Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

An Efficient Ensemble Deep Learning Model for Binary and Multi-Class Classification of Chest-Xray Images with Web Interface

Dr. M Devi Sri Nandhini[1]*, Dhanishga Kalai[2], S Arun[3] & Dr. G Pradeep[4]

1.Assistant Professor - III, School of Computing, SASTRA Deemed University, Thirumalaisamudram, Tanjore, Tamil Nadu, India. *Corresponding Author Email: nandhini.avcce@gmail.com
2,3.PG Student, School of Computing, SASTRA Deemed University, Thirumalaisamudram, Tanjore, Tamil Nadu, India. Email: [2]126176022@sastra.ac.in, [3]12617608@sastra.ac.in
4.Associate Professor, School of Computing, SASTRA Deemed University, Thirumalaisamudram, Tanjore, Tamil Nadu, India. Email: pradeep.g8@gmail.com

Abstract

Chest diseases like pneumonia, tuberculosis, COVID-19 are serious health related problems and dangerous if they are not detected earlier stage. Using Chest X-ray images, manually to check patients takes time and sometimes leads to errors, so an automatic and reliable system is necessary. In this work, three pre-trained models are used to create an ensemble model that is used to classify the chest X-ray images into binary and multi-class categories. The binary model, which used to predict pneumonia and normal cases, achieved accuracy of 96.63% also training accuracy of 98.04% and validation accuracy of 97.44%. The training and validation losses 6.2% and 7.7% shows that the model is stable and does not overfit. It also achieved evaluation results such as 98.36% recall, 97.45% precision, and 97.91% F1-score. The multi-class model classifies six different chest classes and achieved accuracy of 91.48% also training accuracy of 92.65% and validation accuracy of 90.50%. The training loss 18.67%and validation loss 25.17% show good learning, generalization, and evaluation results shows 91.48% recall, 91.52% precision, and 91.47% F1-score. Overall, the proposed system shows good performance and helps to obtain more accurate diagnosis.

Keywords: *Deep Learning, DenseNet121, MobileNetV2, InceptionV3, Ensemble Model, Web Interface, Chest X-ray.*

1. INTRODUCTION

Chest disease such as Pneumonia, tuberculosis and other health problems in chestleads to suffering of people and cause death at serious conditions. There should be early and accurate diagnosis for these diseases to prevent reaching into serious conditions. Chest x-ray images are commonly used for diagnosis, because of the low-cost and easy to access, however understanding the image was still difficult and that needs help of skilled doctors, in many places experts were not found or maybe not available, this may lead to delay in treatments, and may be also leads to incorrect or inconsistency decision [1].

In recent years, deep learning shows more improvement in analysing the medical image processing using Convolutional Neural Network (CNN), particularly capable of learning automatically patterns like understanding image and reducing the needs of manual reasoning.

Pre-trained models are used because they are effectively extract the meaningful patterns from medical images using transfer learning. These models achieve the promising outcomes in chest X-ray classifications[2] [3]. However, focusing on one model can lead to certain limitations, single model can overfit the data, facing problems in variation of image quality and may not have proper data well across different types of chest X-ray images [4]. So, combining

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

these three models building an ensemble model gives as better accuracy and stronger performance.

Ensemble learning has obtained an attention as a real time solution, instead of depending on the single model. It is the combination of multiple models so it produces the more successive performance. These approaches help to reduces the misclassification and getting better performance by utilizing the strength of various architectures. Several studies shown that the ensemble learning produces the stable performance and showing better results in medical image classification [5].

In this research, deep learning approach suggested for classifying chest X-ray images into categories like binary and multi-class classification. The approach combines Densenet121, MobileNetV2 and InceptionV3 to take the merits of the individual strengths. In addition to that pre processing steps and data augmentation steps are involved to get the stable performance with robust prediction. A biased average technique is used to combine the prediction of each model, resulting in better performance and more stable achievements.

2. LITERATURE SURVEY

Many studies have investigated the needs of deep learning techniques for classification and detection of chest related disease using X-ray images. Early research has concentrated on applying Deep Convolutional Neural Network for multi label disease classification. For in case one research proposed an AI-driven method using DenseNet121, InceptionResNetV2 and Resnet152V2 merging with processed data augmentation to determine up to 14 chest-related diseases. The model reaches insistent ROC-AUC score of 0.80, showcasing impact of Deep CNNs separating meaningful attributes from medical images and optimizing the diagnostic process [6].

For additional improvement, ensemble learning methods have been initiated. A stochastic ensemble-based CNN model was processed for COVID-19 detection, where multiple deep learning models were merged to learn a stable latent feature representation. This method reaches an accuracy of 0.91 and AUC of 0.97, illustrating that ensemble approach boost generalization and lowering the misclassification in chest X-ray image analysis [7].

Focusing limitations such as data discrepancies and data imbalance, the DeTrac approach was initiated a class decomposition process that separates classes into many sub-classes, this model achieves an accuracy of 93.1%with high sensitivity, demonstrating and managing data distribution problems considerably improving the model performance in chest X-ray image for classification of COVID-19[8].

Latest development has also engrossed on improving diagnostic performance and transparency. A comprehensive study using Densenet121 and ResNet50 illustrating that deep learning models can perform similar to radiologists to identify multiple thoracic diseases. The unification of explainable AI techniques such as SHAP and LIME additionally develops the transparency and the model achieves an AUC of 94%, shaping the model more worthy for real-world medical applications [9].

Further these methods, transfer learning with VGG-based architectures has shown strong implementation in multi-class disease classification in lung using chest X-ray. A model merging VGG19 with CNN layer was initiated to classify disease such pneumonia, COVID19 and lung cancer. This approach achieved an accuracy of 96.48% and AUC of 99.82%, showing

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

that hybrid architectures can efficiently increase the classification accuracy and hold up speedy diagnosis [10].

For additional boost feature extraction, attention-based mechanism has been combined into deep learning models. An attention-enhanced ResNet50 architecture was process to focus on important regions in chest X-ray images for Pneumonia detection using spatial and channel attention mechanisms. The attention mechanism gradually increases the model's ability to get relevant features and improves the performance of model to achieve 98% accuracy and handled imbalance dataset[11].

The importance of interpretability in medical AI has been tackled through models integrating GRAD-CAM and Guided GRAD-CAM techniques. A decision support system merging classification and segmentation was suggested for COVID-19 diagnosis using CT scan images, the model achieves a performance of 98.51% and given visual explanations of predictions, authorizing clinicians to better recognition and trust AI-based decisions [12].

Enhanced deep learning approach for pneumonia detection using chest X-ray images by combining attention mechanisms such as Convolutional Block Attention Module (CBAM) and Squeeze-and-Excitation (SE) network with CNN architectures. The results demonstrates that the CBAM-enhanced CNN reaches a high accuracy of 98.6%, Substantially outperforming the baseline CNN 92.08%, while the SE-based model reaches 96.25% accuracy. The attention mechanism helps the model targets on important regions such as infected lung areas, increasing feature extraction and decreasing overfitting. This research also highlights real-time fitness by improving a lightweight web-based system for medical purpose, significantly in resource-limited settings [13]. Table 1 shows the models used in each research and their accuracy

Table 1: The models and their per

#	Model	Accuracy
1	DenseNet121, InceptionResNetV2, ResNet152V2	0.80
2	Stochastic ensemble of CNNs	91 %
3	DeTraC (Transfer learning, Class decomposition)	93.10 %
4	DenseNet121, ResNet50, SHAP, LIME	94%
5	VGG19+CNN	96.48%
6	ResNet50 with attention mechanisms	98%
7	GGECS (Guided Grad-CAM based Explainable Classification and Segmentation system)	98.51%
8	CNN+CBAM	98.6%

3. METHODOLOGY

3.1 Dataset Description

The dataset used in this research consists of chest X-ray images available publicly in Kaggle. Here the classification tasks are included as binary and multi-class classification. In binary classification, the dataset is categorized into Normal and Pneumonia with a total of 5,856 images [14].

In multi-class classification the dataset is categorized into six that isCOVID-19, Emphysema, Normal, Pneumonia-Bacterial, Pneumonia-Viral and Tuberculosis with a total of 18,036 [15]. Class imbalance was controlled using class weighting techniques. The bar graphs of Fig. 1 and Fig. 2 represent the total number of images in each of binary and multi-class dataset.

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

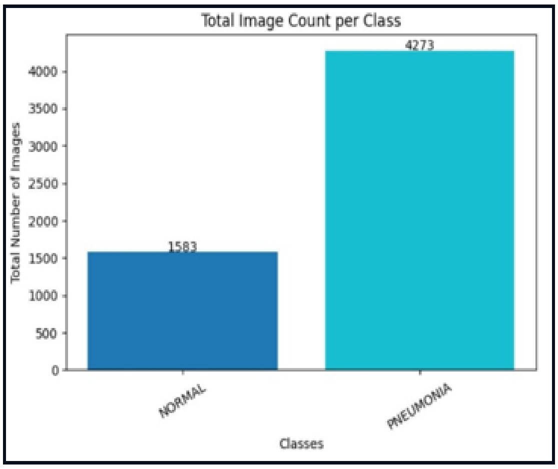

Fig 1: Binary Dataset

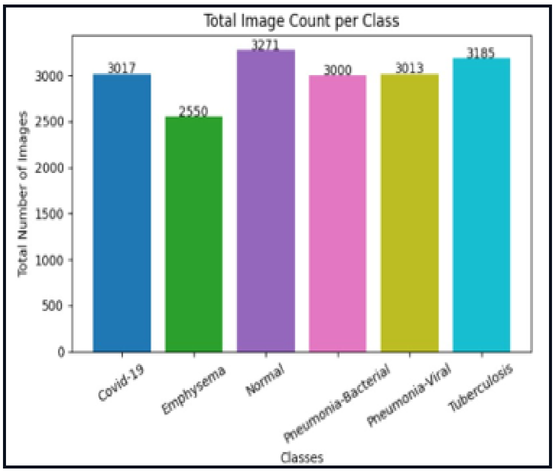

Fig 2: Multi-class Dataset

3.2 Data Pre-processing and Augmentation

All chest-Xray input images are pre-processed to verify consistency across different types of models, the images are resized to 224x224 pixels, which is needed by the particular pre-trained model. Pixel values are normalized to range of 0 to 1 to increase training stability. Even though chest X-ray images are normally gray scale, they are modified into three-channel images to equalize the model input requirements.

These methods verify uniform input and help the model learn meaningful features. Data augmentation was applied to improve the various training data and decrease overfitting. Techniques involved such as rotation, horizontal flipping, zooming, and shifting are used to create variations of the original images. These transformations replicate real-world variations in medical image conditions. Augmentation was used only to the training dataset and to control integrity of validation and test sets. As a result, the model enhances more robust and stable. Here Fig.3 and Fig. 4 shows the sample images before and after data pre-processing and augmentation applied [16].

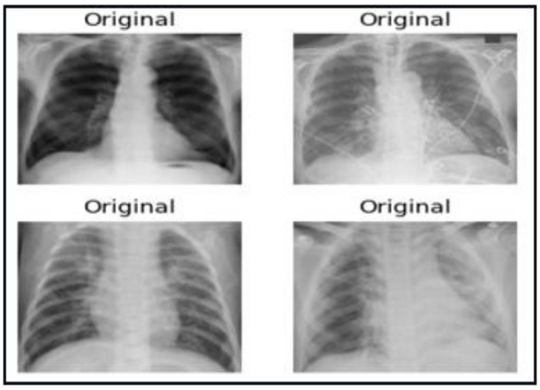

Fig 3: Original Images

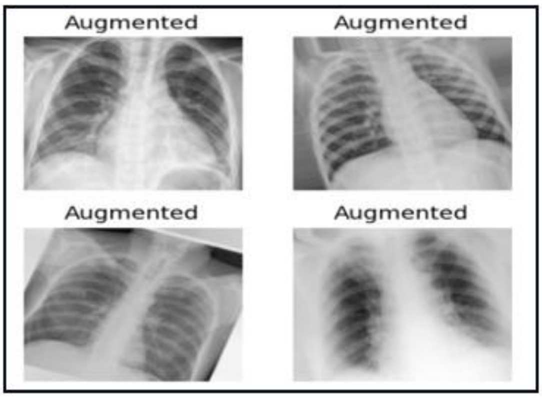

Fig 4: Augmented Images

3.3 Model Architecture

The proposed model is built by combining three pre-trained models such DenseNet121[17], MobileNetV2[18]and InceptionV3[19]. The input for the model is given as

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

chest X-ray image of size 224 x 224 x 3.The received input image is given to the three pre-trained model at the same time. Already a large dataset (Image Net) is trained in these models and useful features are extracted from the chest x-ray images using these models. Each of model focuses on different features such as detailed patterns, lightweight features, multi-scale information. Combining all model features together makes the model to learn and perform better also extracts more information from the chest X-ray images.

To reduce overfitting and make the model to perform well a fine-tuning technique is used. First, frozen all the layers of the three models so the learned patterns are not disturbed. Then, only using the last correctly needed number of layers of each model to train for binary and multi-class classification. This helps the model to adjust the important features need for the medical image dataset while keeping the other basic extracted features stable[20].

After extracting all features and outputs using the three models, they are converted into a smaller vector using Global Average Pooling. These converted vectors are then combined into a single feature vector using concatenation. This is called feature fusion, where the features of all three models are combined. Then these are passed to the batch normalization to make the training stable and steady. A dropout layer is added to prevent overfitting. Also fully connected layers uses ReLU activation to learn patterns. ReLU uses formula (1)

$$f(x) = \max(0, x) \rightarrow (1)$$

At last, the model classifies the input and give as the output. For binary classification, a single neuron with sigmoid activation is used. For multi-class classification, multiple neurons with softmax activation are used. This helps the model to handle both types of classification by using the same structure. Formula of sigmoid and softmax is given in (2) and (3)

$$\sigma = \frac{1}{1 + e^{-x}} \rightarrow (2)$$

$$\sigma(z_i) = \frac{e^{\{z_i\}}}{\sum_{j=1}^{n} e^{\{z_j\}}} \rightarrow (3)$$

3.4 Feature Extraction

Feature extraction is the important step in the proposed model to learn patterns from chest X-ray images. Here we are extracting features using three different pre-trained models. This helps in focusing and capturing more important features for the chest X-ray image. Each model extracts different features so, the models learn well, understand hard patterns, and gives as better performance.

DenseNet121 is very effective in capturing and identifying small patterns also fine changes in the lung regions. These features help in detecting the early signs of diseases that are not clearly visible. This model gives more detailed features. MobileNetV2 is used to extract lightweight spatial features efficiently. It uses less parameters so it is faster compared to others. It captures spatial features between different regions. It balances accuracy and performance. InceptionV3 focus on multi-scale features by analysing the image in different practice. It uses multiple filter size to identify small and big patterns. It can focus on both local and global features effectively. After all features are extracted, Global Average Pooling (GAP) is applied to all three model’s output. This converts the feature maps into a smaller feature vector by taking the average of each model’s value. This helps to keep the important features entirely in

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

each of the model. These are used to learn different patterns and make the model more efficient[21].

3.5Ensemble Learning Approach

The proposed model uses an ensemble learning approach; it merges all the extracted features of the pre-trained models such as DenesNet121, MobileNetV2, InceptionV3. Each model has their own extracted different type of patterns from the chest X-ray image and they are combined using concatenation. This creates a single compact feature vector that contains the important patterns of all three models. The merged feature vector is passed to the classification layer. In binary classification, the outputs get the probabilities of normal or pneumonia. In multi-class classification, the outputs match to probabilities across six classes. These predictions are merged using a weighted averaging technique to make the actual result. This model decreases the boundaries of particular model and increases overall accuracy. Ensemble learning approach improves robustness and makes sure to have the stable performance across the various types of datasets[22].

3.6 Working of the Model

First, get the input chest X-ray image and resize it. Next, pre-processing and augmentation are applied to make the image quality to improve. Then the processed image is sent to the three models at the same time. Each model extracts features from the input image. These extracted features are merged and form as a single representation. Now this merged feature is passed to dense layer for classification. Finally, the model classifies the image for binary classification using sigmoid and for multi-class classification using softmax. The work flow of the model is shown inFig. 5

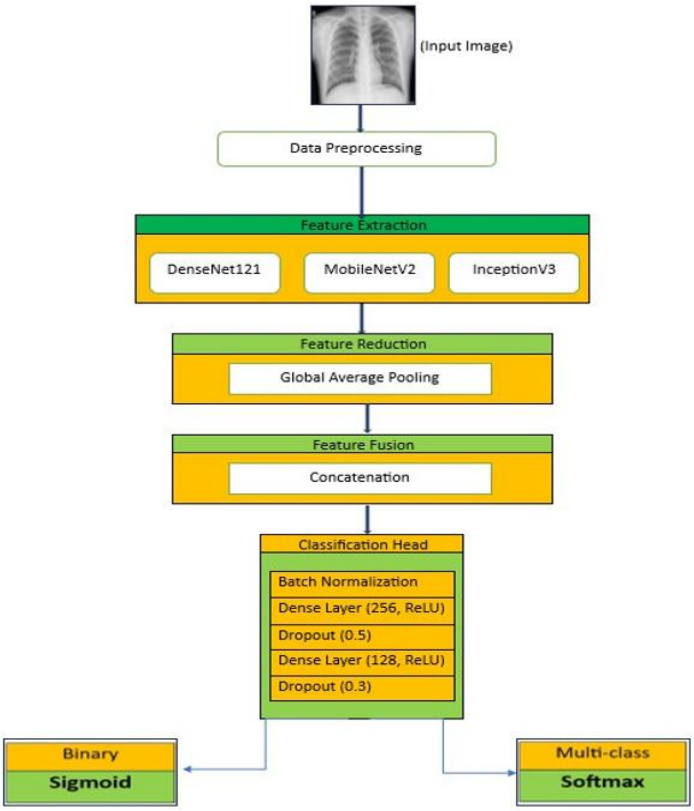

Fig 5: Workflow of the Proposed Model

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

4. RESULTS AND DISCUSSION

4.1 System & Experimental Setup

The proposed model is trained and tested using Kaggle Notebook environment with NVIDIA Tesla T4 GPU with 16 GB RAM. Python with TensorFlow, Keras libraries are used to develop the model. The binary dataset is divided into 80%, 10%, 10% and multi-class dataset is kept as it is respectively. Used batch size is 32 for both and learning rate 0.0001 also with adaptive learning rate to make the model stable. 10 epoch for binary and 15 for multi-class classification with Cross-entropy loss using Adam optimizer.

4.2Evaluation Overview

To completely value the performance of the proposed ensemble model, evaluation metrics were used such as accuracy, recall, precision, specificity, and F1-score these metrics are used to check the performance and make it detailed the understanding of the model's forecasting abilities, especially in medical image classification tasks.

Accuracy indicates the inclusive effectiveness of the model by measuring the ratio of correctly classified samples to the total number of samples, it mathematically calculated using equation (4)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \rightarrow (4)$$

Precision evaluates the proportion of correctly predicted positive cases from all cases predicted as positive and it is calculated using the equation (5)

$$Precision = \frac{TP}{TP + FP} \rightarrow (5)$$

Recall or Sensitivity quantifies the proportion of actual positive cases that are correctly identified by the model. It is calculated using equation (6)

$$Recall = \frac{TP}{TP + FN} \rightarrow (6)$$

F1-Scoreis the harmonic mean of precision and recall, making a single metric that balances both false positive and false negative, it is calculated using the equation (7)

$$F1\ Score = 2 * \frac{Precison * Recall}{Precision + Recall} \rightarrow (7)$$

All the TP, TN, FP, FN are used to represent True Positive, True Negative, False Positive and False Negative respectively where TP and TN indicate correctly classified instances, whereas FP and FN indicate misclassifications.

For additional evaluation the performance of the given model, the Receiver operating characteristics (ROC) curve and the Area Under the Curve (AUC) are applied, these metrics makes a comprehensive analysis of the model's classification capability with different decision threshold, providing important in medical chest X-rays images classification task. The AUC makes a single scalar value that utilizes the performance of the model across all the thresholds. A higher AUC values shows that the model is better at differentiating classes, particularly, it throws back the probability that the model ranks a differently chosen positive case higher that a particularly chosen negative cases.

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

4.3 Binary Classification

Here we are going to discuss and compare our proposed model results with the each of the single pre-trained model using the Classification Performance, Confusion matrix, ROC Curve and the Training and Validation graphs [23].

4.3.1 Classification Performance

Table2 shows the performance of the three pre-trained models and the proposed model using the evaluation metrics. Our proposed model has performed better than the other three pre-trained models and achieves accuracy of 96.93%.

Table 2: Evaluation Metric of All Models

Model	Accuracy	Recall	Precision	F1- Score
DenseNet121	92.66%	90.65%	99.23%	94.75%
MobileNetV2	94.54%	98.13%	94.59%	96.33%
InceptionV3	94.88%	94.16%	98.77%	96.41%
Proposed Model	**96.93%**	**98.36%**	**97.45%**	**97.91%**

4.3.2 Confusion Matrix

Fig. 5a shows the confusion matrix of DenseNet121 similar Fig. 5b, Fig. 5c shows matrix of MobileNetV2, InceptionV3.The confusion matrix of the proposed model shown in Fig. 5d and the model have a high number of correctly classified instances. Finally, the confusion matrix analysis of the model executes reliable changes with minimal misclassification and helps to get high values observed in accuracy, precision, and recall. In matrix 0 denotes Normal and 1 denotes Pneumonia class.

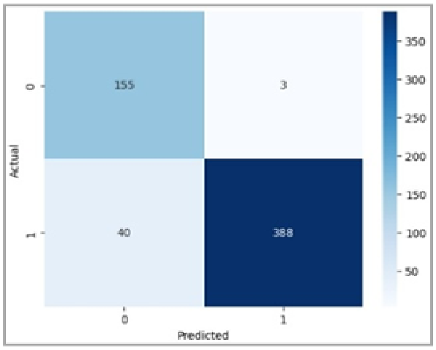

Fig 5a: DenseNet121

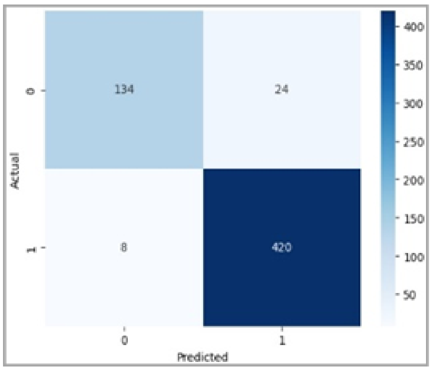

Fig 5b: MobileNetV2

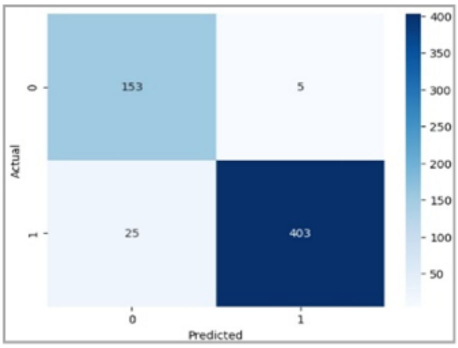

Fig 5c: InceptionV3

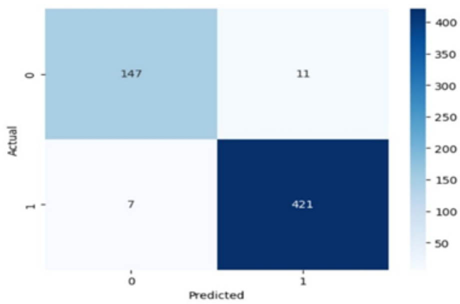

Fig 5d: Proposed Model

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

4.3.3 ROC Curve AUC Analysis

Fig. 6a, Fig. 6b, Fig. 6c shows ROC-AUC score of the pre-trained models. Fig. 6d. shows the ROC-AUC score of the Proposed model, indicating its excellent discriminating capability across various classification thresholds.

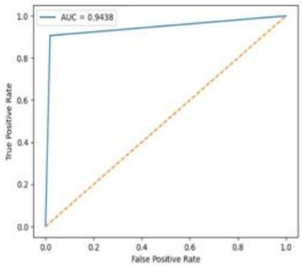

Fig 6a: DenseNet121

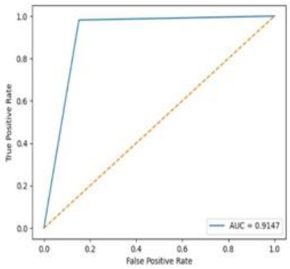

Fig 6b: MobileNetV2

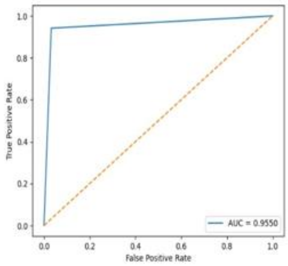

Fig 6c: InceptionV3

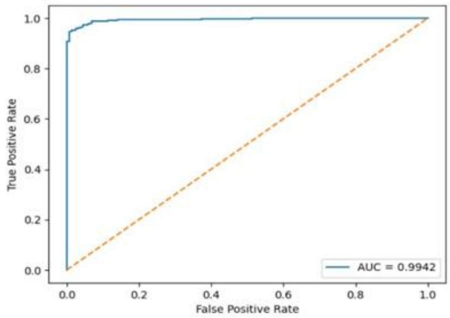

Fig 6d: Proposed Model

4.3.4 Performance of Training and Validation

Fig. 7a and Fig. 7b shows the training and validation of DenseNet121 similarly Fig. 8a and Fig. 8b, Fig. 9a and Fig. 9b shows the training and validation of MobileNetV2 and InceptionV3. Fig. 10a shows that the proposed model accomplishes a training accuracy of 98.04% and a validation accuracy of 97.44% this shows strong learning capability and effective generalization on unobserved data, the model records a training loss of 6.2% and validation loss of 7.7% as shown in Fig. 10b, indicating stable convergence with very small quantity of overfitting.

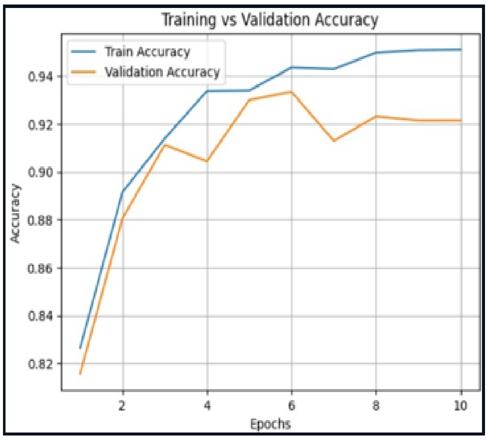

Fig 7a: Train & Val Acc of DenseNet121

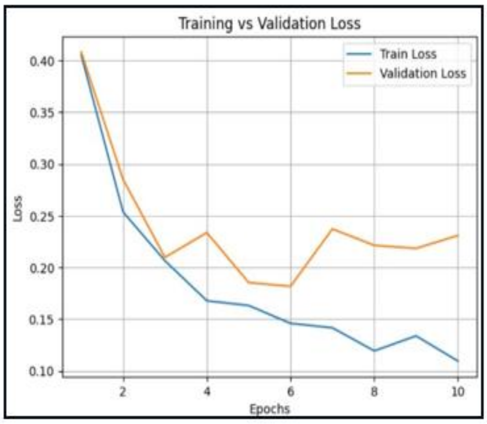

Fig 7b: Train & Val Loss of DenseNet121

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

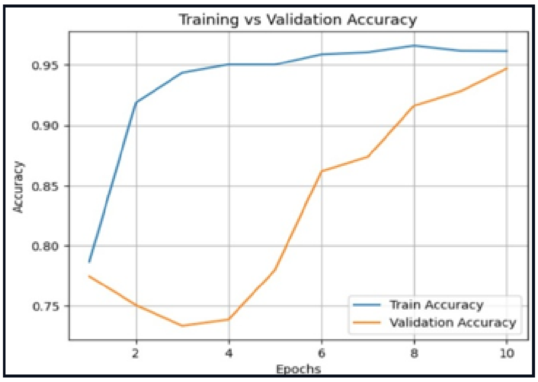

Fig 8a: Train & Val Acc of MobileNetV2

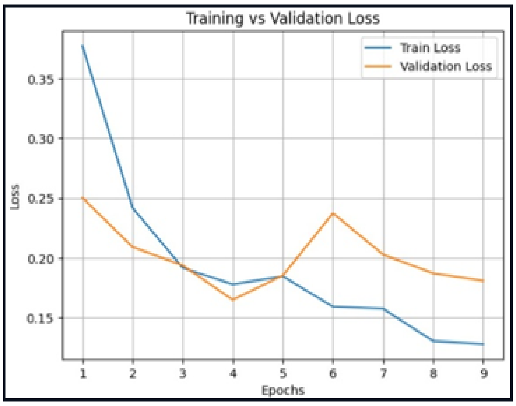

Fig 8b: Train & Val Loss of MobileNetV2

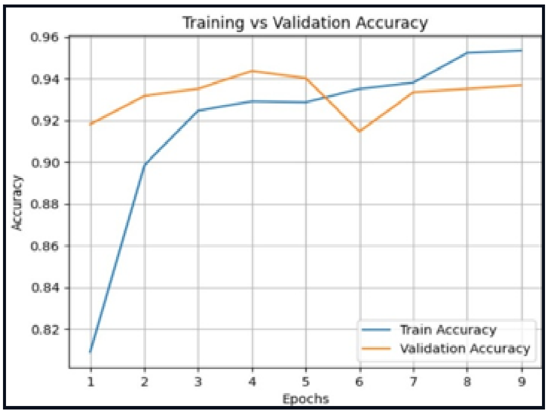

Fig 9a: Train & Val Acc of InceptionV3

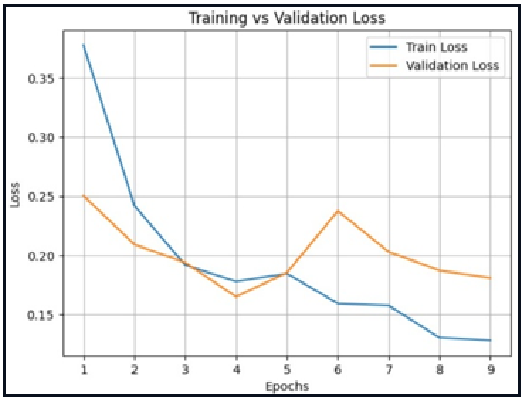

Fig 9b: Train & Val Loss of InceptionV3

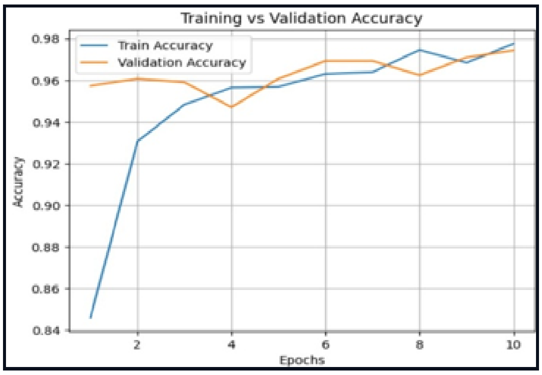

Fig 10a: Train & Val Acc of Proposed Model

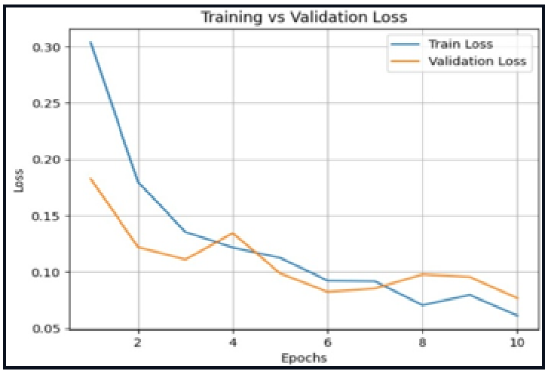

Fig 10b: Train & Val Loss of Proposed Model

4.4 Multi-Class Classification

Here we are going to discuss and compare our proposed model results with the each of the single pre-trained model using the Classification Performance, Confusion matrix, ROC Curve, Class-wise Performance and the Training and Validation graphs [24].

4.4.1 Performance of Classification

Table 3 shows the performance of the three pre-trained models and the proposed model using the evaluation metrics. Our proposed model has performed better than the other three pre-trained models and achieves accuracy of 91.48%.

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

Table 3: Evaluation Metrics of Pre-trained and Proposed model

Model	Accuracy	Recall	Precision	F1- Score
DenseNet121	88.26%	88.26%	88.63%	88.02%
MobileNetV2	87.74%	87.74%	88.25%	87.55%
InceptionV3	85.66%	85.66%	85.76%	85.47%
Proposed Model	**91.48%**	**91.48%**	**91.52%**	**91.47%**

4.4.2 Confusion Matrix

Fig. 11a shows the confusion matrix of DenseNet121 similar Fig. 11b, Fig. 11c shows matrix of MobileNetV2, InceptionV3.The confusion matrix shown in Fig. 11d is the proposed model, makes a class-wise evaluation of the model's performed across six classes, the diagonal features of the matrix get a high number to accurate predictions for all classes, demonstrating the model correctly differentiate disease categories. It also shows that no single class controls the prediction errors, demonstrating that the model is not biased toward any specific class. This coordinates with the steady precision, recall and f1-score values acquired during evaluation. In the confusion matrix 0 denotes Covid-19, 1 denotes Emphysema, 2 denotes Normal, 3 denotes Pneumonia-Bacterial, 4 denotes Pneumonia-Viral and 5 denotes Tuberculosis

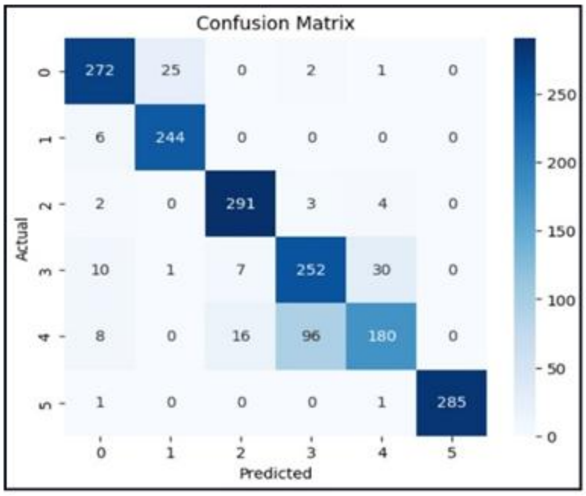

Fig 11a: DenseNet121

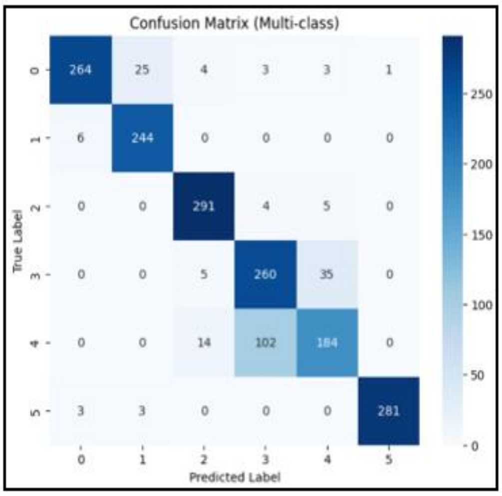

Fig 11b: MobileNetV2

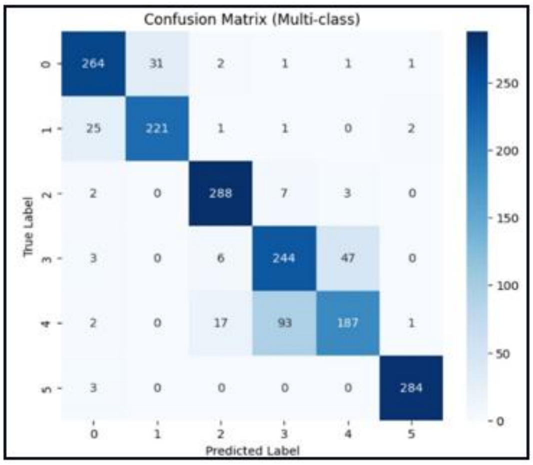

Fig 11c: InceptionV3

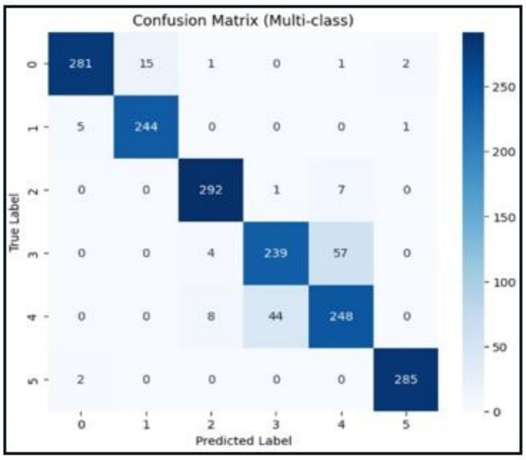

Fig 11d: Proposed Model

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

4.4.3 ROC Curve

The ROC-AUC scores of each classes shows the model's strength, also the model shows that each class is well distinguish from other in feature space. Fig. 12a, Fig. 12b, Fig. 12c and Fig. 12dshows the comparison between the pre-trained models DenseNet121, MobileNetV2, InceptionV3 and the proposed model

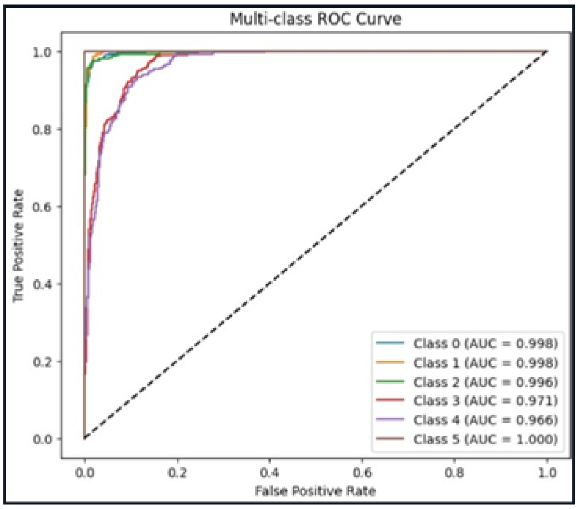

Fig 12a: DenseNet121

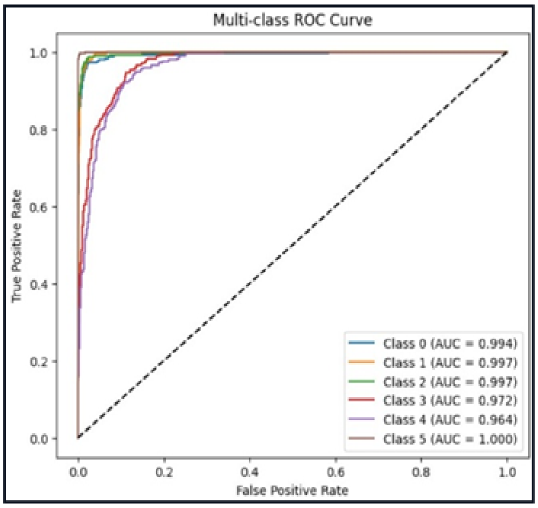

Fig 12b: MobileNetV2

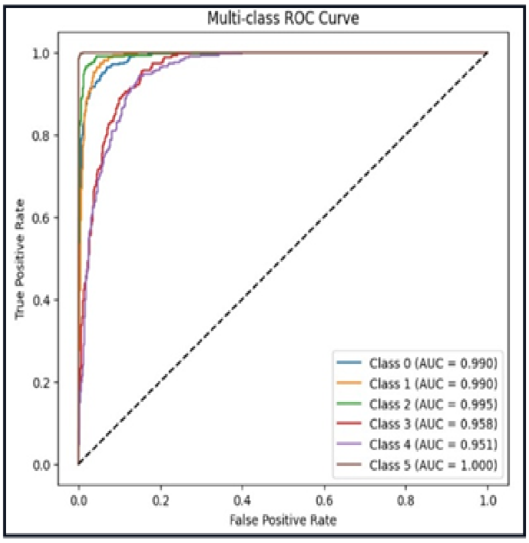

Fig 12c: InceptiionV3

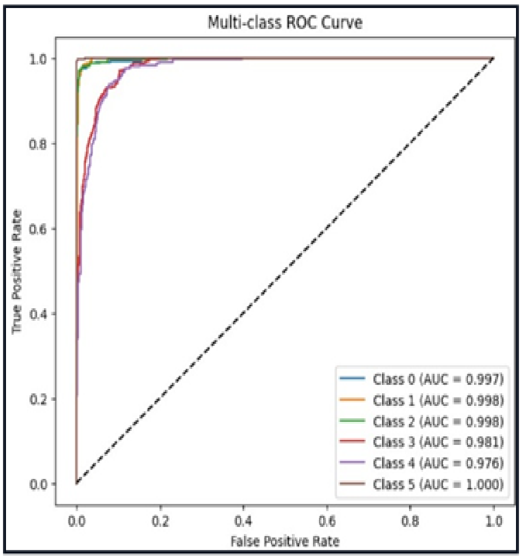

Fig 12d: Proposed Model

4.4.4 Class-wise Performance

The Proposed model of multi-class achieves a best performance in each of the classes. Particularly Tuberculosis has shown a higher performance compared to others. Although Pneumonia subclasses face some challenges due similarity in both classes. Table 4 shows the performance using evaluation metrics of each class.

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

Table 4: Evaluation Metrics of Each Classes

Class	Precision	Recall	F1-Score
Covid-19	0.98	0.94	0.96
Emphysema	0.94	0.98	0.96
Normal	0.96	0.97	0.97
Pneumonia-Bacterial	0.84	0.80	0.82
Pneumonia-Viral	0.79	0.83	0.81
Tuberculosis	0.99	0.99	0.99

4.4.5 Training and validation Analysis

In Fig. 13a and Fig. 13b shows DenseNet121 Training and Validation similarly Fig. 14a&Fig. 14b and Fig. 15a&Fig. 15b shows the Training and Validation of MobileNetV2 and InceptionV3. Fig.16a the proposed model shows a training accuracy of 92.65% and validation accuracy of 90.50%, which gives a strong performance among various six classes.

Fig. 16b shows the training loss 18.67% and validation loss 25.17% indicates a reducing trend during the model training, validation loss approximately higher than training loss, this shows expected and indicates well convergence and good generalization.

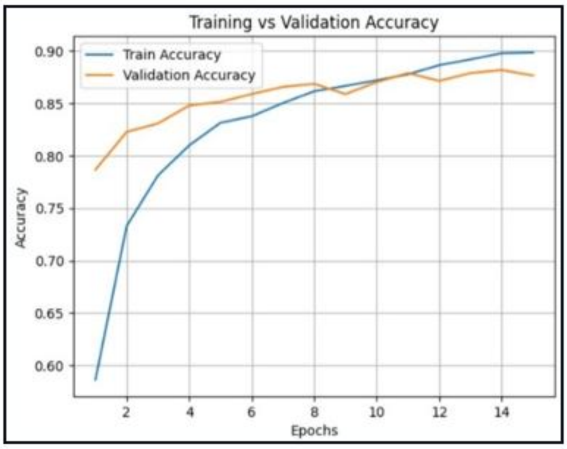

Fig 13a: DenseNet121 Accuracy

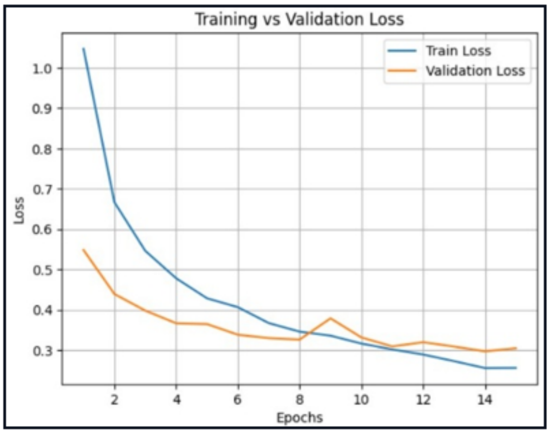

Fig 13b: DenseNet121 Loss

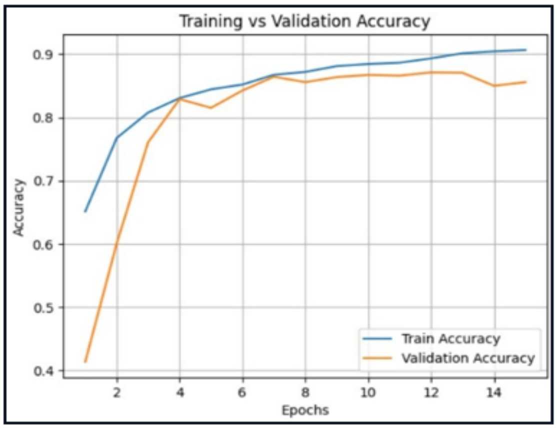

Fig 14a: MobileNetV2 Accuracy

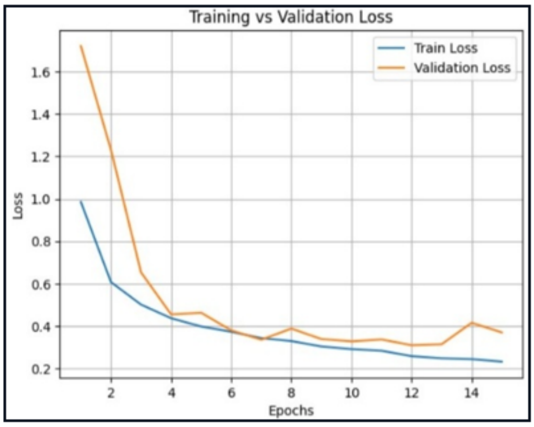

Fig 14b: MobileNetV2 Loss

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

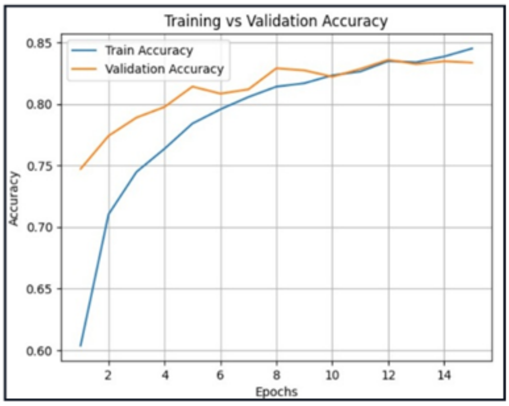

Fig 15a: InceptionV3 Accuracy

Fig 15b: InceptionV3 Loss

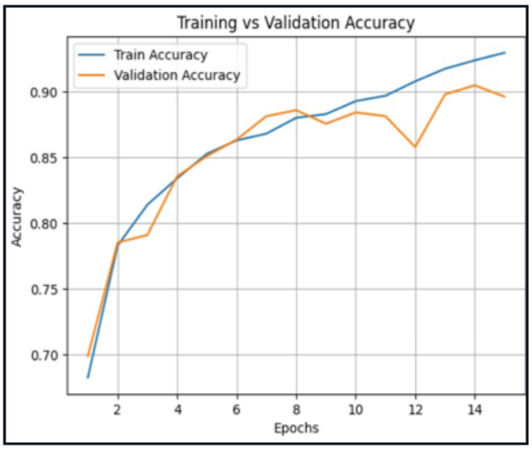

Fig 16a: Proposed Model Accuracy

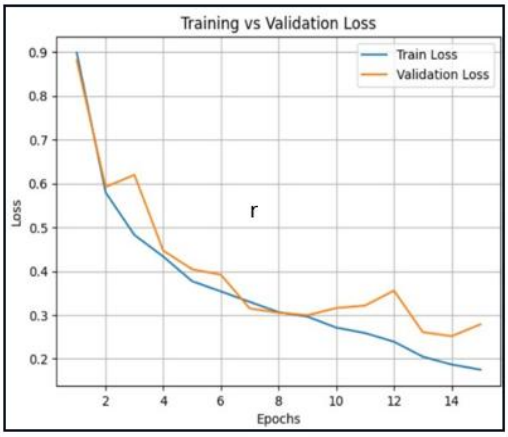

Fig 16b: Proposed Model Loss

4.5 Summary of Both Models

In the binary proposed model, the model accomplishes accuracy of 96.93%. With regarding of classification performance, the model gives a precision of 97.45%, showing a low rate of false positive predictions, and a recall of 98.36%, showing its ability to correctly predicting the majority of positive cases. The F1-score of 97.91% shows a balanced trade-off between precision and recall. In multi-class proposed model, the model shows the accuracy of 91.48%.

In spite of the increased complexity of multi-class problem. The precision of 91.52% shows that the model makes a low false positive rate across all classes, Likewise the recall of 91.48% shows that the model successfully indicates most occurrence of each class, even with limited cases The F1-score of 91.47% shows that the model has a balanced trade-off between precision and recall for all categories, indicating uniform performance without bias toward any specific class.

5.WEB APPLICATION

5.1 Binary classification Interface

Fig 17 shows the interface of binary classification that contains an upload button and the result area shows the classification of patient chest X-ray either pneumonia or normal

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

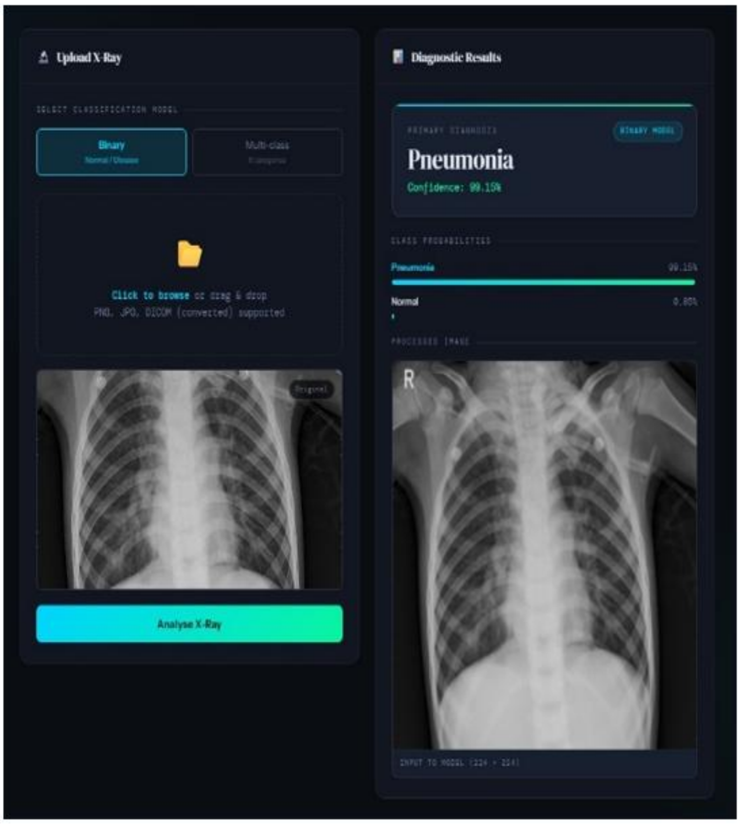

Fig 17: Binary Classification Interface

5.2 Multi-class classification Interface

Fig 18 shows the interface of multi-class classification with the upload section and result section class of the patient's X-ray into one among the six-classes.

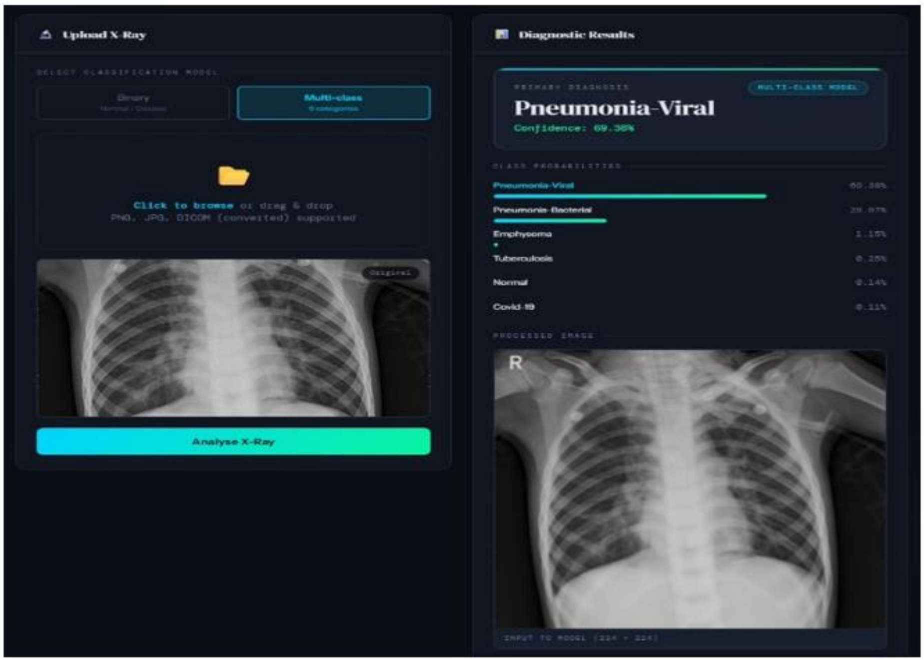

Fig 18: Multi-Class Classification Interface

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

6. CONCLUSION AND FUTURE WORK

The ensemble model merges DenseNet121, MobileNetV2 and InceptionV3 to get classification performance on medical images, with the help of feature-level fusion, the model attains diverse and complementary features from multiple architectures, the partial fine-tuning technique helps to modify the model to the dataset by preventing overfitting. The binary classification model gets higher performance with simpler techniques. the multi-class model also exhibits strong and balanced result with all six classes, evaluation metrics such as accuracy, precision, recall and f1-score verifies the model's reliability. The low validation loss and high validation accuracy gets good generalization. ROC-AUC gets additionally validate the model's stability and robustness. Finally, the model is suitable for real-world healthcare applications. For future work can focus on increasing dataset diversity, controlling class imbalance, including explainable AI techniques, enhancing the model for real-time deployment, and improving performance with the attention mechanisms.

References

1) Debnath SK, Debnath M, Srivastava R. Opportunistic etiological agents causing lung infections: emerging need to transform lung-targeted delivery. Heliyon 2022;8 (12):e12620.
2) Hashmi, M. F., Katiyar, S., Keskar, A. G., Bokde, N. D., & Geem, Z. W. (2020). Efficient pneumonia detection in chest xray images using deep transfer learning. *Diagnostics*, *10*(6), 417.
3) El Asnaoui, K. (2021). Design ensemble deep learning model for pneumonia disease classification. *International Journal of Multimedia Information Retrieval*, *10*(1), 55-68.
4) Tay, Y. X., Kothan, S., Kada, S., Cai, S., & Lai, C. W. K. (2021). Challenges and optimization strategies in medical imaging service delivery during COVID-19. *World journal of radiology*, *13*(5), 102.
5) Prakash, J. A., Ravi, V., Sowmya, V., & Soman, K. P. (2023). Stacked ensemble learning based on deep convolutional neural networks for pediatric pneumonia diagnosis using chest X-ray images. *Neural Computing and Applications*, *35*(11), 8259-8279.
6) Albahli, S., Rauf, H. T., Algosaibi, A., & Balas, V. E. (2021). AI-driven deep CNN approach for multi-label pathology classification using chest X-Rays. *PeerJ Computer Science*, *7*, e495.
7) Arora, R., Bansal, V., Buckchash, H., Kumar, R., Sahayasheela, V. J., Narayanan, N., ... & Raman, B. (2021). AI-based diagnosis of COVID-19 patients using X-ray scans with stochastic ensemble of CNNs. *Physical and Engineering Sciences in Medicine*, *44*(4), 1257-1271.
8) Abbas, A., Abdelsamea, M. M., & Gaber, M. M. (2021). Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Applied Intelligence*, *51*(2), 854-864.
9) Sufian, M. A., Hamzi, W., Sharifi, T., Zaman, S., Alsadder, L., Lee, E., ... & Hamzi, B. (2024). AI-driven thoracic X-ray diagnostics: transformative transfer learning for clinical validation in pulmonary radiography. *Journal of personalized medicine*, *14*(8), 856.

Volume 65 | Issue 05 | May 2026
DOI: 10.5281/zenodo.20178387

10) Malik, H., Anees, T., Din, M., & Naeem, A. (2023). CDC_Net: Multi-classification convolutional neural network model for detection of COVID-19, pneumothorax, pneumonia, lung Cancer, and tuberculosis using chest X-rays. *Multimedia Tools and Applications*, *82*(9), 13855-13880.

11) Li, D. (2024). Attention-enhanced architecture for improved pneumonia detection in chest X-ray images. *BMC Medical Imaging*, *24*(1), 6.

12) Panwar, H., Gupta, P. K., Siddiqui, M. K., Morales-Menendez, R., Bhardwaj, P., & Singh, V. (2020). A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. *Chaos, Solitons & Fractals*, *140*, 110190.

13) Potharaju, S., Tambe, S. N., Dasari, K., Srikanth, N., Venkatarao, R., & Tambe, S. (2025). Enhanced X-ray image Classification for Pneumonia Detection Using Deep Learning Based CBAM and SE Mechanisms. *Intelligence-Based Medicine*, 100299.

14) https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia

15) https://www.kaggle.com/datasets/mohamedasak/chest-x-ray-6-classes-dataset

16) Monshi, M. M. A., Poon, J., Chung, V., & Monshi, F. M. (2021). CovidXrayNet: Optimizing data augmentation and CNN hyperparameters for improved COVID-19 detection from CXR. *Computers in biology and medicine*, *133*, 104375.

17) Arulananth, T. S., Prakash, S. W., Ayyasamy, R. K., Kavitha, V. P., Kuppusamy, P. G., & Chinnasamy, P. (2024). Classification of paediatric pneumonia using modified DenseNet-121 deep-learning model. *IEEE Access*, *12*, 35716-35727..

18) An evaluation of lightweight deep learning techniques in medical imaging for high precision COVID-19 diagnostics. Healthc.Anal. 2022, 2, 100096

19) Mujahid, M., Rustam, F., Álvarez, R., Luis Vidal Mazón, J., Díez, I. D. L. T., & Ashraf, I. (2022). Pneumonia classification from X-ray images with inception-V3 and convolutional neural network. *Diagnostics*, *12*(5), 1280.

20) Yanar, E., Hardalaç, F., & Ayturan, K. (2025). Pelm: A deep learning model for early detection of pneumonia in chest radiography. *Applied Sciences*, *15*(12), 6487.

21) Varshni, D., Thakral, K., Agarwal, L., Nijhawan, R., & Mittal, A. (2019, February). Pneumonia detection using CNN based feature extraction. In *2019 IEEE international conference on electrical, computer and communication technologies (ICECCT)* (pp. 1-7). IEEE.

22) Masud, M., Bairagi, A. K., Nahid, A. A., Sikder, N., Rubaiee, S., Ahmed, A., & Anand, D. (2021). A pneumonia diagnosis scheme based on hybrid features extracted from chest radiographs using an ensemble learning algorithm. *Journal of Healthcare Engineering*, *2021*(1), 8862089.

23) Hasija, S., Akash, P., Hemanth, M. B., Kumar, A., & Sharma, S. (2022). A novel approach for detection of COVID-19 and Pneumonia using only binary classification from chest CT-scans. *Neuroscience Informatics*, *2*(4), 100069.

24) Alshmrani, G. M. M., Ni, Q., Jiang, R., Pervaiz, H., & Elshennawy, N. M. (2023). A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images. *Alexandria Engineering Journal*, *64*, 923-935.