

Multiclass Classification of Chest X-rays based Pulmonary Disorder Using a Specialized VGG-19 Deep Neural Network

Vazralu M.¹, Madijagan M.²

School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

E-mail: ¹mvazralu1206@gmail.com, ²madijagan.m@vit.ac.in

Abstract

Respiratory infections such as COVID-19, tuberculosis (TB) and pneumonia, remain important global health challenges, often requiring rapid and accurate diagnosis to prevent complications. Due to the visual similarities in chest X-ray (CXR) images, distinguishing between these diseases can be complex. In this study, we proposed, a deep learning (DL)-based model utilizing a customized VGG-19 architecture for multiclass classification of lung diseases, including COVID-19, pneumonia, TB, and healthy cases. A total of 5,928 CXR images were collected from open-access platforms, comprising COVID-19, pneumonia, TB, and normal cases. The dataset was pre-processed using bilateral filtering for noise suppression and Multiscale Retinex for image enhancement. Furthermore, data augmentation and image resizing were also applied to increase robustness. When compared with the state-of-the-art techniques, the proposed method achieved a classification accuracy of 98.48% in identifying various lung disorders, with precision at 97% and an F1-score of 96%, indicating that it is an appropriate technique for computerized lung disease diagnosis in clinical environments.

Keywords: Chest X-ray, Lung Disease Detection, VGG-19, COVID-19, Pneumonia, Tuberculosis, Transfer Learning.

1. Introduction

Respiratory diseases such as COVID-19, pneumonia and TB are still major causes of death worldwide. Serious medical problems can occur if these illnesses are not diagnosed early especially in developing countries with inadequate access to health care. Chest X-rays are one of the most common methods to provide a simple way to find many kinds of lung disorders and they are also less expensive. Because of their easy availability, they have become a key part of traditional diagnostics, in situations where there are limited resources.

Even for an expert radiologist, interpreting CXR images is a challenging task because different diseases exhibit various overlapping patterns. For example, the small visual similarities that may be present between COVID_19 and the early stages of pneumonia can make it problematic to visualize manually and prone to errors. Physicians need computer-aided diagnostic (CAD) systems that can help in automatically analyzing CXR images and accurately diagnosing lung conditions. [1,2]

Convolutional Neural Networks (CNNs) are particularly compatible with medical image examination due to their capability to automatically learn complex spatial hierarchies

from raw pixel information. They are good at detecting fine textures and structural changes like opacities, lesions, and consolidations that might be difficult to detect through manual observation. They can efficiently manage noise, varying illumination conditions and large datasets which makes them useful for diagnostic tasks involving chest radiography. Recent studies have highlighted the strengths of using deep learning (DL) models for the automatic classification of chest related diseases from CXR images. [3-5]

The proposed method introduced a customized DL model to classify CXR images into four categories namely COVID-19, TB, pneumonia and healthy. Unlike conventional methods the proposed method enhances classification by combining the pertained Visual Geometry Group-19(VGG-19) backbone with a Feedforwarded Neural Network.

The key contribution of this approach is to implement a two-step image pre-processing. In step one we used bilateral filtering to suppress noise while preserving structural boundaries. In step two, we used Multi-Scale Retinex (MSR) to improve the visibility of low-contrast regions in radiographs. Additionally, Real-Time data augmentation was applied during training to increase input variability and enhance the model's generalization capability. Evaluated on a diverse dataset of 5,928 CXR images, the proposed method achieved consistently high classification performance, supporting its potential as a decision-support tool in medical screening applications. The main aim of this study is to build architecture and evaluate a DL-based multiclass classification system capable of accurately identifying COVID-19, pneumonia, TB, and healthy cases from CXR images, using a robust and optimized CNN pipeline.

The remainder of this paper is organized as follows: Section 2 presents the related literature and background on DL-based approaches for lung disease classification. Section 3 describes the dataset, preprocessing techniques, and the architecture of the proposed model. Section 4 details the experimental setup, outcomes, and performance analysis. Finally, Section 5 concludes the study with key decisions and future research guidelines.

2. Related Work

DL methods, particularly CNNs, have become prevalent in CXR analysis for lung disease recognition. Researchers have explored a variety of architectures—from fine-tuned VGG models to hybrid frameworks incorporating attention modules and Vision Transformers (ViT) to detect pneumonia, COVID-19, TB, and additional lung illnesses. A detailed comparison of existing DL approaches for lung disease classification is provided in Table 1, including author contributions, model architectures, performance metrics, and constraints.

Table 1. Summary of Related Work on DL-based Lung Disease Classification Using Chest X-rays

Author(s)	Method	Performance	Limitation
Rajpurkar et al.[6]	CheXNet (121-layer DenseNet on ChestX-Ray14)	Radiologist-level Pneumonia detection (~F1-score superiority) (MedRxiv, arXiv)	Focused on Pneumonia only

Rahman et al.[7]	Deep CNNs + segmentation for TB detection	~99% (TB vs normal)	Single-disease classification, no multiclass focus (arXiv)
El Asnaoui et al.[8]	ResNet50, MobileNetV2, Inception-ResNetV2, VGG19	>96% (ResNet50)	Limited to binary classification tasks
Narin et al.[9]	Fine-tuned ResNet50	Up to 99.7%	Small, imbalanced datasets; only binary classification
Khan et al.[10]	Xception-based CNN for 4-class CXR classification	89.60%	Public datasets only; moderate class separation
Bharati et al.[11]	Hybrid CNN (VGG + Spatial Transformer)	73%	Trained on NIH dataset; no COVID-19 cases included
Gao et al.[12]	Vision Transformer on CT scans	F1-score: 0.76	Not based on CXR; limited modality generalization
Krishnan & Krishnan[13]	Fine-tuned Vision Transformer on CXR	97.61%	Binary only; small test set
Qi et al.[14]	ViT with cross-attention using enhanced/original CXR	96.21% (10% labeled data)	High training complexity; ViT requires more compute
Hadhoud et al[15]	Hybrid ResNet-50 + ViT-b16 on CXR for TB & Pneumonia	Qualitative improvement	Accuracy not explicitly reported
Yulvina et al[16]	CNN + ViT on 133-image TB dataset	91.10%	Trained on small (133-image) dataset
Slika et al[17]	Vision Transformer regression on CXR	Pearson 0.923; MAE ~0.512	Regression only; not classification
Kumar et al.[18]	VGG19 + PCA + ICEA	98.88%	Complex architecture; lacks external validation
Kabir et al.[19]	Knowledge distillation for CXR	~97% accuracy, ~94% precision	Training complexity; limited external validation

Ou et al.[20]	U-Net variants and segmentation for TB lesions	Mean Precision ~0.88, Recall ~0.75, F1 ~0.81, MIoU ~0.70, Accuracy ≈1.0	Focused on tuberculosis only; no multiclass classification
Sharma et al.[21]	Deep CNN segmentation & classification for TB	High accuracy reported	Lacks multiclass evaluation and comparisons
Oltu et al.[22]	Compact Convolutional Transformer (CCTCOVID)	99.20%	Relies on well-labeled datasets; high compute requirements
Kishore et al.[23]	SWAV self-supervised + KD on ResNet student	Improved accuracy, reduced GFLOPs	Evaluated largely on binary datasets

A review of the prior studies shows the effectiveness of DL models in identifying lung diseases using X-rays. However, many of these approaches focus on binary classification, insufficient preprocessing and a lack of diverse datasets. To overcome the above limitations, we proposed a custom-made VGG-19-based model designed to classify multiple lung diseases. The proposed technique uses advanced methods in preprocessing images, reducing noise and transfer learning. The technique aims to deliver improved accuracy and useful robustness in real-time diagnostic situations.

3. Methodology

5,928 CXR images are collected from the publicly accessible Kaggle dataset for our suggested method [24]. The dataset is divided into four categories: normal cases (1,802), COVID-19 (1,626), pneumonia (1,800), and TB (700). Since the posterior-anterior (PA) view offers a more accurate anatomic representation and medical clarity than the anterior-posterior (AP) view, it is the suggested view in clinical radiology. This consistency supports reliable automatic classification and aids in suitable understanding. The dataset was divided as 70% for training and 30% for testing, for the proposed model's development and validation. Table 2 provides an overview of the image distribution for each class and subset.

Table 2. Class-wise Distribution of CXR Images for Training and Testing in the Proposed Framework

Class	Total Samples	Training Samples (70%)	Testing Samples (30%)
COVID-19	1,626	1,138	488
Pneumonia	1,800	1,260	540
Tuberculosis	700	490	210
Normal	1,802	1,261	541
Total	5,928	4,149	1,779

3.1 Pre-processing Techniques

All the medical images must be preprocessed before being applied to the DL applications. Several preprocessing techniques have been used to improve the images in the suggested model so that they are readable and appropriate for feature extraction and classification.

3.1.1 Image Resizing

Every image of the chest was set to a consistent resolution of 224×224 pixels in order to comply with the input requirements of the projected architecture. This uniform input size ensures consistent processing of all samples while maintaining the critical anatomical components required for accurate feature extraction and classification.

3.1.2 Bilateral Filtering for Noise Reduction

Bilateral filtering was used to minimize unwanted noise while maintaining key information. Bilateral filtering reduces noise without misrepresenting important anatomical boundaries as it takes into consideration both spatial proximity and pixel intensity, in contrast with standard filters that may smooth both noise and significant edges. This is particularly beneficial for preserving characteristics like penetrations and nodules in the lung. Figure 1 displays the bilateral filtering's input and output.

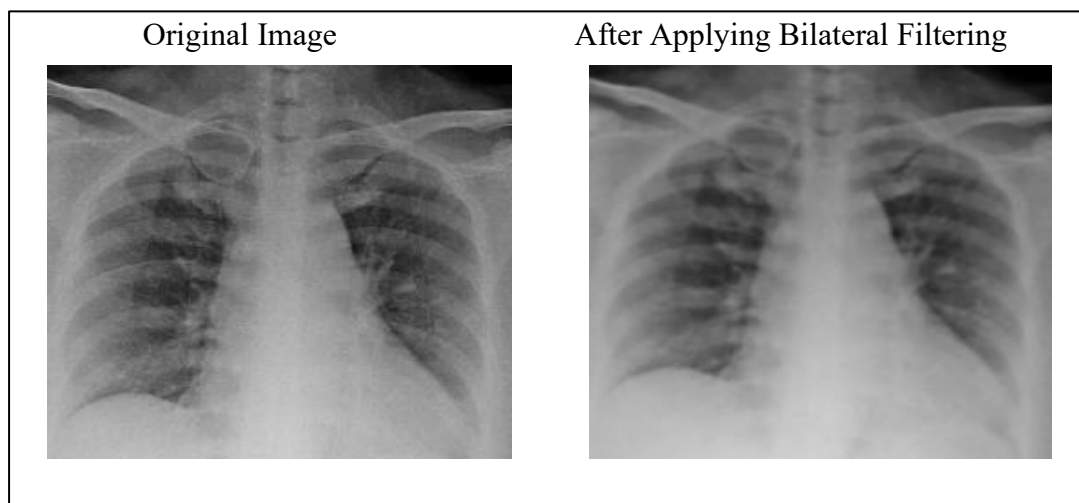


Figure 1. The Result of Applying the Bilateral Filter

3.1.3 Contrast Enhancement Using Multi-Scale Retinex (MSR)

Multi-Scale Retinex (MSR) was used to make subtle features in the X-ray images clearer. This technique, which is based on the Retinex theory regarding human vision, integrates many dynamic range compression scales to enhance image contrast. [25] It aids in improving local contrast, especially in areas where anomalies linked to disease may be harder to identify. Figure 2 below depicts input and output of the Multi-Scale Retinex enhancement.

3.1.4 Image Normalization

Every pixel value in the image was changed to fall between 0 and 1. In addition to ensuring uniformity across all input data, this normalization stage helps the model train faster and more consistently.

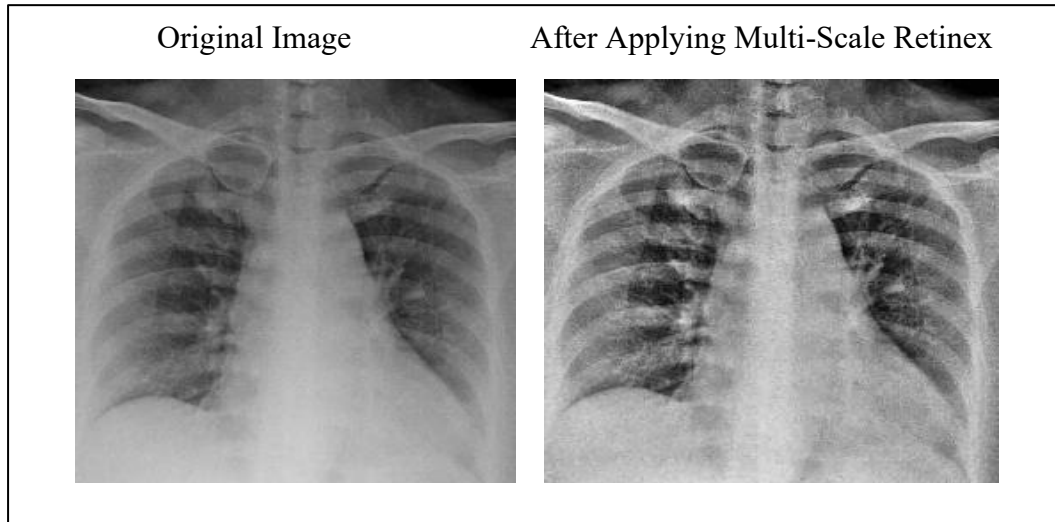


Figure 2. The Result of Applying the MSR Enhancement

3.1.5 Data Augmentation

Real-time data augmentation was employed during training to enhance model generalization and reduce overfitting. Various changes were applied to the input images, such as horizontal flipping, random rotations, zooming, and shearing. By representing inconsistency in real clinical imaging, these changes help the model learn more robust features.

3.2 Proposed Method

Optimized for multiclass classification of lung diseases using CXR images, the proposed model is a modified version of the popular VGG-19 architecture. COVID-19, pneumonia, tuberculosis, and normal cases are the four categories it is intended to identify. The model incorporates modifications to better fit medical imaging applications while utilizing the strengths of deep CNNs.

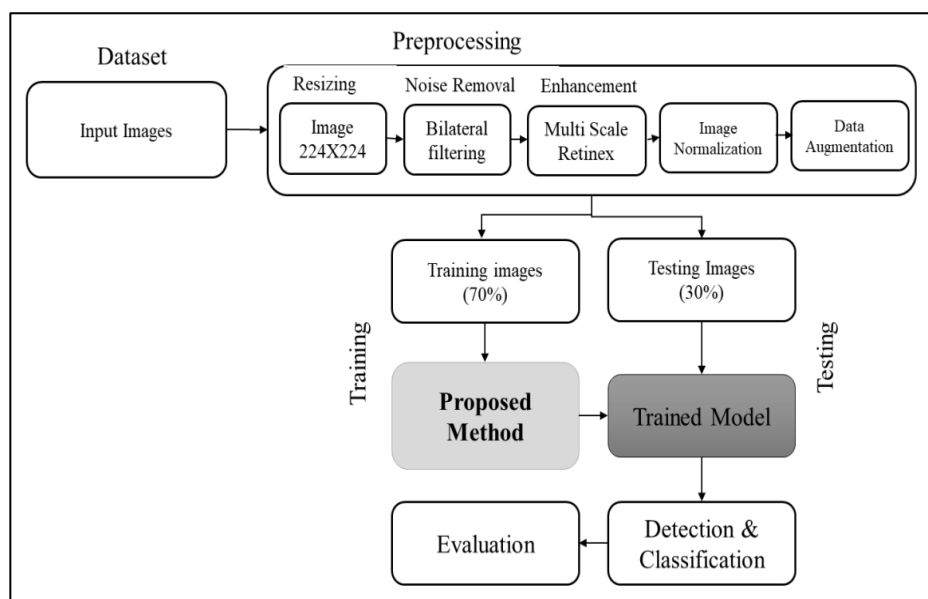
3.2.1 Architectural Overview

The pre-trained VGG-19 model, which was initially trained on the ImageNet dataset, is used as the convolutional basis in the proposed method. This enables the network to make use of rich hierarchical feature representations that have been acquired from a substantial image corpus. A customized classifier head was used in place of the original fully connected layers to adapt the model for the lung disease classification task.

Table 3. Layer-wise Architecture of the Proposed Model

Layer Type	Details
Input Layer	$224 \times 224 \times 3$ RGB image
Convolutional Blocks	16 convolutional layers from pre-trained VGG-19 (frozen during training)
Activation Function	ReLU after each convolutional layer
Pooling Layers	MaxPooling2D after each block to reduce spatial dimensions
Flatten Layer	Converts final feature maps into a 1D vector
Dense Layer 1	512 neurons with ReLU activation
Dropout Layer 1	Dropout rate = 0.5 to prevent overfitting
Dense Layer 2	256 neurons with ReLU activation
Dropout Layer 2	Dropout rate = 0.5
Output Layer	Dense layer with 4 neurons + Softmax activation (for 4-class classification)

The layer wise architecture presented in Table 3 combines the proven feature extraction capabilities of the VGG-19 convolutional backbone with a lightweight, customized classification head. The decision to retain the pretrained VGG-19 convolutional layers allows the model to leverage generalized features learned from the ImageNet dataset, which are transferable to medical imaging. By freezing these layers, training time is reduced and the risk of overfitting is minimized, especially given the moderate size of the dataset. The added fully connected layers and dropout units provide task-specific learning and enhance generalization, while the final SoftMax layer enables precise multiclass prediction across the four lung disease categories. Figure 3 presents the overall plan of the proposed classification framework, while Figure 4 presents the detailed architectural structure of the planned model.

**Figure 3.** Block Diagram of the Proposed Approach

3.2.2 Training Configuration

The proposed model was developed using a supervised learning approach aimed at improving performance across multiple disease categories. TensorFlow and Keras were used for training in a Python environment with GPU support for faster model convergence and less calculation time. An NVIDIA GeForce RTX 3060 GPU with 12 GB of VRAM and an Intel Core i7 processor with 32 GB of RAM were used to train the model. The average inference time for CXR images is approximately 46 milliseconds, which shows that the proposed model is suitable for real-time applications.

The convolutional base remained with fixed weights to utilize prior knowledge and improve generalization, while only the output and fully connected layers of VGG-19 were trained on the given dataset. This proposed method is useful only when a moderately large CXR dataset is available because it removes the need for extensive training from scratch. Table 4 provides a summary of the parameters used for training the model for optimization.

Table 4. Summary of Training Parameters Used for Training the Proposed Model

Parameter	Value	Description
Loss Function	Categorical Cross-Entropy	Suitable for multiclass classification with mutually exclusive classes
Optimizer	Adam	Adaptive optimizer combining momentum and learning rate adjustment
Learning Rate	0.0001	Chosen for stable and gradual convergence
Batch Size	32	Number of samples processed per training step
Epochs	100	Full passes through the training data
Validation Split	0.2 (20%)	Proportion of training data used for validation

4. Experimental Results and Performance Evaluation

The results of the proposed model on the multiclass CXR dataset are shown in this section. On a 70:30 train-test split, the model was trained using transfer learning, with the convolutional base (VGG-19) remaining frozen and the final fully connected layers being fine-tuned. A number of conventional classification metrics were used to assess performance.

4.1 Evaluation Metrics

The following metrics were computed in order to evaluate the model's competency.

Accuracy: Proportion of correctly predicted instances over the total number of predictions.

$$Accuracy = \frac{\text{No. of images correctly classified}}{\text{Total no of images}} \quad (1)$$

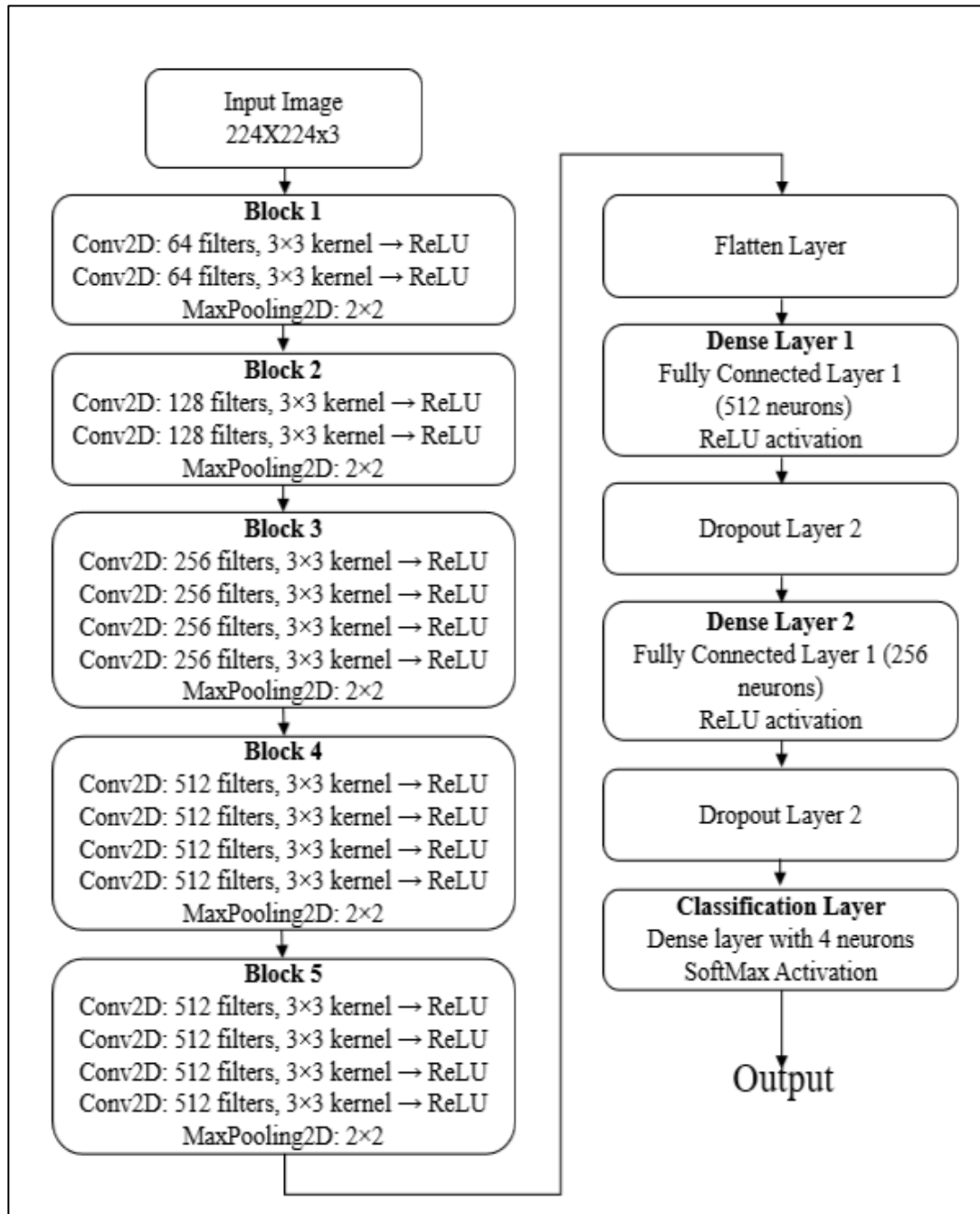


Figure 4. Shows the Architecture of the Proposed Method

Precision: The ratio of the total number of true positives to the total number of false positives.

$$\text{Precision} = \frac{\text{sum of all true positives (Tp)}}{\text{Sum of all True Positives(TP)+All False Positives (FP)}} \quad (2)$$

Recall (Sensitivity): The ratio of the total number of true positives to the total number of false negatives and true positives.

$$\text{Recall} = \frac{\text{sum of all true positives (Tp)}}{\text{sum of all true positives(Tp)+ All False Negatives (FN)}} \quad (3)$$

F1-Score: Harmonic mean of recall and precision, showing equilibrium between the two.

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4)$$

4.2 Results on Test Set

All four of the test set's lung disease categories showed consistently good classification performance from the suggested model. The model's strong generalization ability is demonstrated by the individual class accuracies, which varied from 98.20% for TB to 98.88% for normal cases (Table 5). The model's ability to differentiate between these clinically similar conditions was demonstrated by the 98.49% accuracy rate for both the COVID-19 and pneumonia classes, as well as its high precision (0.97) and recall (0.95). Although the accuracy was a little lower, the TB class showed reliable disease-specific feature detection with high recall (0.90) and precision (0.94). The normal class had the highest accuracy, with precision and recalls of 0.98, indicating the model's reliability in detecting healthy cases. Each class's F1-score varied from 0.92 to 0.98, indicating balanced performance free from class bias.

Based on the high recall, precision, and F1-scores observed in all classes, the proposed model seems to be a good fit for classifying lung diseases using CXR images. The consistent performance demonstrates the effectiveness of transfer learning, architectural design, and preprocessing methods.

Table 5. Class-wise Performance Metrics of Proposed Model on Test Dataset

Class	TP	TN	FP	FN	Accuracy (%)	Precision	Recall	F1-Score
COVID-19	470	1281	10	18	98.43	0.98	0.96	0.97
Pneumonia	520	1231	8	20	98.43	0.98	0.96	0.97
Tuberculosis	190	1557	12	20	98.2	0.94	0.9	0.92
Normal	530	1229	9	11	98.88	0.98	0.98	0.98
Average					98.49	0.97	0.95	0.96

The model had some misclassifications even with its excellent performance. The false negative findings (18 and 20, respectively) in the COVID-19 and pneumonia classes indicate that it is hard to distinguish overlapping features, especially in the initial phase. TB also had 20 false negatives and 12 false positives, which are most likely due to its visual similarity with pneumonia. Since there was minimal misclassification in the regular class, the model's performance in discerning healthy cases correctly was better validated. Such errors indicate that separability among visually dissimilar conditions might be improved in later iterations, even though overall high performance is achieved.

4.2.1 Accuracy

The new model performs better than current DL methods for lung disease classification. Table 6 illustrates how although models such as CoroNet and VDSNet obtained 89.60% and 73.00% accuracy, respectively, the Multi-Feature Vision Transformer (ViT) boosted accuracy to 96.21%. Conversely, the proposed method's remarkable accuracy of 98.49% was achieved

due to the effective utilization of transfer learning, a pre-trained VGG-19 backbone, and an optimized preprocessing pipeline. These optimizations led to improved feature extraction and generalization for all four classes. The graphical accuracy comparison of the proposed method for multiclass classification is depicted in Figure 5.

Table 6. Accuracy Comparison of Proposed Method with Existing DL Models

Model	Accuracy (%)
CoroNet [10]	89.60%
VDSNet (VGG + STN)[11]	73.00%
Multi-Feature ViT Model [14]	96.21%
Proposed Method	98.49%

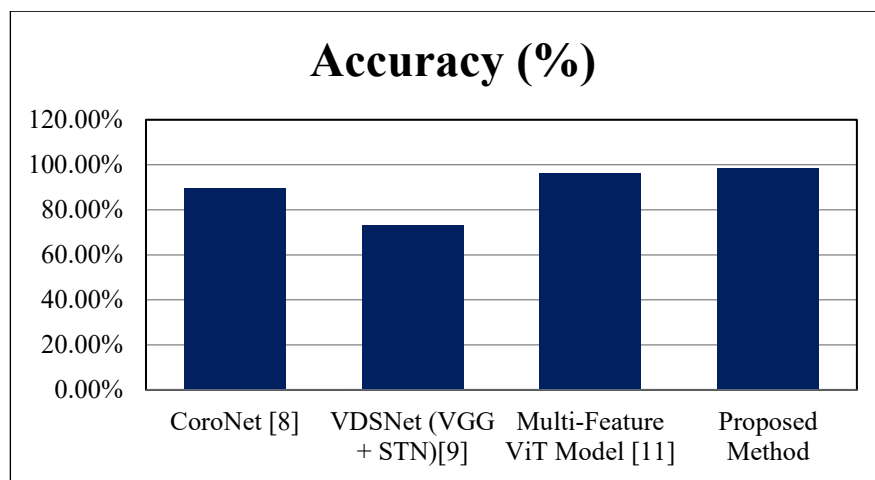


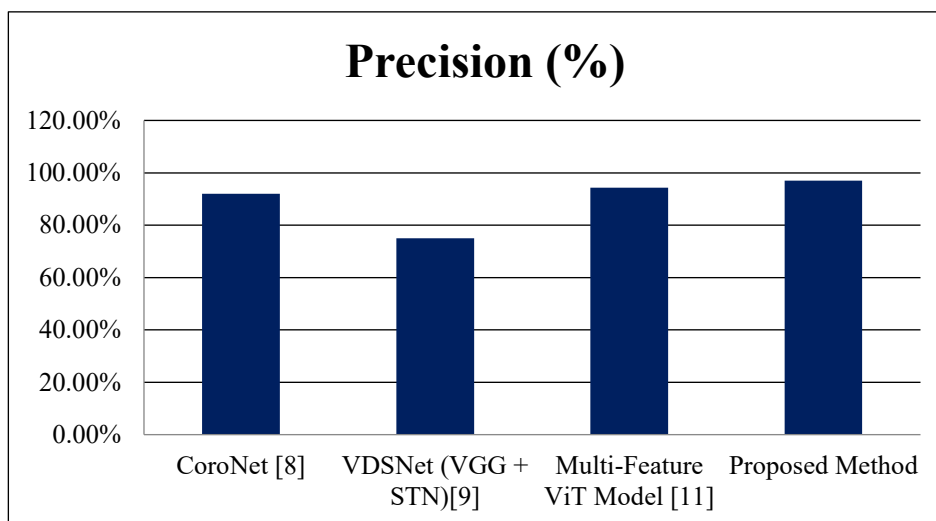
Figure 5. Accuracy Comparative Performance Metrics of Proposed Method for Multiclass Classification

4.2.2 Precision

Comparison values of precision for various DL models have been presented in Table 7. The proposed model achieved the highest precision of 97.00%, demonstrating its ability to accurately classify disease cases and minimize false positives. While CoroNet achieved 92.00% accuracy and the Multi-Feature ViT model lagged behind at 94.30%, both were practical but less accurate than the suggested approach. On the contrary, VDSNet, which integrates VGG with Spatial Transformer Networks, had the lowest accuracy at 75.00%, according to the scenario that it did not have high capacity for feature extraction of complex features and did not contain full preprocessing mechanisms. Due to its design decisions, transfer learning using pretrained weights, and improved preprocessing techniques like bilateral filtering and MSR that preserve image details and restore feature extraction, the Proposed Method performs better than the state-of-the-art. Figure 6 illustrates the graphical comparison of the proposed method for Multiclass Classification.

Table 7. Precision Comparison of Proposed Method with State-art Methods

Model	Precision (%)
CoroNet [10]	92.00%
VDSNet (VGG + STN)[11]	75.00%
Multi-Feature ViT Model [14]	94.30%
Proposed Method	97.00%

**Figure 6.** Precision Comparative Performance Metrics of Proposed Method for Multiclass Classification

5. Conclusion

In this paper, we have suggested a personalized DL model in order to categorize CXR images into four categories. The framework is built on the VGG-19 network and enhanced by incorporating a feedforwarded neural network personalized for the classification task. To restore the important features of CXR images a two-step preprocessing methodology is implemented. To advance the generalization ability, data augmentation and transfer learning were employed. The evaluation showed that the recommended method achieved a 98.49% classification accuracy, a 96% F1-score across all categories, and a good precision score of 97%. The proposed model performed well while comparing with other models such as VDSNet, CoroNet, and Multi-feature ViT particularly in terms of removing false positives and achieving accurate results. These results support its real-time application for medical screening. Nevertheless, the model's ability to generalize across a wide range of real-world clinical settings may be limited if only publicly available datasets are used. Additionally, other imaging methods like CT scans that could further the investigation are not included in the current framework; only chest X-rays are. This framework may be expanded to include CT scan examinations in the future and combined with artificial intelligence (AI) tools that can interpret the results to help physicians.

References

- [1] Mwendo, Irad, Kinyua Gikunda, and Anthony Maina. "Deep transfer learning for detecting Covid-19, Pneumonia and Tuberculosis using CXR images--A Review." *arXiv preprint arXiv:2303.16754* (2023).
- [2] Yadav, Sapna, Syed Afzal Murtaza Rizvi, and Pankaj Agarwal. "Detection of lung diseases for Pneumonia, tuberculosis, and covid-19 with artificial intelligence tools." *SN Computer Science* 5, no. 3 (2024): 303.
- [3] Ahmed, Mohammed Salih, Atta Rahman, Faris AlGhamdi, Saleh AlDakheel, Hammam Hakami, Ali AlJumah, Zuhair AlIbrahim, Mustafa Youldash, Mohammad Aftab Alam Khan, and Mohammed Imran Basheer Ahmed. "Joint diagnosis of Pneumonia, COVID-19, and tuberculosis from Chest X-Ray images: A deep learning approach." *Diagnostics* 13, no. 15 (2023): 2562.
- [4] Venkataramana, Lokeswari, D. Venkata Vara Prasad, S. Saraswathi, C. M. Mithumary, R. Karthikeyan, and N. Monika. "Classification of COVID-19 from tuberculosis and Pneumonia using deep learning techniques." *Medical & Biological Engineering & Computing* 60, no. 9 (2022): 2681-2691.
- [5] Wang, Wenkong, Weijie Huang, Quanli Lu, Jiyang Chen, Menghua Zhang, Jia Qiao, and Yong Zhang. "Attention mechanism-based deep learning method for hairline fracture detection in hand X-rays." *Neural Computing and Applications* 34, no. 21 (2022): 18773-18785.
- [6] Rajpurkar, Pranav, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding et al. "Chexnet: Radiologist-level Pneumonia detection on Chest X-Rays with deep learning." *arXiv preprint arXiv:1711.05225* (2017).
- [7] Rahman, Tawsifur, Amith Khandakar, Muhammad Abdul Kadir, Khandaker Rejaul Islam, Khandakar F. Islam, Rashid Mazhar, Tahir Hamid et al. "Reliable tuberculosis detection using Chest X-Ray with deep learning, segmentation and visualization." *Ieee Access* 8 (2020): 191586-191601.
- [8] El Asnaoui, Khalid, and Youness Chawki. "Using X-ray images and deep learning for automated detection of coronavirus disease." *Journal of Biomolecular Structure and Dynamics* 39, no. 10 (2021): 3615-3626.
- [9] Narin, Ali, Ceren Kaya, and Ziyne Pamuk. "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks." *Pattern Analysis and Applications* 24, no. 3 (2021): 1207-1220.
- [10] Mandindi, Nkcubeko Umzubongile Siphamandla. "Detecting and Understanding COVID-19 Misclassifications: A Deep Learning and Explainable AI Approach." Master's thesis, University of the Witwatersrand, Johannesburg (South Africa), 2023.
- [11] Bharati, Subrato, Prajoy Podder, and M. Rubaiyat Hossain Mondal. "Hybrid deep learning for detecting lung diseases from X-ray images." *Informatics in Medicine Unlocked* 20 (2020): 100391.

- [12]Zavrak, Sultan. "Enhancing covid-19 diagnosis through vision transformer-based analysis of Chest X-Ray images." arXiv preprint arXiv:2306.06914 (2023).
- [13]Krishnan, Koushik Sivarama, and Karthik Sivarama Krishnan. "Vision transformer based COVID-19 detection using Chest X-Rays." In 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC), IEEE, 2021, 644-648.
- [14]Qi, Xiao, David J. Foran, John L. Noshier, and Ilker Hacıhaliloglu. "Multi-feature vision transformer via self-supervised representation learning for improvement of covid-19 diagnosis." In Workshop on Medical Image Learning with Limited and Noisy Data, pp. 76-85. Cham: Springer Nature Switzerland, 2022.
- [15]Hadhoud, Yousra, Tahar Mekhaznia, Akram Bennour, Mohamed Amroune, Neesrin Ali Kurdi, Abdulaziz Hadi Aborujilah, and Mohammed Al-Sarem. "From binary to multi-class classification: A two-step hybrid cnn-vit model for chest disease classification based on x-ray images." *Diagnostics* 14, no. 23 (2024): 2754.
- [16]Yulvina, Rizka, Stefanus Andika Putra, Mia Rizkinia, Arierta Pujitresnani, Eric Daniel Tenda, Reyhan Eddy Yunus, Dean Handimulya Djumaryo, Prasandhya Astagiri Yusuf, and Vanya Valindria. "Hybrid Vision Transformer and Convolutional Neural Network for Multi-Class and Multi-Label Classification of Tuberculosis Anomalies on Chest X-Ray." *Computers* 13, no. 12 (2024): 343.
- [17]Slika, Bouthaina, Fadi Dornaika, Hamid Merdji, and Karim Hammoudi. "Lung Pneumonia severity scoring in Chest X-Ray images using transformers." *Medical & Biological Engineering & Computing* 62, no. 8 (2024): 2389-2407.
- [18]Kumar, Rahul, Cheng-Tang Pan, Yi-Min Lin, Shiue Yow-Ling, Ting-Sheng Chung, and Uyanahewa Gamage Shashini Janesha. "Enhanced Multi-Model Deep Learning for Rapid and Precise Diagnosis of Pulmonary Diseases Using Chest X-Ray Imaging." *Diagnostics* 15, no. 3 (2025): 248.
- [19]Kabir, Md Mohsin, M. F. Mridha, Ashifur Rahman, Md Abdul Hamid, and Muhammad Mostafa Monowar. "Detection of COVID-19, Pneumonia, and tuberculosis from radiographs using AI-driven knowledge distillation." *Heliyon* 10, no. 5 (2024).
- [20]Ou, Chih-Ying, I-Yen Chen, Hsuan-Ting Chang, Chuan-Yi Wei, Dian-Yu Li, Yen-Kai Chen, and Chuan-Yu Chang. "Deep learning-based classification and semantic segmentation of lung tuberculosis lesions in Chest X-Ray images." *Diagnostics* 14, no. 9 (2024): 952.
- [21]Sharma, Vinayak, Sachin Kumar Gupta, and Kaushal Kumar Shukla. "Deep learning models for tuberculosis detection and infected region visualization in Chest X-Ray images." *Intelligent Medicine* 4, no. 2 (2024): 104-113.
- [22]Oltu, Burcu, Selda Güney, Seniha Esen Yuksel, and Berna Dengiz. "Automated classification of Chest X-Rays: a deep learning approach with attention mechanisms." *BMC Medical Imaging* 25, no. 1 (2025): 71.
- [23]Kishore, Jaydeep, Akshita Jain, K. Krishna Koushika, Pawan Kumar Mishra, Shekhar Karanwal, and Surendra Solanki. "Enhancing medical diagnosis on Chest X-Rays:

knowledge distillation from self-supervised based model to compressed student model." *Discover Computing* 28, no. 1 (2025): 118.

- [24] Chowdhury, Muhammad EH, Tawsifur Rahman, Amith Khandakar, Rashid Mazhar, Muhammad Abdul Kadir, Zaid Bin Mahbub, Khandakar Reajul Islam et al. "Can AI help in screening viral and COVID-19 Pneumonia?." *Ieee Access* 8 (2020): 132665-132676.
- [25] Kumar, Chinthakindi Kiran, Gaurav Sethi, and Kirti Rawal. "Adapting to the Dark: A Novel Adaptive Low Light Illumination Correction Algorithm for Video Sequences in Wireless Communications." *International Journal of Electrical and Electronics Research* 11, no. Special issue (2023): 01-09.

Annexure

VGG19 is a well-known extension of the VGGNet family, designed by the Visual Geometry Group at the University of Oxford. Compared to VGG16, it employs a deeper structure with a total of 19 layers, including 16 convolutional layers, three fully connected layers, and five max-pooling operations. The layer wise information is shown in the below table 8. A schematic illustration of the VGG16 architecture is presented in the following figure.

Table 8. Showing the Layer wise Information of the VGG-19 Architecture

Layer No.	Layer Type	Kernel Size	No. of Filters / Neurons	Output Size (H × W × D)
1	Conv2D	3 × 3	64	224 × 224 × 64
2	Conv2D	3 × 3	64	224 × 224 × 64
3	MaxPooling	2 × 2	–	112 × 112 × 64
4	Conv2D	3 × 3	128	112 × 112 × 128
5	Conv2D	3 × 3	128	112 × 112 × 128
6	MaxPooling	2 × 2	–	56 × 56 × 128
7	Conv2D	3 × 3	256	56 × 56 × 256
8	Conv2D	3 × 3	256	56 × 56 × 256
9	Conv2D	3 × 3	256	56 × 56 × 256
10	Conv2D	3 × 3	256	56 × 56 × 256
11	MaxPooling	2 × 2	–	28 × 28 × 256
12	Conv2D	3 × 3	512	28 × 28 × 512
13	Conv2D	3 × 3	512	28 × 28 × 512
14	Conv2D	3 × 3	512	28 × 28 × 512
15	Conv2D	3 × 3	512	28 × 28 × 512
16	MaxPooling	2 × 2	–	14 × 14 × 512
17	Conv2D	3 × 3	512	14 × 14 × 512
18	Conv2D	3 × 3	512	14 × 14 × 512
19	Conv2D	3 × 3	512	14 × 14 × 512
20	Conv2D	3 × 3	512	14 × 14 × 512
21	MaxPooling	2 × 2	–	7 × 7 × 512
22	Fully Connected	–	4096	1 × 1 × 4096
23	Fully Connected	–	4096	1 × 1 × 4096
24	Fully Connected	–	1000 (ImageNet classes)	1 × 1 × 1000
25	Softmax	–	–	1 × 1 × 1000



Figure 7. Visualizing VGG-19: Layers, Convolutions, and Max-Pooling