and ResNet50 are more suitable for lung disease classification tasks than ViT-B/16.

Keywords— Lung disease classification, deep learning, transfer learning, Vision Transformer, ResNet50, DenseNet121, medical image classification, federated learning.

I. INTRODUCTION (*Heading 1*)

Lung diseases, including pneumonia, tuberculosis, COVID-19, and other pulmonary disorders, continue to pose serious challenges to healthcare systems worldwide. Early and accurate diagnosis is essential for selecting appropriate treatment and improving patient outcomes. Chest X-ray (CXR) is one of the most commonly used imaging techniques for examining lung conditions because it is easily accessible, relatively inexpensive, and can be obtained quickly. However, interpreting CXR images manually can be difficult, particularly when different diseases produce similar visual patterns or when abnormalities are not clearly visible. These challenges have led to growing interest in artificial intelligence (AI)-based computer-aided diagnosis systems. Deep learning

ISSN: 3049-3129 (Online)

methods, especially convolutional neural networks (CNNs), have shown considerable potential in this area by automatically learning meaningful features from medical images. Models such as ResNet and DenseNet121 have been widely used because of their ability to learn effective hierarchical representations through residual and dense feature connections. More recently, Vision Transformers (ViTs) have gained attention in medical image analysis because they divide an image into smaller patches and use self-attention mechanisms to learn relationships between different regions of the image.

Transfer learning has further improved the application of deep learning to medical image classification by using pretrained models to be adapted to specific medical imaging tasks. Models which are trained initially on large datasets such as ImageNet can offer useful features and reduce the time of training. Nevertheless, comparing the performance of existing lung disease classification models remains challenging, as previous studies often use different datasets, disease categories, preprocessing techniques, data-splitting strategies, augmentation methods, and optimization settings. As a result, it is difficult to determine whether the reported performance differences are primarily due to the model architecture or the experimental conditions. A systematic comparison under a common experimental setup is therefore important for understanding the relative strengths of CNN-based and Transformer-based models.

In this study, a comparative transfer learning approach is developed for the classification of five lung disease classes from chest X-ray images using ViT-B/16, ResNet-50, and DenseNet-121. In the following section, the existing research are discussed in Chapter II, dataset and model architecture are explored in chapter III , following that the result of ever model is explained and compared in Chapter IV and finally the conclusion of the proposed work is represented in the Chapter V.

II. Literature Review

This survey detailed the influence of deep learning frameworks, ranging from basic convolutional neural networks to advanced architectures like ResNet and DenseNet in classification of lung diseases automatically.

Vardhan et al. [1] proposed a deep transfer learning framework for estimating lung opacity from chest X-ray images. The study employed VGG-16, ResNet-50, and CheXNet-121 using different segmentation and data-balancing strategies. All the models are trained with 30,586 out of 38,365 chest radiographs images. The best five-fold model achieved a precision of 65.89%, recall of 66.30%, F1-score of 65.82%, and macro-averaged MAE of 0.3944 using CheXNet-121 with double-stage sampling and no ROI segmentation. The study demonstrated the importance of preprocessing and class-balancing strategies in addition to network architecture.

Shamrat et al. [2] proposed MobileLungNetV2, a fine-tuned MobileNetV2 model for multi-class lung disease classification using the ChestX-ray dataset. Images are pre-processed with CLAHE, Gaussian filtering, and data augmentation. The initial MobileNetV2 achieved 91.6% accuracy, while the fine-tuned MobileLungNetV2 achieved 96.97% accuracy, with overall precision, recall, and specificity of 96.71%, 96.83%, and 99.78%, respectively.

Al-Sheikh et al. [3] proposed three models viz customized CNN, AlexNet, and VGG16 to classify COVID-19, pneumonia, and normal cases using both chest X-ray and CT images. The proposed customized CNN achieved 98.60% accuracy, 98.40% sensitivity, and 98.50% specificity on the X-ray dataset, while the CT dataset produced 98.80% accuracy, 98.50% sensitivity, and 98.40% specificity.

Singh et al. [4] studied a Vision Transformer for pneumonia detection from chest X-ray images. The model utilized self-attention for capturing global relations among image regions. The proposed model trained and tested on a publicly available Kaggle chest X-ray pneumonia dataset was employed. The proposed ViT framework achieved 97.61% accuracy, 95% sensitivity, and 98% specificity. This study illustrates the capability of transformer-based architectures to capture global contextual information, which may be challenging to model with traditional CNNs.

Ko et al. [5] employed different Vision Transformer architectures like ViT, FastViT, and CrossViT integrated with various optimization methods viz. using Adam, AdamW, NAdam, RAdam, SGDW, and Momentum optimizers for lung disease prediction. The study used 19,003 chest X-ray images covering COVID-19, viral pneumonia, bacterial pneumonia, MERS, SARS, tuberculosis, and normal cases. In the evaluation of the models, ViT with RAdam achieved 95.87% accuracy on the balanced dataset, whereas FastViT with NAdam achieved the best accuracy of 97.63% on the imbalanced dataset.

Li et al. [6] proposed a Multi-Branch Residual Attention Network (MBRANet) for multi-label thoracic disease classification. The architecture integrates a ResNet-50 backbone with coordinate attention, employs feature pyramid-based multi-scale feature fusion, and utilizes a multi-branch feature classifier incorporating class-specific residual attention. The evaluation of the model was conducted using the datasets ChestX-Ray14, CheXpert, MIMIC-CXR, and IU X-Ray. The mean AUC values reported for these datasets were 0.841, 0.895, 0.805, and 0.745, respectively. Specifically, for the ChestX-Ray14 dataset, ResNet-50 was determined to be the most effective backbone, with a mean AUC of 0.841.

Bennour et al. [7] developed three deep learning frameworks—CovCXR-Net, MDCXR3-Net, and MDCXR4-Net to classify images into 4 categories such as COVID-19, pneumonia, pulmonary opacity, and normal. The research utilized datasets including the COVID-19 Image Data Collection, Chest X-Ray Images Pneumonia dataset, and COVID-19 Radiography Database. In the experiment with MDCXR4-Net, 4,000 images were analysed, with each category containing 1,000 images. The models achieved accuracies of 99.09%, 97.74%, and 90.37% for their respective benchmarks.

Mahamud et al. [8] introduced a framework for explainable transfer learning, employing an improved DenseNet-201 to classify COVID-19, pneumonia, and tuberculosis using chest X-ray images. Data preprocessing was done with a denoising autoencoder, CLAHE, gamma correction, data augmentation, and XAI techniques such as SHAP, LIME, Grad-CAM, and Grad-CAM++. The model performance reached 99.20% of accuracy and 99% of both precision and recall.

Charan et al. [9] explored transfer learning for multi-class lung disease prediction using fused chest X-ray datasets which includes tuberculosis, COVID-19, and pneumonia images. The study related MobileNetV2, VGG16, InceptionNet, ResNet50, and EfficientNet and incorporated Local Binary Pattern (LBP)-based textural features. ResNet50 achieved 97.1% accuracy which is higher than the remaining models.

A two-stage hybrid CNN–ViT framework [10] was proposed for chest disease classification using X-ray images. The study considered Normal, Tuberculosis, Viral Pneumonia, and Bacterial Pneumonia classes and compared several pretrained architectures. For multi-class classification, ViT-B/16 achieved 97.25% accuracy, 97.10% precision, 97.66% recall, and 97.38% F1-score. ResNet-50 achieved 93.65% accuracy and DenseNet-121 achieved 90.43%. The ensemble model achieved 96.18% accuracy, 96.80% precision, 96.00% recall, and 96.40% F1-score. These results highlight the exceptional performance of ViT-B/16 in multi-class lung disease classification.

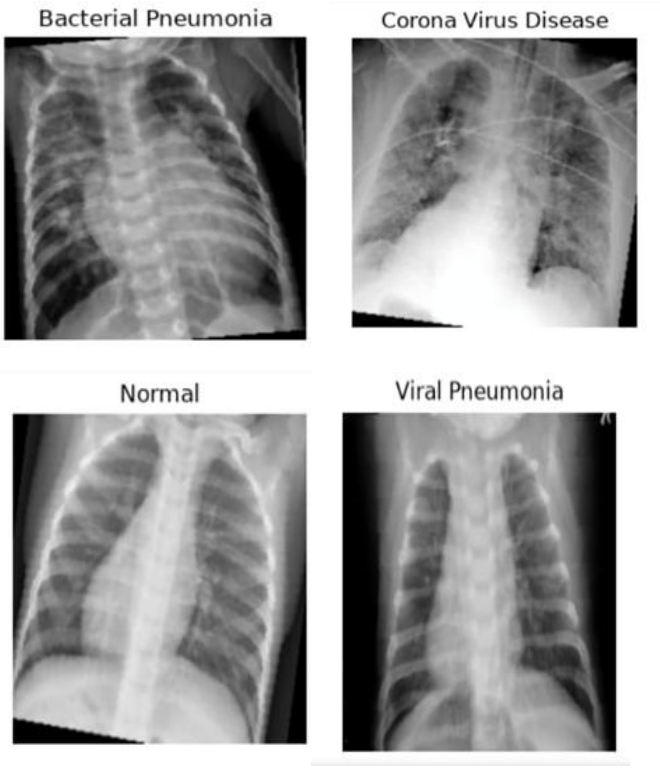
In 2025, Fu et al. [11] developed LungMaxViT, an explainable hybrid transformer model that integrates a CNN block along with squeeze-and-excitation mechanisms and a multi-axis transformer. This model was evaluated using the COVID-QU-Ex and ChestX-ray14 datasets. For the COVID-19 dataset, LungMaxViT achieved an accuracy of 96.8%, an AUC of 98.3%, and an F1-score of 96.7%. On the ChestX-ray14 dataset, it recorded an AUC of 93.2% and an F1-score of 70.7%. The authors also compared it with ResNet50, MobileNetV2, ViT, and MaxViT, showing the benefits of integrating CNN-based local features with transformer-based global representations.

From the observation, transfer learning and vision transformed occupied predominant place in the lung disease classification. Even there are several studies, there is a gap of research to propose a model with vision transformer to classify the most common diseases as covid-19, viral pneumonia, bacterial pneumonia, tuberculosis and Normal using both X-ray and CT-scan images.

III. Dataset and Methodology

A. *Dataset*

In this work, a publicly available Lungs Disease Dataset from Kaggle is taken. The dataset has 10095 Chest X-Rays images of Viral Pneumonia, Bacterial Pneumonia, Covid-19, Tuberculosis and normal with the count of 2008, 2009, 2031, 2034 and 2013 respectively. The Fig.1 represents the X-ray images of various lung diseases.



ISSN: 3049-3129 (Online)

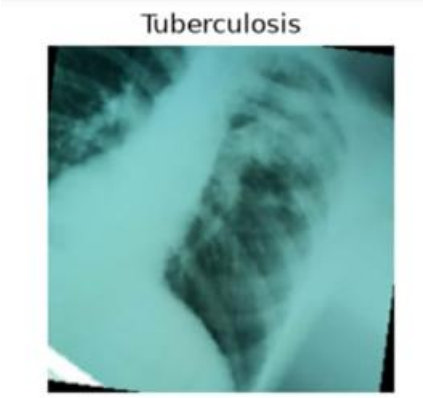


Fig 1. X-Ray images of various lung disease

B. Data Preprocessing

All input images are resized to pixels, while random horizontal flipping, rotation, and resized cropping are applied to the training images to improve model generalization. ImageNet-based normalization is then performed to match the requirements of the pretrained models. The dataset is divided into training and validation sets using an 80:20 ratio.

C. Resnet50

ResNet50 architecture is pre-trained CNN model which is trained with ImageNet. In this study, it is employed as a transfer learning model for lung disease classification. The network is initialized with the default pre-trained weights and the original fully connected classification layer of ResNet50 is replaced with a new linear layer whose output dimension corresponds to the number of target lung disease classes. Through this modification, the network is adapted with the learned features to the specific requirements of the proposed lung disease classification task.

The proposed ResNet50 model is used as a comparative deep learning model for lung disease classification. ResNet50 model is trained using Cross-Entropy Loss and AdamW optimizer. It is fully connected layer is replaced with a new linear layer respected to four diseases.

D. DenseNet121

DenseNet121 is employed as a transfer learning model for lung disease classification. The model is initialized using its default pre-trained weights learned from ImageNet dataset. The classification layer of DenseNet121 is replaced with a new fully connected linear layer, with the output dimension corresponding to the number of target lung disease classes.

DenseNet121 model consists of densely connected layers where the output of each layer enabling feature reuse and improving flow of information. The model learns features from Chest-X Ray images and performs multi class classification.

E. Vision Transformer Base/16 (ViT-B/16)

ViT-B/16 architecture is implemented for lung disease classification using a transfer learning approach. The model is initialized with the default pre-trained weights to utilize visual feature representations learned from ImageNet dataset. The final classification head of the pre-trained ViT-B/16 model is replaced with a new fully connected linear layer with the output dimension of the number of target lung disease classes.

IV. Result and Discussion

All the proposed models are trained and evaluated on a GPU-enabled system using CUDA acceleration. Specifically, the experiments are accomplished on an NVIDIA H200 NVL graphics processing unit (GPU) with approximately 139.8 GB of dedicated memory.

To ensure a fair comparison, all three models are trained using the same experimental settings, including cross-entropy loss, AdamW optimization with a learning rate of 0.0001 and a ReduceLROnPlateau learning-rate scheduler for 20 epochs.

The Fig 8 represents the validation accuracy of each model in every epoch. The result illustrates that all the three models show an overall improvement during the training. Among the three models, DenseNet121 achieves high validation accuracy.

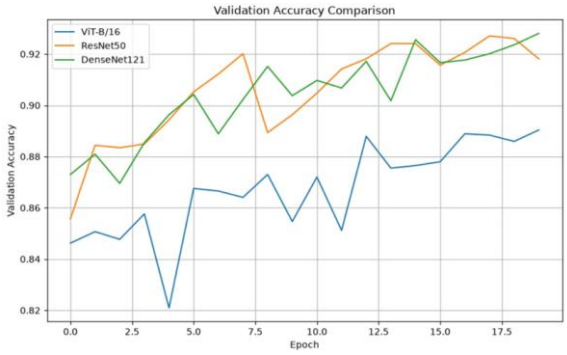


Fig 2. Validation accuracy of different deep learning models.

All the models are tested with test images with the count of 2025. The Fig 3, Fig.5 and Fig.7 represents the confusion matrices of all the models to shows the performance of models in classification of each class. The diagonal elements of confusion matrix show the count of correctly classified images, whereas the off-diagonal elements indicate misclassified samples.

Fig 4, Fig 6 and Fig 8 present the classification report of the ResNet50, DenseNet121 and VIT-B/16 models for the five-class lung disease classification task. The report summarizes the class-wise performance using precision, recall, F1-score, and support for each disease class.

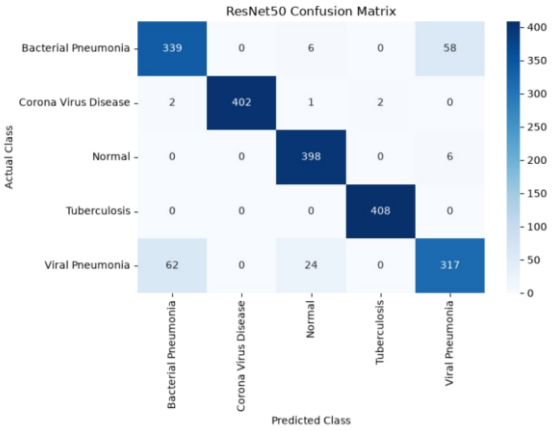


Fig 3. Confusion matrix of ResNet50

```
============================================================
ResNet50
============================================================
Accuracy : 0.9205
Precision: 0.9196
Recall   : 0.9205
F1 Score : 0.9197

Classification Report:

                      precision    recall  f1-score   support

 Bacterial Pneumonia     0.8412    0.8412    0.8412       403
Corona Virus Disease     1.0000    0.9877    0.9938       407
              Normal     0.9277    0.9851    0.9556       404
          Tuberculosis   0.9951    1.0000    0.9976       408
      Viral Pneumonia    0.8320    0.7866    0.8087       403

            accuracy                         0.9205      2025
           macro avg     0.9192    0.9201    0.9194      2025
        weighted avg     0.9196    0.9205    0.9197      2025
```

Fig 4. Classification performance of the ResNet50 model on the test dataset.

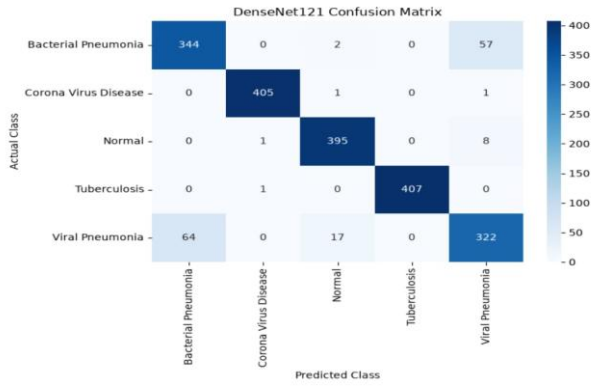


Fig 5. Confusion matrix of DenseNet121

```
============================================================
DenseNet121
============================================================
Accuracy : 0.9249
Precision: 0.9243
Recall   : 0.9249
F1 Score : 0.9245

Classification Report:

                      precision    recall  f1-score   support

 Bacterial Pneumonia     0.8431    0.8536    0.8483       403
Corona Virus Disease     0.9951    0.9951    0.9951       407
              Normal     0.9518    0.9777    0.9646       404
          Tuberculosis   1.0000    0.9975    0.9988       408
      Viral Pneumonia    0.8299    0.7990    0.8142       403

            accuracy                         0.9249      2025
           macro avg     0.9240    0.9246    0.9242      2025
        weighted avg     0.9243    0.9249    0.9245      2025
```

Fig 6. Classification performance of the DenseNet121 model on the test dataset.

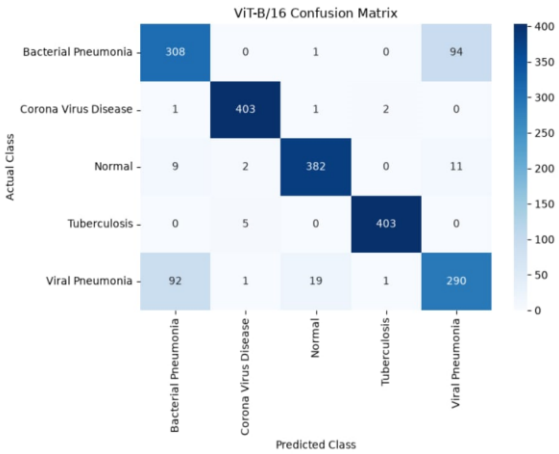


Fig 7. Confusion matrix of VIT-B/16

```
============================================================
ViT-B/16
============================================================
Accuracy : 0.8820
Precision: 0.8818
Recall   : 0.8820
F1 Score : 0.8819

Classification Report:

                      precision    recall  f1-score   support

 Bacterial Pneumonia     0.7512    0.7643    0.7577       403
Corona Virus Disease     0.9805    0.9902    0.9853       407
              Normal     0.9479    0.9455    0.9467       404
          Tuberculosis   0.9926    0.9877    0.9902       408
      Viral Pneumonia    0.7342    0.7196    0.7268       403

            accuracy                         0.8820      2025
           macro avg     0.8813    0.8815    0.8813      2025
        weighted avg     0.8818    0.8820    0.8819      2025
```

Fig 8. Classification performance of the VIT-B/16 model on the test dataset.

Every model is compared in terms of different metrices such as Accuracy, Precision, Recall and F1-Score. The Fig 9. Shows the performance of the three models in different performance metric.

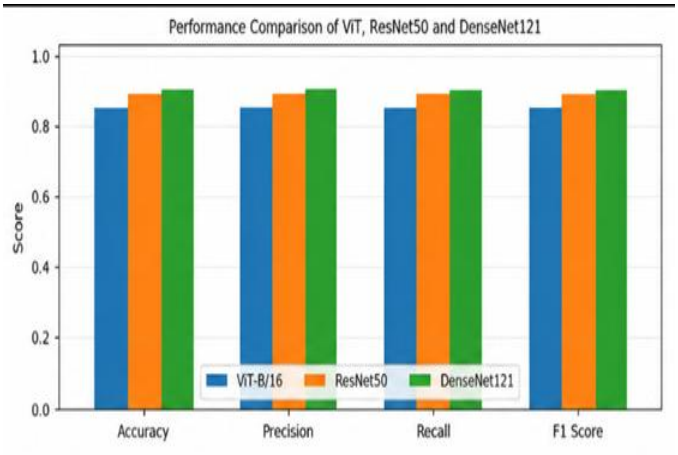


Fig 9. Performance of ViT-B/16, ResNet50, and DenseNet121 based on accuracy, precision, recall, and F1-score.

V. CONCLUSION

This study explores the application of pretrained deep learning architectures for multi-class lung disease classification including CNN architecture and Vision Transformer. Three architectures, namely ViT-B/16, ResNet50, and DenseNet121, are comparatively evaluated using accuracy, precision, recall, and F1-score. From the experimental results among the three models, we observe that DenseNet121 provides best overall classification performance with accuracy of 92.49% and F1-score of 92.45%. ResNet50 also achieves competitive performance as accuracy of 92.05% and F1-score of 91.97%, while ViT-B/16 achieves accuracy of 88.18%. A detailed analysis of ResNet50 revealed excellent performance in detecting tuberculosis, with a recall of 100% and an F1-score of 99.76%, though it was less effective for viral pneumonia.

The performance of models is evaluated with multiple metrics rather than relying solely on accuracy to find the appropriate pretrained model for further development of an efficient and reliable computer-aided lung disease classification system. In the future, the proposed framework may be extended federated learning framework to reduce the sharing of sensitive medical images.

REFERENCES

[1] A. Vardhan, A. Makhnevich, P. Omprakash, D. Hirschorn, M. Barish, S. L. Cohen, *et al.*, "A radiographic, deep transfer learning framework, adapted to estimate lung opacities from chest X-rays," *Bioelectronic Medicine*, vol. 9, Art. no. 1, 2023, doi: 10.1186/s42234-022-00103-0.

[2] F. J. M. Shamrat, S. Azam, A. Karim, K. Ahmed, F. M. Bui, and F. De Boer, "High-precision multiclass classification of lung disease through customized MobileNetV2 from chest X-ray images," *Computers in Biology and Medicine*, vol. 155, Art. no. 106646, 2023, doi: 10.1016/j.compbiomed.2023.106646.

[3] M. H. Al-Sheikh, O. Al Dandan, A. S. Al-Shamayleh, H. A. Jalab, and R. W. Ibrahim, "Multi-class deep learning architecture for classifying lung diseases from chest X-Ray and CT images," *Scientific Reports*, vol. 13, Art. no. 19373, 2023, doi: 10.1038/s41598-023-46147-3.

[4] S. Singh, M. Kumar, A. Kumar, B. K. Verma, K. Abhishek, and S. Selvarajan, "Efficient pneumonia detection using Vision Transformers on chest X-rays," *Scientific Reports*, vol. 14, Art. no. 2487, 2024, doi: 10.1038/s41598-024-52703-2.

[5] J. Ko *et al.*, "Optimization of vision transformer-based detection of lung diseases from chest X-ray images," *BMC Medical Informatics and Decision Making*, 2024, doi: 10.1186/s12911-024-02591-3.

[6] X. Li *et al.*, "Automated thorax disease diagnosis using multi-branch residual attention network," *Scientific Reports*, vol. 14, Art. no. 11865, 2024, doi: 10.1038/s41598-024-62813-6.

[7] A. N. Patel, R. Murugan, G. Srivastava, P. K. R. Maddikunta, G. Yenduri, T. R. Gadekallu, and R. Chengoden, "An explainable transfer learning framework for multi-classification of lung diseases in chest X-rays," *Alexandria Engineering Journal*, vol. 98, pp. 328–343, 2024, doi: 10.1016/j.aej.2024.04.072.

[8] U. Chutia, A. S. Tewari, J. P. Singh, and V. K. Raj, "Classification of lung diseases using an attention-based modified DenseNet model," *Journal of Imaging Informatics in Medicine*, vol. 37, no. 4, pp. 1625–1641, 2024, doi: 10.1007/s10278-024-01005-0.

[9] A. Hage Chehade, N. Abdallah, J.-M. Marion, M. Hatt, M. Oueidat, and P. Chauvet, "Advancing chest X-ray diagnostics: A novel CycleGAN-based preprocessing approach for enhanced lung disease classification in ChestX-Ray14," *Computer Methods and Programs in Biomedicine*, vol. 259, Art. no. 108518, 2025, doi: 10.1016/j.cmpb.2024.108518.

[10] A. Hage Chehade, N. Abdallah, J.-M. Marion, M. Oueidat, and P. Chauvet, "Evaluating the impact of view position in X-ray imaging for the classification of lung diseases," *Physical and Engineering Sciences in Medicine*, vol. 48, no. 3, pp. 1225–1236, 2025, doi: 10.1007/s13246-025-01579-1.

[11] V. Anand, M. Shuaib, I. Khan, M. Ullah, and S. Alam, "Secure pulmonary diagnosis using transformer-based approach to X-ray classification with KL divergence optimization," *Frontiers in Medicine*, vol. 12, Art. no. 1716066, 2025, doi: 10.3389/fmed.2025.1716066.