International Journal of Basic and Applied Sciences, 14 (7) (2025) 149-160

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/3gepr148
Research paper



Optimized Silicosis Detection Using a Feature-Enriched Multi-Model Ensemble Framework with An Enhanced U-Net Architecture

Shivaanivarsha. N [1,2] *, Kavipriya. P [1]

[1] *Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, India*
[2] *Department of Electronics and Communication Engineering, Sri Sairam Engineering College, Chennai, India*
**Corresponding author E-mail: varsha.ece@sairam.edu.in*

Received: August 23, 2025, Accepted: October 23, 2025, Published: November 5, 2025

Abstract

Silicosis remains a major occupational health concern worldwide. This Research proposes an ensemble learning model combined with an improved U-Net for silicosis detection and classification. The dataset included 1300 healthy lung images and 1328 augmented silicosis images, split 80:20 for training and testing. The ensemble model integrated Inception V2, Inception V3, MobileNet, and ResNet50 for robust classification. Enhanced U-Net with skip connections and optimized hyperparameters achieved a validation accuracy of 98.33% and a Dice coefficient of 0.967. The ensemble outperformed standalone models, achieving an F1-score of 95.57%, a precision of 96.42%, and an MCC of 0.9275. Training was conducted using an NVIDIA i7 CPU, with R-Studio for analytics and visualization. Comparative analysis confirmed the model's superiority, aiding radiologists in faster decision-making.

Keywords: *Crystalline silica, Deep learning, Ensemble Model, Lung disease, Silicosis, U-Net*

1. Introduction

Silicosis, a serious occupational lung disease caused by prolonged exposure to crystalline silica dust, is rising due to a lack of awareness. Workers in mining, construction, and quarry industries are at high risk, with silica particles smaller than 10 micrometers posing severe health hazards. Over 10 million workers in India face silicosis risks, with OSHA setting the Permissible Exposure Limit at 0.1 mg/m³ for pure quartz silica [1-3].

Silica exists in three forms: Crystalline silica (found in soil, sand, and granite), Microcrystalline silica (composed of tiny quartz crystals), and Amorphous silica (less hazardous but still concerning). Silicosis is categorized into four types based on severity: Chronic, Accelerated, Complicated, and Acute silicosis. Diagnosis relies on imaging techniques like Chest X-ray, CT scan, Bronchoscopy, Lung Biopsy, and Sputum tests, with X-ray being the most commonly used method [4-8].

Due to limited annotated datasets for silicosis, CNN architectures such as VGG, DenseNet, and ResNet have been widely used, with hybrid models integrating transformers or RNNs for improved detection. Attention-based CNNs enhance feature extraction, while transfer learning refines detection accuracy. U-Net-based segmentation helps identify fibrotic lung patches, aiding severity assessment.

This research employs an ensemble learning model combining classification and segmentation techniques to enhance silicosis detection accuracy.

2. Literature Review

Several studies have explored AI and image processing techniques for detecting lung diseases, including silicosis, pneumoconiosis, and lung cancer. Early methods focused on conventional segmentation and feature extraction. For instance, Gao et al. (2007) [9] proposed a fully automated lung segmentation approach for detecting nodules in CT images using anisotropic diffusion, thresholding, and morphological smoothing. Chanda and Sarkar (2020) [10] applied adaptive thresholding and Gabor filtering for pattern recognition in diseased lung regions, achieving over 90% precision. Texture-based feature extraction using the Grey Level Co-Occurrence Matrix (GLCM) was explored by Saad et al. (2021) [11], demonstrating promising results despite segmentation inconsistencies. Deep learning methods further improved detection accuracy: Zhang et al. (2021) [12] developed a predictive model for COVID-19 screening, while Shivaanivarsha and Kavipriya (2023) [13] combined ResNet50 and Xception models for general medical image analysis. Pneumoconiosis-specific studies include Devnath et al. (2022a, 2022b) [4-5], Li et al. (2024) [3], and Yang et al. (2021) [8], which utilized ensemble learning, multi-dimensional deep features, and CNN-based subregion segmentation, achieving up to 0.973 accuracy and outperforming radiologists.

Copyright © N Shivaanivarsha et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recent advancements in transformer-based architectures, such as Vision Transformers (Liu et al., 2024 [2]), have demonstrated superior global context modeling and attention-based feature extraction in medical imaging, enabling more robust detection of subtle lung abnormalities. Ethical considerations in AI diagnostics, including potential biases in demographic representation, imaging modalities, and data augmentation strategies, are increasingly emphasized to ensure fair, generalizable, and clinically safe deployment of AI systems (Shivaanivarsha & Kavipriya, 2023 [1]). Collectively, these studies underscore the role of CNNs, transformers, transfer learning, and advanced segmentation techniques in enhancing the accuracy and reliability of automated lung disease detection.

3. Implementation

Deep learning is a popular technique used in medical imaging analysis. For Analysis and detection of disease, it gives more accurate results. The overall steps followed in in Figure 1.

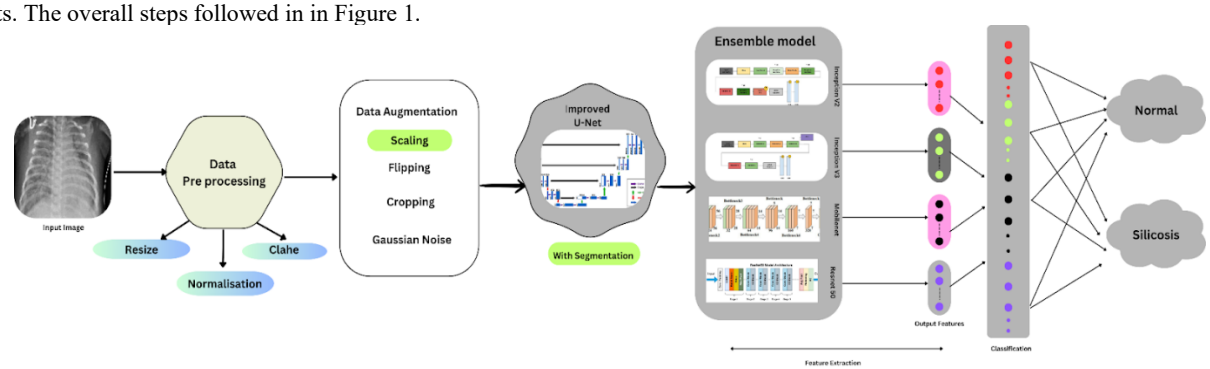


Fig. 1: Overall Flow Diagram of Proposed Model

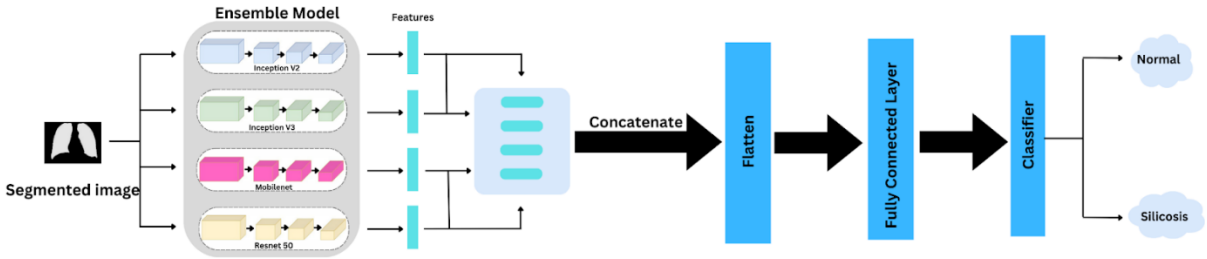


Fig. 2: Proposed Ensemble Model block diagram

The Ensemble Model Framework (Figure 2) for silicosis classification integrates segmentation from the Improved U-Net with four pre-trained DL models: MobileNet, ResNet 50, Inception V2, and Inception V3. Segmented X-ray or CT images serve as input, with each model extracting distinct silicosis-specific features. These extracted features are concatenated into a single vector, processed by a fully connected layer, and classified into Normal or Silicosis. The ensemble approach enhances accuracy by capturing both local and global patterns while U-Net ensures precise lung region focus. This scalable framework improves diagnostic reliability and adaptability across datasets and imaging modalities.silicosis development in the upper lobes of lung nodules is shown in the figure 3

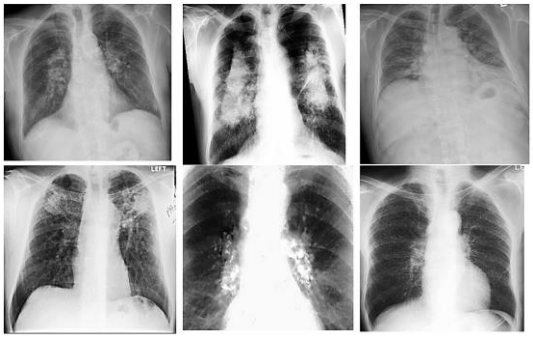


Fig. 3: Silicosis Chest X-ray images of complicated with progressive massive fibrosis (Source: Radiopaedia.org)

3.1 Methodology

This methodology emphasizes the two main techniques for detecting silicosis. The first method focuses on detecting and classifying silicosis with pre-trained models the other approach is based on an ensemble. Although the preprocessing and feature extraction processes vary, both methods identify and predict silicosis using DL models. The suggested system uses a hybrid model design, data augmentation, and sophisticated pre-processing to improve detection performance. The classification approach employs two techniques. The first is segmentation, which combines ensemble networks utilizing Inception V2, Inception V3, mobilenet, and ResNet-50 with the enhanced U-Net segmented areas. The final blocks are trained for improved task adaptability using these models as the foundation.

In the second segmentation method, the model extracts global features from the full medical image instead of segmenting individual areas and then feeds these data straight into a classification model. Chest X-ray images are taken for training purposes.

The dataset was collected from online repositories, comprising 1300 healthy lung images from Kaggle and 1328 silicosis-affected images from Radiopedia.org, with augmentation applied to balance classes. Images were resized to 256×256 pixels, normalized to [0,1], and split into 80% training and 20% testing sets. Pre-processing included noise reduction, contrast enhancement using CLAHE, and adaptive fil-

tering to preserve lung structures. Gaussian noise was added for model robustness. A novel improved U-Net was used for segmentation, integrating local and global sampling, skip connections, and residual convolutional blocks. The segmented lung areas were then classified to determine silicosis severity.

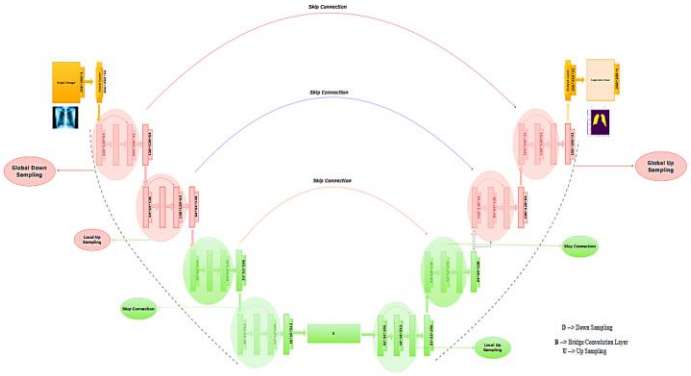


Fig. 4: Proposed Improved U-net diagram for segmentation

The proposed architecture (Fig. 4) enhances silicosis segmentation and classification by improving the conventional U-Net framework. It incorporates robust skip connections, local upsampling, and both global downsampling and upsampling pathways to preserve spatial details. The encoder systematically extracts hierarchical features using convolutional and pooling operations, capturing both local and global patterns. It consists of four sequential blocks with 3×3 convolutional layers and 2×2 max-pooling, reducing spatial dimensions while emphasizing deep semantic features. The bridge layer connects the encoder and decoder, ensuring a smooth transition while preserving high-level feature representations.

Table 1: Dimension transformations of Global Down / Up Sampling

Proposed Improved U-net model	Global Down Sampling	Global Upsampling
Transformations of Dimensions	Block 1 - $256 \times 256 \times 64 \rightarrow 128 \times 128 \times 64$ Block 2 - $128 \times 128 \times 128 \rightarrow 64 \times 64 \times 128$ Block 3 - $64 \times 64 \times 256 \rightarrow 32 \times 32 \times 256$ Block 4 - $32 \times 32 \times 512 \rightarrow 16 \times 16 \times 512$	Block 1 - $16 \times 16 \times 1024 \rightarrow 32 \times 32 \times 512$ Block 2 - $32 \times 32 \times 512 \rightarrow 64 \times 64 \times 256$ Block 3 - $64 \times 64 \times 256 \rightarrow 128 \times 128 \times 128$ Block 4 - $128 \times 128 \times 128 \rightarrow 256 \times 256 \times 64$
Transformations of Dimensions	Block 1 - $256 \times 256 \times 64 \rightarrow 128 \times 128 \times 64$ Block 2 - $128 \times 128 \times 128 \rightarrow 64 \times 64 \times 128$ Block 3 - $64 \times 64 \times 256 \rightarrow 32 \times 32 \times 256$ Block 4 - $32 \times 32 \times 512 \rightarrow 16 \times 16 \times 512$	Block 1 - $16 \times 16 \times 1024 \rightarrow 32 \times 32 \times 512$ Block 2 - $32 \times 32 \times 512 \rightarrow 64 \times 64 \times 256$ Block 3 - $64 \times 64 \times 256 \rightarrow 128 \times 128 \times 128$ Block 4 - $128 \times 128 \times 128 \rightarrow 256 \times 256 \times 64$
Transformations of Dimensions	Block 1 - $256 \times 256 \times 64 \rightarrow 128 \times 128 \times 64$ Block 2 - $128 \times 128 \times 128 \rightarrow 64 \times 64 \times 128$ Block 3 - $64 \times 64 \times 256 \rightarrow 32 \times 32 \times 256$ Block 4 - $32 \times 32 \times 512 \rightarrow 16 \times 16 \times 512$	Block 1 - $16 \times 16 \times 1024 \rightarrow 32 \times 32 \times 512$ Block 2 - $32 \times 32 \times 512 \rightarrow 64 \times 64 \times 256$ Block 3 - $64 \times 64 \times 256 \rightarrow 128 \times 128 \times 128$ Block 4 - $128 \times 128 \times 128 \rightarrow 256 \times 256 \times 64$
Transformations of Dimensions	Block 1 - $256 \times 256 \times 64 \rightarrow 128 \times 128 \times 64$ Block 2 - $128 \times 128 \times 128 \rightarrow 64 \times 64 \times 128$ Block 3 - $64 \times 64 \times 256 \rightarrow 32 \times 32 \times 256$ Block 4 - $32 \times 32 \times 512 \rightarrow 16 \times 16 \times 512$	Block 1 - $16 \times 16 \times 1024 \rightarrow 32 \times 32 \times 512$ Block 2 - $32 \times 32 \times 512 \rightarrow 64 \times 64 \times 256$ Block 3 - $64 \times 64 \times 256 \rightarrow 128 \times 128 \times 128$ Block 4 - $128 \times 128 \times 128 \rightarrow 256 \times 256 \times 64$
Transformations of Dimensions	Block 1 - $256 \times 256 \times 64 \rightarrow 128 \times 128 \times 64$ Block 2 - $128 \times 128 \times 128 \rightarrow 64 \times 64 \times 128$ Block 3 - $64 \times 64 \times 256 \rightarrow 32 \times 32 \times 256$ Block 4 - $32 \times 32 \times 512 \rightarrow 16 \times 16 \times 512$	Block 1 - $16 \times 16 \times 1024 \rightarrow 32 \times 32 \times 512$ Block 2 - $32 \times 32 \times 512 \rightarrow 64 \times 64 \times 256$ Block 3 - $64 \times 64 \times 256 \rightarrow 128 \times 128 \times 128$ Block 4 - $128 \times 128 \times 128 \rightarrow 256 \times 256 \times 64$

At the core of the architecture, the bridge layer functions as a bottleneck, seamlessly linking the encoder and decoder while focusing on extracting deep semantic features. This layer refines the input dimensions from 16×16×512 to 16×16×1024 through convolutional operations. Capturing global semantic information is crucial for accurate segmentation, particularly in complex medical imaging tasks.

To ensure that the reconstructed feature maps retain both high spatial resolution and contextual details from the encoder, the architecture integrates upsampling and skip connections. The decoder gradually increases spatial dimensions, reconstructing the segmentation map from the encoded features. By utilizing convolutional layers, skip connections, and upsampling, the model achieves precise and high-resolution segmentation.

The final layer of the model employs a 1×1 convolution to transform the enhanced feature mappings into the required number of output channels. This process generates a segmentation map with pixel-wise predictions. The resulting dimensions (256×256×C) depend on the number of classes in the segmentation task, where C=1 for binary segmentation.

The proposed framework utilizes an Enhanced U-Net for accurate lung segmentation, followed by a parallel ensemble of four pre-trained CNNs (Inception V2, Inception V3, MobileNet, and ResNet-50) for silicosis classification. The parallel ensemble was chosen over se-

quential strategies because it allows for common feature selection across architectures, reduces average prediction risk, and simplifies training and scaling. The liver segmentation areas provide accurate and noise-free input, and the fused ensemble features are passed through dense layers for final classification. This approach performed best, with an average validation accuracy of 98.48%, an F1 score of 96.93%, an MCC of 0.9275, and an AUC of 0.97, indicating robust and reliable silicosis detection.

3.2 Feature Extraction and Classification

A deep CNN processes pre-processed images, extracting and classifying features to identify silicosis-related patterns. Unlike segmentation-based methods, this approach analyzes entire images directly. The proposed ensemble model integrates four CNN architectures—Inception V2, Inception V3, MobileNet, and ResNet50—using a heterogeneous (parallel) ensemble approach for classification. Cross-validation ensures robustness and prevents overfitting. Features are extracted, concatenated, and passed through a dense layer. The method employs Stacking, where base models (level-0) generate predictions, and a meta-model (level-1) optimally combines them. This parallel transfer learning strategy enhances classification accuracy and future predictive performance. The final classification ensemble processes the lung region segmented by the improved U-Net model, ensuring a refined and accurate classification of silicosis.

$$\mathrm{FE} = [\mathrm{FE}_{\mathrm{InceptionV2}}, \mathrm{FE}_{\mathrm{InceptionV3}}, \mathrm{FE}_{\mathrm{Mobilenet}}, \mathrm{FE}_{\mathrm{ResNet50}}] \quad (1)$$

where [.] denotes the vector concatenation, and 'FE' is the ensemble function.

$$\mathrm{FE} = \frac{\mathrm{FE}_{\mathrm{InceptionV2}}, \mathrm{FE}_{\mathrm{InceptionV3}}, \mathrm{FE}_{\mathrm{Mobilenet}}, \mathrm{FE}_{\mathrm{ResNet50}}}{4} \quad (2)$$

The above equations [2-3] are used to determine the average for each element.
The ensemble model was optimized using appropriate loss functions and optimization techniques, leveraging the strengths of four CNN architectures to enhance accuracy and robustness in silicosis classification. Fully connected layers were customized for improved classification. Hyperparameter tuning was performed, adjusting epochs, optimizer, batch size, loss function, and activation functions. Categorical cross-entropy was used as the loss function, with SGD as the optimizer. Feature maps from convolutional layers were analyzed to visualize the model's processing. The model optimized hyperparameters iteratively for better performance, using sigmoid activation and Adam optimizer to enhance training efficiency.

3.3 Performance evaluation

A confusion matrix, a useful technique for evaluating classification performance, was used to analyze the model's performance. Key performance indicators, including accuracy, precision, recall, F1 score, and area under the curve (AUC), were extracted from the confusion matrix using the following formulas. When assessing an ensemble model, there are standard metrics.
One of the most widely used and straightforward metrics is accuracy, which is defined in Equation 4.

$$\mathrm{Accuracy} = \frac{\mathrm{T_{Pos}} + \mathrm{T_{Neg}}}{\mathrm{T_{Pos}} + \mathrm{T_{Neg}} + \mathrm{F_{Pos}} + \mathrm{F_{Neg}}} \quad (3)$$

Equation 5 defines recall, sometimes referred to as sensitivity, as the ensemble model's capacity to detect positive samples.

$$\mathrm{Recall} = \frac{\mathrm{T_{Pos}}}{\mathrm{T_{Neg}} + \mathrm{F_{Pos}}} \quad (4)$$

where positive indicates how many observations are positive, and genuine positive indicates how many observations are truly positive.
Precision is another well-known effectiveness parameter. It measures the proportion of cases that are identified as positive. The precision equation is defined formally as 6.

$$\mathrm{Precision} = \frac{\mathrm{T_{Pos}}}{\mathrm{T_{Pos}} + \mathrm{F_{Pos}}} \quad (5)$$

Similarly, specificity quantifies the model's ability to detect negative samples. The definition of the equation is 7.

$$\mathrm{Specificity} = \frac{\mathrm{T_{Neg}}}{\mathrm{T_{Neg}} + \mathrm{F_{Pos}}} \quad (6)$$

where negative indicates how many negative observations there are, and true negative (TN) indicates how many real negative observations there are.
Recall metrics and precision are frequently traded off. The second measure frequently declines when an attempt is made to improve the first. To quantify this trade-off, F-Measure computes the harmonic mean of precision and recall. which is defined in equation 8.

$$\mathrm{F1-Score} = \frac{2 \times (\mathrm{Precision} \times \mathrm{Recall})}{\mathrm{Precision} + \mathrm{Recall}} \quad (7)$$

Intersection over union (IoU) and the dice coefficient were employed as performance indicators to evaluate the improved U-Net model's effectiveness. which id denoted in equation 9.

$$\mathrm{Dice\ Coeficient} = \frac{2 .|\ \mathrm{Pr} \cap \mathrm{Gt}|}{|\mathrm{Pr}| + |\mathrm{Gt}|} \quad (8)$$

Where Pr and Gt are the predictive and ground truth.

$$TPR = T_{Pos} / (T_{Pos} + F_{Neg}) \quad (9)$$

$$FPR = F_{Pos} / (F_{Neg} + T_{Neg}) \quad (10)$$

$$AUC = \frac{1}{2} \left(\frac{T_{Pos}}{T_{Pos} + F_{Neg}} + \frac{T_{Neg}}{T_{Neg} + F_{Pos}}\right) \quad (11)$$

where TPR and FPR are the true positive rate and false positive rate, respectively, denoted in equations [10-12]. AUC stands for the area under a curve defined in equation 14.
By calculating the percentage of accurately detected silicosis cases among all instances that were anticipated to be abnormal, NPV provides insight into false negatives (FNs) for silicosis, given in equation 13.

$$NPV = (T_{Neg}) / (T_{Neg} + F_{Neg}) \quad (12)$$

4. Results and Discussion

The efficiency of the proposed model and the pre-trained model comparison is presented in this section. In general, segmentation of medical images improves accuracy in medical diagnosis. The proposed model was created particularly to solve the difficulties in classifying and segmenting silicosis. The benefits of the proposed models are downsampling and up-sampling hierarchically, improved feature extraction on several levels, and improved segmentation and classification by aiding in the recovery of lost spatial information. The model's comprehension of global context is ensured by rich feature learning at the bottleneck. The model is well-suited to a variety of medical imaging datasets. Figure 5 shows the sample segmented mask and image of the collected dataset. Different lung regions and possible silicosis-related abnormalities, like fibrotic scarring or nodular lesions, are highlighted by the images' varied intensity and structure.

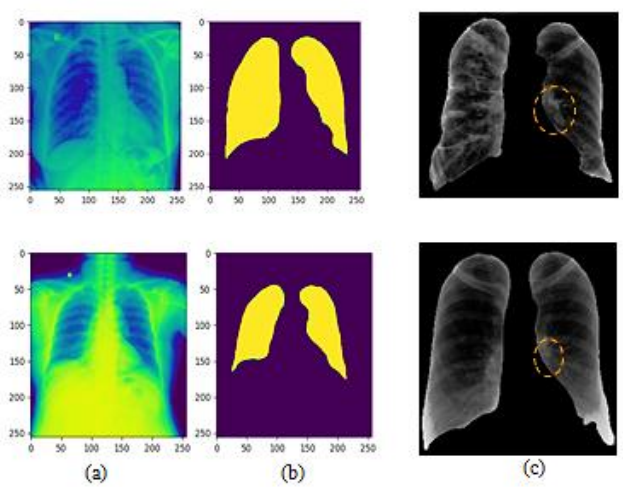


Fig. 5: Improved Unet Results visualization of the dataset (a) Original image, (b) Segmentation mask, (c) Segmented images.

The model can learn spatial and structural information since the segmentation mask clearly distinguishes the sick from healthy lung regions. The large overlap between the reference mask and the projected segmentation shows that the model successfully reproduces the ground truth segmentation. The segmented regions show that the model can accurately localize silicosis by following the contours of areas affected by the disease. During the training and validation stages, the Improved U-Net model was assessed using various evaluation metrics, as shown in Table 1. The ability of the Improved U-Net model to precisely segregate areas impacted by silicosis is seen visually in Figure 5. To quantify the progression of disease, plan treatments, and support automated diagnosis systems, precise segmentation is essential in silicosis detection. As this image illustrates, accurately identifying sick regions lowers the possibility of incorrect diagnosis and aids radiologists in making decisions. The high Dice coefficient (0.97+) and accuracy (97%) attained during model validation further demonstrate its efficacy in the automated categorization and detection of silicosis. The model is established as a reliable tool for medical imaging in silicosis investigation by the combination of qualitative and quantitative results.

Table 2: Evaluation metrics of Improved U-net segmentation method

Evaluation Parameters	Improved Unet Training	Improved Unet Validation
Loss	0.698	0.291
Dice Coef	0.967	0.9701
Accuracy	98.26	98.33
Learning Rate	Ir: 16525e-05	
Optimizer & Epochs	Adam & 40	



Fig. 6: A comparative analysis of the training and validation performance of the proposed segmentation model

Figure 6 and Table 2 provide a comparative analysis of the training and validation evaluation metrics for the proposed segmentation model. In the context of silicosis detection in chest radiographs, these metrics offer insights into the model's performance, consistency, and ability to generalize. The loss metric measures how closely the model's predictions align with the actual labels, with lower loss values indicating better model alignment and performance.

The model's ability to reduce errors during the learning process is demonstrated by the training loss (TL). The model has successfully identified patterns in the training dataset, as indicated by a value of 0.698. One popular similarity statistic for image segmentation tasks is the Dice Coefficient. Values nearer 1 indicate higher segmentation accuracy; it spans from 0 to 1.0.967 is the training dice coefficient. The model's accuracy in identifying silicosis-affected areas in training photos is demonstrated by this high value. A marginally greater level of validation. Excellent generalization and consistency are indicated by the dice coefficient (0.9701), which guarantees that the model's segmentation capacity extends to invisible radiographs. The assessment measures show how well the Improved U-Net model detects silicosis. The model is ideal for medical applications due to its high Dice coefficients, low loss values, and exceptional accuracy. The model's capacity to successfully learn features from the training data can be seen by its 96.26% training accuracy (TA), and its robustness can be seen by its 97.18% VA, which shows that the model performs well on unseen validation data. The evaluation measures show that the Improved U-Net model is effective in detecting silicosis. The model's high Dice coefficients, low loss values, and exceptional precision make it a good fit for medical applications. The model is ideal for real-world applications because of its low validation loss (VL) and excellent VA, which demonstrate its dependability on unseen data.

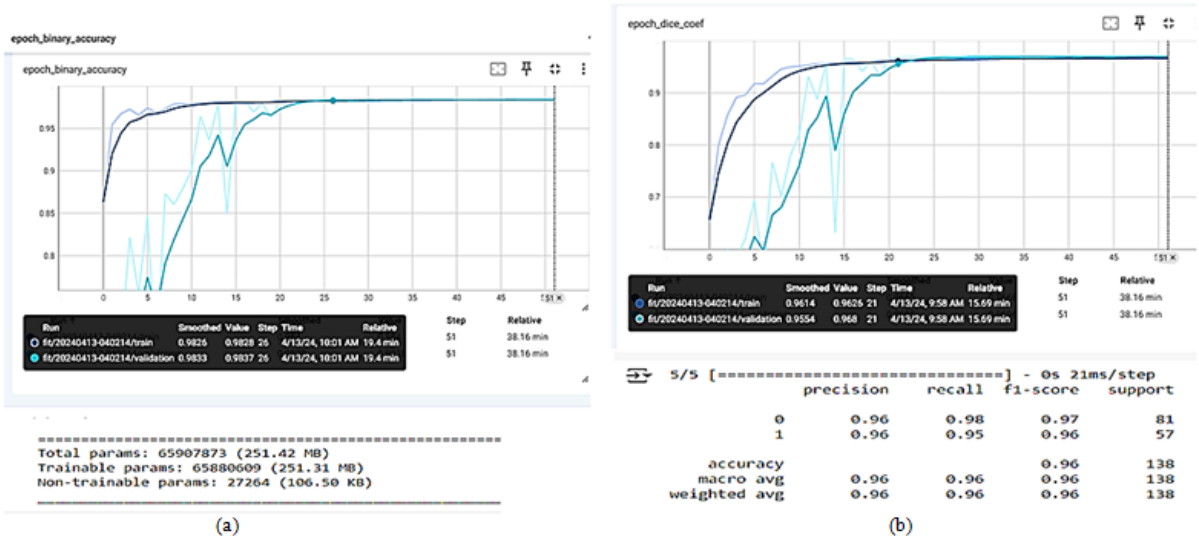
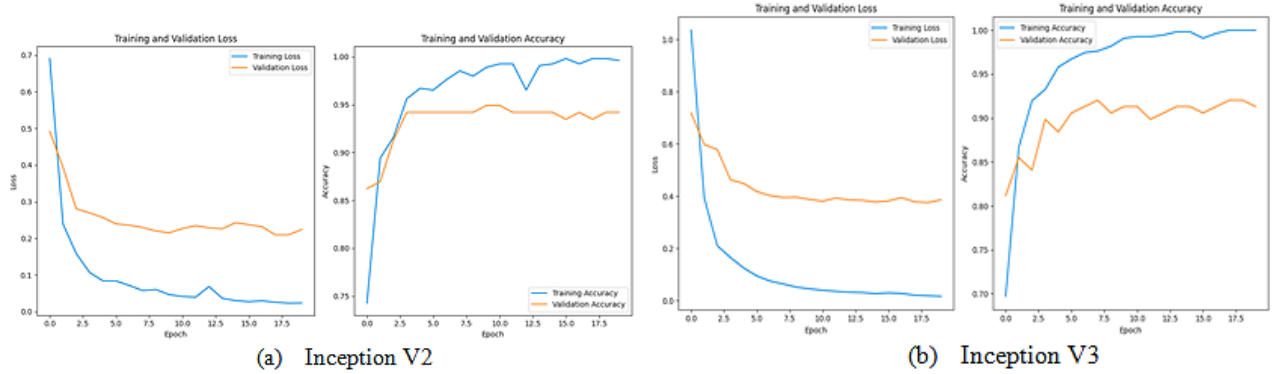


Fig. 7: (a) Accuracy Convergence (b) Dice Coefficient Convergence Curve of the Improved U-Net Model

The accuracy and dice coefficient convergence curves for the Improved U-Net model's training and validation stages are shown in Figure 7. The complexity of the improved U-Net model is indicated by the total parameters, which are 65,907,873 (251.42 MB). Its capacity to identify complex patterns in silicosis detection is shown by the huge number of parameters. The majority of the model parameters are optimized during training, which is essential for learning, as seen by the trainable parameters of 65,880,609 (251.31 MB). The binary accuracy trend over epochs for both training (solid black curve) and validation (light blue curve) is shown by the curve on the left. The graph shows that the model learns quickly in the early epochs, as evidenced by a significant rise in accuracy, as well as stabilization. The model has converged to an ideal solution when training and VA plateau after about 20 epochs. A TA of 0.9862 and a VA of 0.9833 were attained. These high numbers demonstrate how resilient the model is and how well it generalizes to unknown validation data. The Dice coefficients for training (solid black curve) and validation (light blue curve) throughout epochs are displayed in the curve on the right. We may infer from the graph that, like accuracy, the Dice coefficient rises substantially in the early epochs, demonstrating the model's capacity to precisely segment regions afflicted by silicosis. The Dice coefficient stabilizes after about 25 epochs, guaranteeing accurate dataset segmentation. The model's outstanding segmentation performance is confirmed by the validation dice coefficient of 0.9701 and the training dice coefficient of 0.967. The convergence curves show that the Improved U-Net model achieves optimal performance with little overfitting by efficiently learning the underlying patterns for both classification and segmentation tasks. The model's generalization capacity is confirmed by its consistent performance across training and validation datasets. Its excellent accuracy values and Dice coefficient make it perfect for clinical applications in silicosis identification.



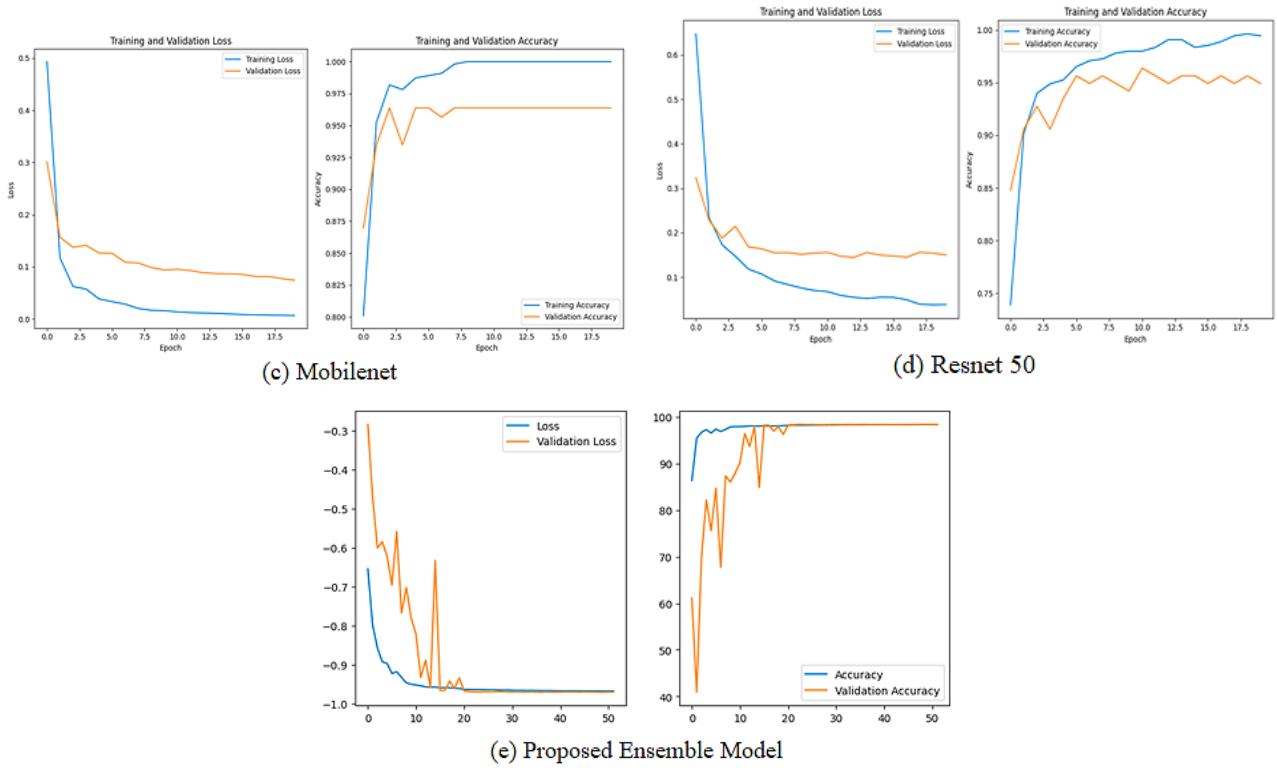


Fig. 8: Training and VA curve of all the models

Table 3: Training and VA /loss Comparison of all the models

Model	Average VA	Average VL	Average TA	TL
Inception V2	94.20	0.2240	0.99	0.0237
Inception V3	91.30	0.3853	0.99	0.0162
Mobilenet	96.38	0.0744	0.99	0.0068
Resnet 50	94.93	0.1503	0.99	0.0386
Proposed Model (5-fold) (Modified Unet+Ensemble)	98.48	-0.9696	98.40	-0.966

Table 3 provides a detailed comparison of various DL models for silicosis detection in chest radiographs, with corresponding accuracy, loss, and confusion matrices shown in Figure 8. The model's average VA, determined during the five-fold cross-validation phase, measures its performance on unseen data. A lower VL indicates better performance, reflecting the model's ability to predict expected outputs for validation data. The average TA represents the accuracy on the training dataset, while the TL indicates the training error, with lower values suggesting better convergence.

The proposed model, which integrates a modified U-Net architecture with an ensemble strategy, outperforms individual models (Inception V2, Inception V3, MobileNet, and ResNet50) in terms of performance. With an impressive average VA of 98.48%, the proposed model demonstrates superior generalization to new data. It also shows exceptional prediction confidence and convergence with a VL of -0.9696 (negative loss indicates logarithmic or normalized corrections). Training measures further validate its resilience, with 98.40% accuracy and a loss of -0.966.

In comparison, Inception V2 provides a reliable but less efficient option, achieving 94.20% VA and a VL of 0.2240. Inception V3, however, performs the weakest among the models, with the lowest VA (91.30%) and the highest VL (0.3853), making it less suitable for silicosis detection in this study. MobileNet stands out with the best performance, achieving 96.38% VA and the lowest VL (0.0744), benefiting from its lightweight design. ResNet50 also performs well, especially in feature extraction for complex medical imaging tasks, with 94.93% VA and a VL of 0.1503.

All models, except the proposed one, achieve 99% TA. The slight decrease in the proposed model's TA (98.40%) is a result of the ensemble's focus on reducing overfitting by balancing training and validation performance. The enhanced performance of the proposed model underscores the importance of combining ensemble learning, which leverages the strengths of multiple models, with a modified U-Net tailored for segmentation. The high VA of the proposed model ensures reliable silicosis identification, critical for timely clinical diagnosis.

Table 4: Evaluation Metrics of Silicosis lung classification by all the models

Performance Metrics	Inception V2	Inception V3	Mobilenet	Resnet 50	Proposed Model (Modified Unet+Ensemble)
Precision	95.06	91.56	94.61	94.25	96.34
Recall	95.06	93.82	94.66	95.06	97.53
F1 score	95.06	92.68	95.59	95.65	96.93
Specificity	0.9288	87.71	93.41	93.74	94.73
NPV	0.9298	0.9090	0.9093	0.9310	0.9642

A detailed examination of several models for the classification of silicosis lung is given in Table 4. The proposed model (Modified U-Net + Ensemble) demonstrates superior performance in detecting silicosis in chest radiography, achieving the highest precision (96.34%), recall (97.53%), F1-score (96.93%), specificity (94.73%), and NPV (96.42%). Its ability to minimize false positives and false negatives

ensures reliable classification. While Inception V2 performs consistently well with a precision, recall, and F1-score of 95.06%, its specificity (92.88%) and NPV (92.98%) are lower than those of MobileNet and the proposed model. Inception V3 shows the weakest performance, with a precision of 91.56%, an F1-score of 92.68%, and a specificity of 87.71%, indicating its limitations in classifying silicosis. MobileNet exhibits balanced performance but has a slightly lower recall (94.66%) compared to the proposed model. ResNet-50 provides well-rounded results but falls short of the proposed ensemble approach. The high recall and NPV of the proposed model make it a highly effective tool for early silicosis detection, reducing the risk of false negatives. By integrating an ensemble technique with a customized U-Net, the model enhances diagnostic accuracy, minimizes human error, and accelerates the detection process, making it a valuable asset for radiologists.

Table 5: Evaluation Metrics of Normal lung classification by all the models

Performance Metrics	Inception V2	Inception V3	Mobilenet	Resnet 50	Proposed Model (Modified Unet+Ensemble)
Precision	92.98	90.90	93.93	93.13	96.42
Recall	92.98	87.71	93.41	93.73	94.73
F1 score	92.98	89.28	92.12	93.91	95.57
Specificity	95.06	93.82	94.66	95.06	97.53
NPV	0.9504	0.9156	0.9261	0.9325	0.9634

The proposed model (Modified U-Net + Ensemble) outperforms individual CNN architectures in classifying normal lungs with high precision (96.42%), recall (94.73%), F1-score (95.57%), specificity (97.53%), and NPV (96.34%). Compared to Inception V2, Inception V3, MobileNet, and ResNet-50, the ensemble model consistently achieves better accuracy, reducing misclassifications. By integrating multiple feature extraction techniques through ensemble learning, the model ensures a robust balance between identifying normal lungs and accurately diagnosing silicosis, as illustrated in Figure 9.

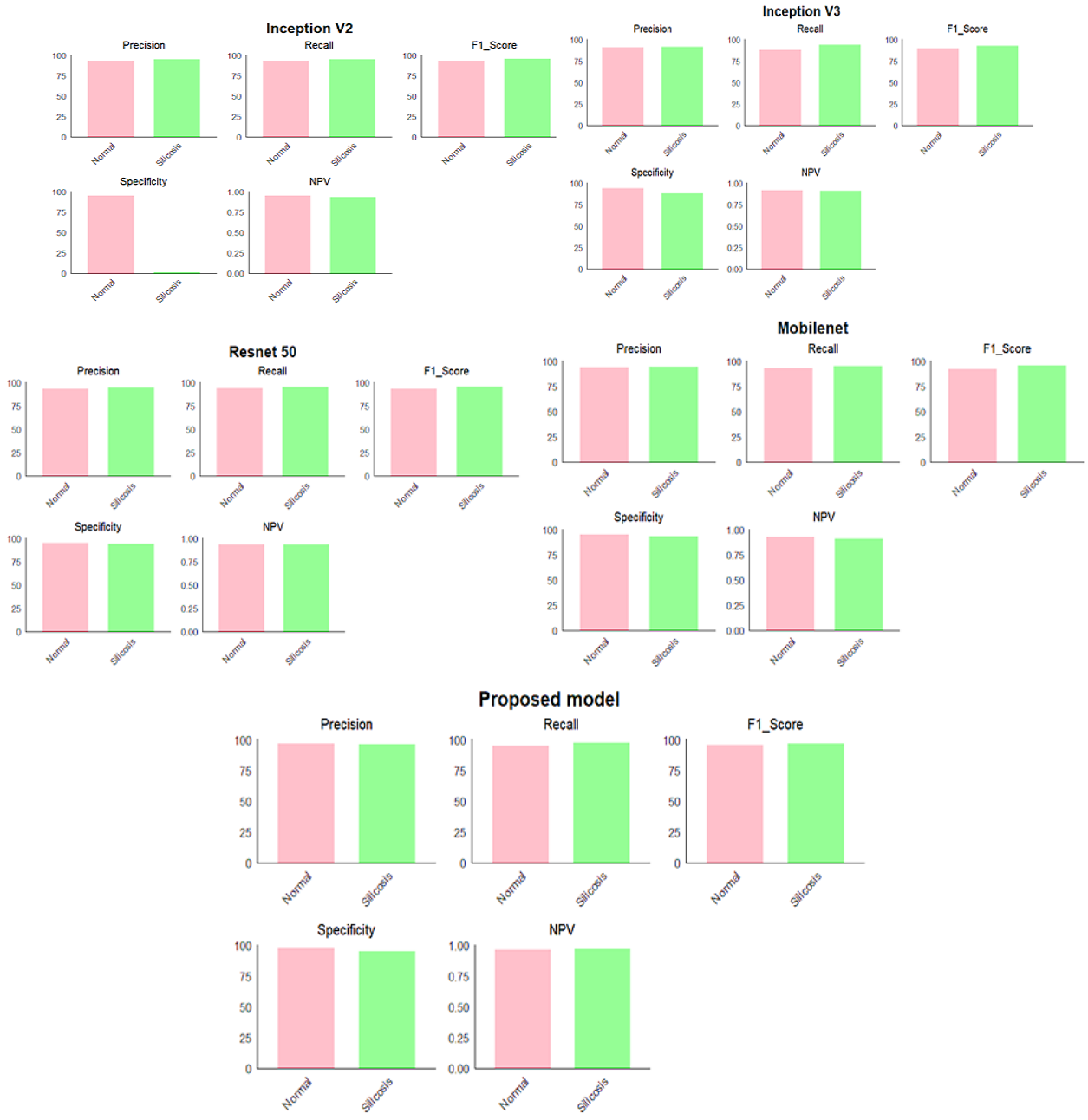


Fig. 9: A comparative analysis of the evaluation metrics for all the models

Table 6: Mathew Correlation coefficient comparison of all models

Performance Metrics	Inception V2	Inception V3	Mobilenet	Resnet 50	Proposed Model (Modified Unet+Ensemble)
Mathew Correlation Coefficient	0.8840	0.8260	0.9113	0.8910	**0.9275**

The proposed model (Modified U-Net + Ensemble) consistently outperforms individual CNN architectures, achieving the highest precision of 96.42%, recall of 94.73%, F1-score of 95.57%, specificity of 97.53%, and NPV of 96.34%. Compared to Inception V2, Inception V3, MobileNet, and ResNet-50, the ensemble approach demonstrates improved classification accuracy by reducing misclassifications and enhancing reliability. Additionally, the highest Matthews Correlation Coefficient (MCC) of 0.9275 confirms the model's superior ability to balance true and false predictions. Inception V2 and ResNet-50 exhibit stable performance with MCC scores of 0.8840 and 0.8910, respectively, while MobileNet surpasses them with 0.9113. Inception V3, however, shows the weakest performance with an MCC of 0.8260, indicating a higher rate of misclassification. The ensemble model effectively integrates feature extraction from multiple architectures, leveraging their complementary strengths to ensure robust silicosis detection with minimal false positives and false negatives. The suggested model's excellent performance in silicosis classification is validated by its high MCC score (0.9275) ensures balanced predictions, reduces overfitting problems, and improves the overall accuracy of the model. Which is given in Table 6 and depicted in Figure 10.

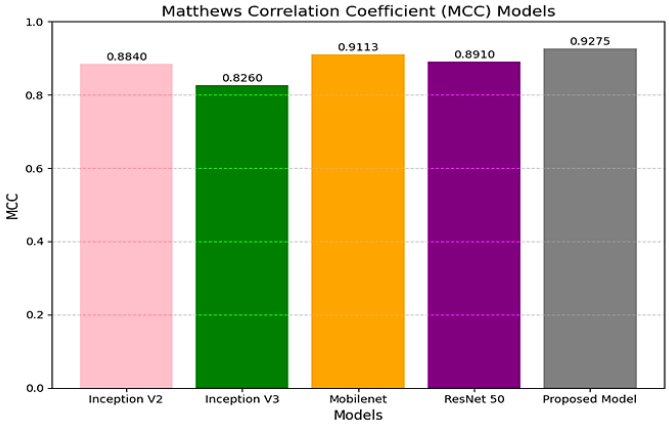


Fig. 10: Comparative analysis of Mathew Correlation coefficient of all the models

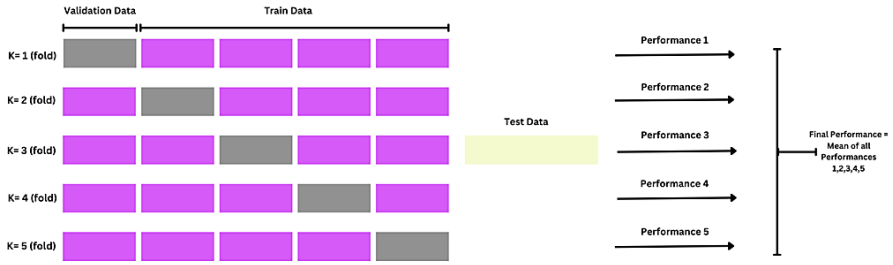
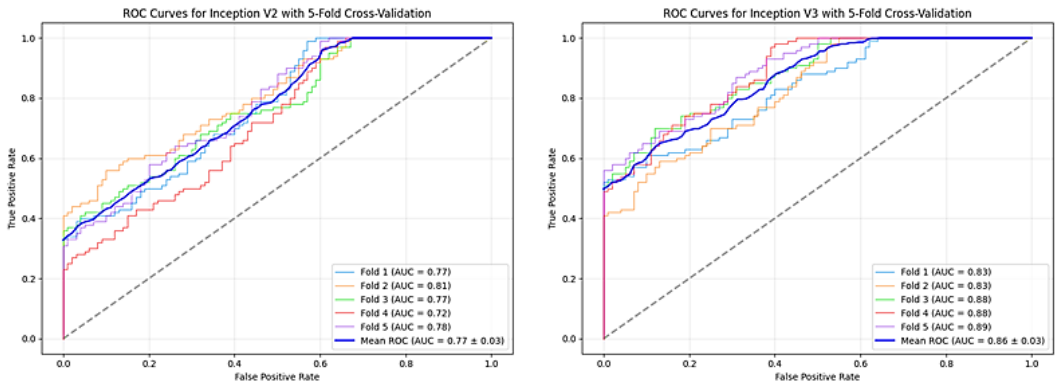


Fig. 11: Five-fold cross-validation

To improve the robustness of the chosen features, prevent overfitting, and ensure reliable model evaluation, we employed five-fold cross-validation, as shown in Figure 11. Cross-validation is especially effective for refining models on smaller datasets with pre-trained networks. During each iteration, one fold serves as the validation set, while the remaining K−1 folds are used for training. The model is trained on the K−1 folds and validated using the excluded fold.



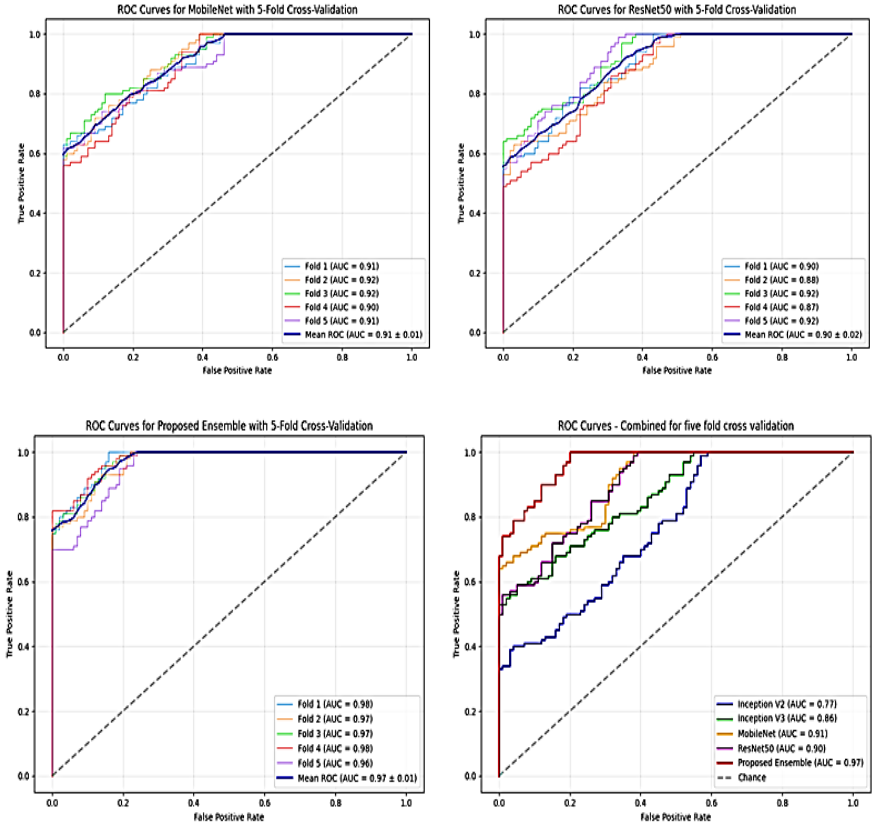


Fig. 12: ROC curve of all the models for five five-fold cross-validation

One important technique for assessing the effectiveness of classification models is the Receiver Operating Characteristic (ROC) curve. At different threshold values, it graphically depicts the trade-off between sensitivity (TPR) and specificity (FPR). In medical applications where precise classification is essential, such as silicosis detection, this graphic is especially useful. To guarantee a reliable assessment of multiple DL models and an ensemble model for silicosis classification in chest radiographs, we use five-fold cross-validation in this study. As seen in the above fig 12, a combined performance metric is obtained by averaging the ROC curves and associated AUC values for each model across all five folds. Inception V2's AUC is 0.77, Inception V3's is 0.86, MobileNet's is 0.91, Resnet50's is 0.90, and the suggested ensemble's is 0.97. With the greatest AUC (0.97), the proposed ensemble model performs better than any single model and demonstrates nearly ideal classification capabilities. With AUC values of 0.91 and 0.90, respectively, MobileNet and ResNet50 perform competitively when compared to other models. With an AUC of 0.86, Inception V3 exhibits a respectable level of performance, but Inception V2 trails slightly with an AUC of 0.77.AUC = 0.50 for random guessing is represented by the dashed diagonal line. Every model performs noticeably better than the chance line, demonstrating their capacity for insightful forecasting. By utilizing their various feature representations and minimizing overfitting, the ensemble approach combines the advantages of individual models. High AUC values are obtained by MobileNet and ResNet50, demonstrating their resilience in obtaining significant features from chest radiographs. The performance of these architectures is influenced by their deeper layers and lightweight design, respectively. While Inception V2 and V3 show respectable performance, their marginally lower AUCs indicate that they could be improved, either with more training data. Five-fold cross-validation is used to guarantee the evaluation's objectivity and that performance indicators accurately represent the models' capacity to generalize to new data. The response curve and values can be seen in Figure 12 and Table 7.

Table 7: Comparative AUC performance analysis of all the models (5-fold cross-validation)

K-Fold	Inception V2	Inception V3	Mobilenet	Resnet 50	Proposed Ensemble Model
	AUC	AUC	AUC	AUC	AUC
1- fold	0.77	0.83	0.91	0.90	0.98
2- fold	0.81	0.83	0.92	0.88	0.97
3- fold	0.77	0.88	0.92	0.92	0.97
4- fold	0.72	0.88	0.90	0.87	0.98
5- fold	0.78	0.89	0.91	0.92	0.96
Mean	0.77 ± 0.03	0.86 ± 0.03	0.91 ± 0.01	0.90 ± 0.02	**0.97 ± 0.01**

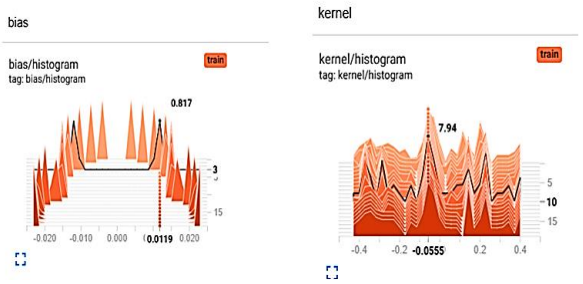


Fig. 13: Bias and Kernel Histogram

The TensorBoard histogram visualization in Figure 13 depicts the distribution of bias and kernel weights during training. The Y-axis of the bias histogram represents the frequency of bias values, while the X-axis indicates bias values, typically ranging between -0.02 and 0.02, centered around zero. The stable, symmetric distribution suggests controlled biases, preventing instability. Similarly, the kernel histogram shows weight distribution across training, with values ranging from -0.4 to 0.4. The centered and uniform distribution indicates balanced weight updates, avoiding vanishing or exploding gradients. These patterns confirm a well-behaved training process without signs of overfitting or underfitting. Research suggests that optimized silicosis detection models achieve over 90% sensitivity and specificity, influenced by architecture, preprocessing, and data quality. The proposed model maintained stability and demonstrated high accuracy. Figure 14 presents a prediction example where the model correctly classified a normal lung X-ray with an output of 0.000631, rounded to 0, confirming accurate detection.

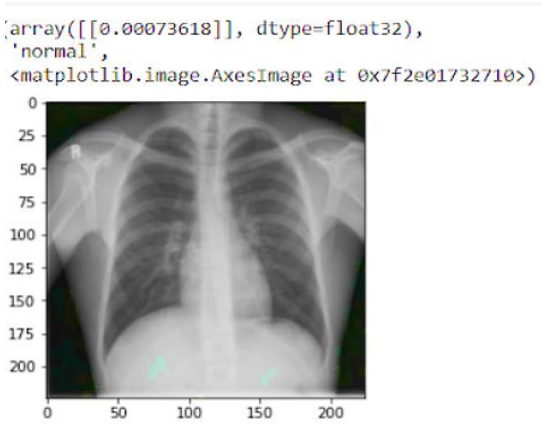


Fig. 14: Prediction of a Normal Lung Image



Fig. 15: Predicting Silicosis lung image

Figure 15 displays the prediction for a silicosis lung X-ray image, where the model predicted a value of 0.9104177, which was rounded to 1. Class 1 represents a Silicosis Lung Image, and the prediction was accurate.

The superiority of the ensemble model over individual architectures lies in its ability to exploit the common strengths of the basic network. The lightweight design of MobileNet allows for efficient feature extraction from low-resolution maps, while the deep architecture of ResNet50 captures complex hierarchical features, and Inception models provide spatial information at different scales. By combining these different faces, the ensemble achieves a more robust and accurate classification. Furthermore, potential dataset biases—such as demographic variations or differences in imaging modalities—were mitigated through data augmentation, including rotation, scaling, and noise injection, thereby improving the model's generalization to unseen data. Therefore, the ensemble not only reduces the risk of overfitting but also ensures consistent performance on different chest radiographic datasets.

5. Conclusion

In this research, we used an ensemble of pre-trained DL models in conjunction with an Enhanced U-Net model to propose an advanced framework for the classification and detection of silicosis. The method achieves state-of-the-art performance by combining robust feature extraction and classification with precise lung region segmentation. By successfully isolating lung areas, the Enhanced U-Net lowers noise and guarantees accurate input for classification. By leveraging Inception V2, Inception V3, MobileNet, and ResNet 50, the ensemble model provides a comprehensive feature representation that captures both localized and global silicosis patterns. Superior outcomes were shown by the suggested framework in terms of F1 score, accuracy, precision, recall, and Matthews Correlation Coefficient (MCC). Using 5-fold cross-validation, the system notably obtained an average VA of 98.48%, demonstrating its generalisability and resilience. Furthermore, combining the segmentation and classification phases improves the data's interpretability and qualifies them for practical clinical uses. R-Studio was used for analysis and visualization in the implementation of the suggested framework. An NVIDIA-powered i7 processor was used for both the training and validation procedures, guaranteeing effective computing and model optimization.

References

[1] Shivaanivarsha, N., & Kavipriya, P. (2023). Identification of silicosis using deep learning model on computed tomography images. In *IEEE 3rd Mysore Sub Section International Conference (MysuruCon)*, Mysuru, India.

[2] Liu, Y., Yu, L., Wu, Q., & Zhang, X. (2024). Identification of high-risk population of pneumoconiosis using deep learning segmentation of lung 3D images and radiomics texture analysis. *Computer Methods and Programs in Biomedicine, 244*, 108006.

[3] Li, X., Zheng, Y., Li, Y., & Zhang, Z. (2024). Deep convolutional network-based chest radiographs screening model for pneumoconiosis. *Frontiers in Medicine, 11*, 1290729.

[4] Devnath, L., Kumar, R., Singh, P., & Pandey, A. (2022). Detection and visualisation of pneumoconiosis using an ensemble of multi-dimensional deep features learned from chest X-rays. *International Journal of Environmental Research and Public Health, 19*(18), 11193.

[5] Devnath, L., Kumar, R., Singh, P., & Pandey, A. (2022). Deep ensemble learning for the automatic detection of pneumoconiosis in coal worker's chest X-ray radiography. *Journal of Clinical Medicine, 11*(18), 5342.

[6] Wang, D., Zhang, Y., Liu, J., & Zhang, J. (2020). Automated pneumoconiosis detection on chest X-rays using cascaded learning with real and synthetic radiographs. In *Proceedings of the Digital Image Computing: Techniques and Applications (DICTA)*.

[7] Xuan, W., Han, Z., & Zheng, L. (2021). Non-invasive detection of silicosis based on array sensing and pattern recognition. *E3S Web of Conferences, 271*, 1–6.

[8] Yang, F., Zhi, M., Chen, X., & Li, Y. (2021). Pneumoconiosis computer aided diagnosis system based on X-rays and deep learning. *BMC Medical Imaging, 21*, 1–7.

[9] Gao, Q., Wang, S., Zhao, D., & Liu, J. (2007). Accurate lung segmentation for X-ray CT images. In *Proceedings of the Third International Conference on Natural Computation*, Haikou, China, 275–279.

[10] Chanda, P. B., & Sarkar, S. K. (2020). Effective and reliable lung segmentation of chest images with medical image processing and machine learning approaches. In *IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI)*, Buldhana, India.

[11] Saad, M. N., Mohsin, M. F., Hamid, H. A., & Muda, Z. (2021). Lung area segmentation on chest X-ray based on texture image feature. In *Proceedings of the 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, Penang, Malaysia, 92–96.

[12] Zhang, L., Rong, R., Li, Q., & Wang, X. (2021). A deep learning-based model for screening and staging pneumoconiosis. *Scientific Reports, 11*(2201), 1–7.

[13] Shivaanivarsha, N., & Kavipriya, P. (2023). A robust approach for the detection of diabetic retinopathy at an early stage using deep CNN. In *Proceedings of the International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence*, Chennai, India.

[14] Ramesh, A., Bhatkhande, A., Avduth, N., Jadhav, B., & Rathi, K. (2022). Automated lung segmentation on chest CT using deep learning. In *Proceedings of the 3rd International Conference for Emerging Technology (INCET)*, Belgaum, India.

[15] Shivaanivarsha, N., Kavipriya, P., & Shyamkumar, M. (2023). An efficient fully automated detection of mucormycosis using three-dimensional deep learning on computer tomography studies. In *Proceedings of the International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*, Chennai, India.