A Silicosis-Focused Hybrid Architecture Leveraging Enhanced Segmentation and Feature Based Classification

Shivaanivarsha Natarajan
Department of Electronics and Communication Engineering
Sathyabama Institute of Science and Technology
India
varsha.ece@sairam.edu.in

Kavipriya Periyasamy
Department of Electronics and Communication Engineering
Sathyabama Institute of Science and Technology
India
kavipriya.ece@sathyabama.ac.in

Abstract: *Long-term exposure to crystalline silica dust causes silicosis, an irreversible occupational lung disease that is currently a major global health concern because of its delayed diagnosis and few available treatments. In this work, a new segmentation-driven hybrid framework for automated silicosis staging and detection from chest radiographs, called SilicoNet (Silicosis+Network), is proposed. Two separate experiments were conducted to validate the framework. In the first experiment, SilicoNet was used to segment lungs from pre-processed and augmented chest radiographs, and its performance was compared to that of the conventional U-Net model. In the second experiment, the outputs of the proposed custom Convolutional Neural Network (CNN) model were systematically compared with those of three popular CNN architectures used for classification: MobileNetV2, InceptionV2, and ResNet50. For validation, evaluation criteria such as the Dice similarity coefficient, Jaccard index, precision, recall, f1-score, accuracy, specificity, matthews correlation coefficient, negative predictive value, and training length were used. The findings show that SilicoNet and the customised CNN perform better than standard baselines, achieving a maximum accuracy of 96.40%. This study is distinctive because it combines improved segmentation and optimised classification, which results in better robustness and generalisability than previous models.*

Keywords: *Deep learning, lung disease, silicosis, segmentation, UNet.*

Received June 3, 2025; accepted October 1, 2025
https://doi.org/10.34028/iajit/23/2/09

1. Introduction

Inhaling crystalline silica dust for an extended period is the main cause of silicosis, a progressive and irreversible occupational lung disease. It results from the lung tissue's reaction to silica dust inhalation. An issue that arises in the lungs that affects their normal function is called a lung disorder. Silica particles act like microscopic knives on the lungs. They cause little cuts and long-term damage that leaves scars or lung fibrosis. A total absence of silica exposure does not stop the development. It continues to pose a serious threat to human health, particularly in poorer nations where industrial safety regulations are frequently insufficient. The radiographic appearance of many nodules, primarily in the superior and posterior segments of the upper lobes, with a diameter of 1-10mm, is characteristic of simple silicosis. The radiographic appearance of extensive opacities with areas of homogeneous consolidation, which mostly affect the superior and middle segments of the lungs, is a characteristic of complicated silicosis, also known as Progressive Massive Fibrosis (PMF). Manually diagnosing silicosis from chest radiographs is still a difficult and subjective task, even with the increasing advances in medical imaging. Inter-observer variability, poor image contrast, and overlapping anatomical structures are common challenges for radiologists. To overcome these constraints, this research suggests a hybrid deep learning architecture that blends the advantages of feature-based classification and improved segmentation. A modified U-Net architecture with global and local sampling paths, attention mechanisms, and residual connections is used in the segmentation module to extract localised pathological features. By making Region-Of-Interest (ROI) identification more accurate, this enhances the downstream classification work. This article aims to save the lives of those who suffer from Silicosis by providing early and precise diagnosis. The main contributions of this research study are:

- The development of a novel UNet architecture that combines global and local sampling modules enables accurate lung region segmentation from chest X-ray images. The localisation of silicosis-affected areas is much enhanced by this segmentation phase.
- To extract multi-level deep features, the segmented lung regions are processed using an InceptionV2, MobilenetV2, ResNet50 models, and the proposed Convolutional Neural Network (CNN) model. This method makes use of each architecture's advantages

to capture a variety of spatial and semantic patterns.

This novel method facilitates a more thorough evaluation and understanding of the patient's health by medical professionals, which can improve screening efficacy and reduce the probability of a misinterpretation. The early detection of silicosis is therefore essential.

2. Literature Review

To find silicosis and other related occupational lung disorders, a lot of research has been done using medical imaging and artificial intelligence. The bulk of previous research, however, concentrates on either segmentation or classification independently, and it usually targets pneumoconiosis or lung conditions more generally rather than silicosis specifically.

This section provides an overview of important efforts that have prepared the way for our suggested hybrid approach, which combines feature-based classification with improved segmentation for the detection of silicosis. A chest radiograph-based early Computer-Aided Diagnosis (CAD) system for silicosis identification [35]. They demonstrated that automated silicosis recognition was still feasible in 2014 with their method, which involved traditional image processing and texture feature extraction. But the absence of deep learning limited its adaptability and endurance. A multi-scale feature mapping method for pneumoconiosis staging was presented by Zhang [34]. Although these studies demonstrated that automation was feasible, they had trouble remaining robust under various imaging settings. Pneumoconiosis identification utilising coal workers' chest radiographs was investigated by Devnath *et al*. [10] using ensemble deep learning models. Their high-performing ensemble, which illustrated the advantages of combining several CNN architectures, confirmed our choice to include Inception V2, Inception V3, and ResNet50 in our proposed ensemble. A deep learning-based model for pneumoconiosis screening and staging was developed by Zhang *et al*. [33]. To detect silicosis, a deep learning model was applied to Computed Tomography (CT) images by Shivaanivarsha and Kavipriya [24], showcasing the viability of using Artificial Intelligence (AI) to detect diseases early. However, segmentation-driven feature analysis was not explored in the study, which solely concentrated on CT imaging. Pre-trained CNNs were used by Devnath *et al*. [9] to use transfer learning for black lung detection. Their research demonstrates how effective well-tuned pertained models are in identifying characteristics in medical images. A transfer learning system specifically designed for silicosis detection utilising chest radiography was developed by Shivaanivarsha *et al*. [25], where localisation accuracy was decreased by the lack of segmentation techniques. Ahmad *et al*. [1] examined machine learning techniques for automated interpretation of chest X-rays in a variety of conditions, indicating a high potential for diagnostic assistance. Gite *et al*. [13] suggested deep learning frameworks that enhanced the segmentation of lung images, which is essential for applications involving downstream classification, and demonstrated better Jaccard and Dice scores than baselines. But unfocused on the identification of silicosis and has limited clinical applicability for evaluation integration with classification models. A systematic study of AI methods for predicting diseases connected to the airways was suggested by Koul *et al*. [18] and emphasised how deep learning can be used to detect lung diseases. The significance of lung segmentation is highlighted in the study [27], which uses machine learning and deep learning approaches after lung segmentation for COVID-19 identification from CT scans. An enhanced transfer learning model for the identification of silicosis was presented by Sharma *et al*. [23], which demonstrated notable gains in accuracy. showed how pretrained models may be employed for occupational lung disease, but primarily relied on transfer learning without addressing segmentation. By using UNet for semantic segmentation and feature transfer from pre-trained CNN models, the research work in [5, 7, 15, 17, 19, 21] improves classification accuracy. Ali *et al*. [3], improved the study border recognition in low-contrast areas by using Generative Adversarial Networks (GANs) for lung segmentation in chest X-rays. Yu *et al*. [31] suggested a hybrid AttendSeg model that combines metaheuristic optimisation and mathematical modelling for breast cancer segmentation using gravitational clustering. An ensemble of deep learning models was employed by Singh *et al*. [26] and by Zaw and Mon [32] to increase and improve classification reliability. The increasing use of hybrid and ensemble AI models for the efficient diagnosis of diseases is highlighted in the literature. Comparative analysis of a few existing methods is listed in Table 1. Our study builds on these foundations by proposing a robust approach for detecting and classifying silicosis. The majority of the research that is now being conducted in the field of medical imaging focuses on diseases such as pneumonia, lung cancer, tuberculosis, and breast cancer. Models such as U-Net, ResNet, Inception have shown promising results regarding these diseases. The detection and classification of silicosis, on the other hand, is the subject of only a small number of studies, which leaves a significant gap in occupational health applications. Gunasekaran and Vivekasaran [14] developed the model for fetal heart disease detection using CNN and VGG 16 and achieved 99.86% accuracy.

Table 1. Comparative analysis.

Reference	Dataset Used	Methodology	Key Findings
Behera *et al.* [6]	Gastrointestinal tumour dataset	Inception ResNetV2 CNN framework	An enhanced tumour classification decision resulted in improved prediction.
Alotaibi [4]	Health care review dataset	Review of ensemble deep learning approaches	Benefits of ensembles across domains are highlighted, and no new theory is suggested.
Ahmed *et al.* [1]	Chest X-ray images	Review of machine learning approaches	It emphasised how ML-assisted radiography provides high sensitivity in common chest anomalies, decreases inter-observer variability, and speeds up diagnostics.
Saifullah *et al.* [22]	Brain tumour datasets	U-net with Resnet50 encoder	Enhanced tumour detection segmentation accuracy. This domain is limited to brain MRIs and does not apply to chest radiography.
Hussain *et al.* [16]	Brain tumour datasets	ResNet-50-CNN	Improved brain tumour detection with a cutting-edge CNN unique to neuroimaging. No framework for lung disease hybrid Segmentation-classification.
Lozano *et al.* [20]	Glaucoma datasets	R-CNN with Resnet50	Enhanced the identification of glaucoma by using segmentation-guided deep learning.
Chhabra and Kumar [8]	Chest X-rays	ResNet50	Increased feature extraction and reported good detection efficiency for pneumonia.
Yang *et al.* [29]	Ultrasound images	Dual branch ResNet50 with U-net segmentation	Increased classification accuracy by the integration of segmentation and classification. only the ultrasound modality was studied.
Xu *et al.* [28]	Review of medical image datasets	Survey of ResNet in medical imaging	A comprehensive analysis of the role that ResNet plays in processes involving segmentation and classification.
Younis *et al.* [30]	MRI images	ResNet50 Model	ResNet-50 has been shown to have excellent performance in the detection of brain tumours. The application is restricted to imaging of the brain.

3. Methodology

In this part, we outline the approach used in our research. Here, we provide a thorough breakdown of the suggested method for detecting silicosis in lungs. First, the data collection process has begun. Images of both healthy and silicosis-affected lung X-ray images are gathered and visualized. To evaluate the condition and decide how to treat various cases, the classification of silicosis disease is crucial. The diagnosis of silicosis is made using a variety of imaging techniques, such as CT, biopsy, bronchoscopy, and X-rays. In this research, we have collected X-ray images. The overall process flow is shown in Figure 1. Data preparation, segmentation with a modified UNet, deep feature extraction, MobilenetV2, InceptionV2, Resnet50, and custom CNN models are used for classification. Subsections go into detail about each stage of the strategy.

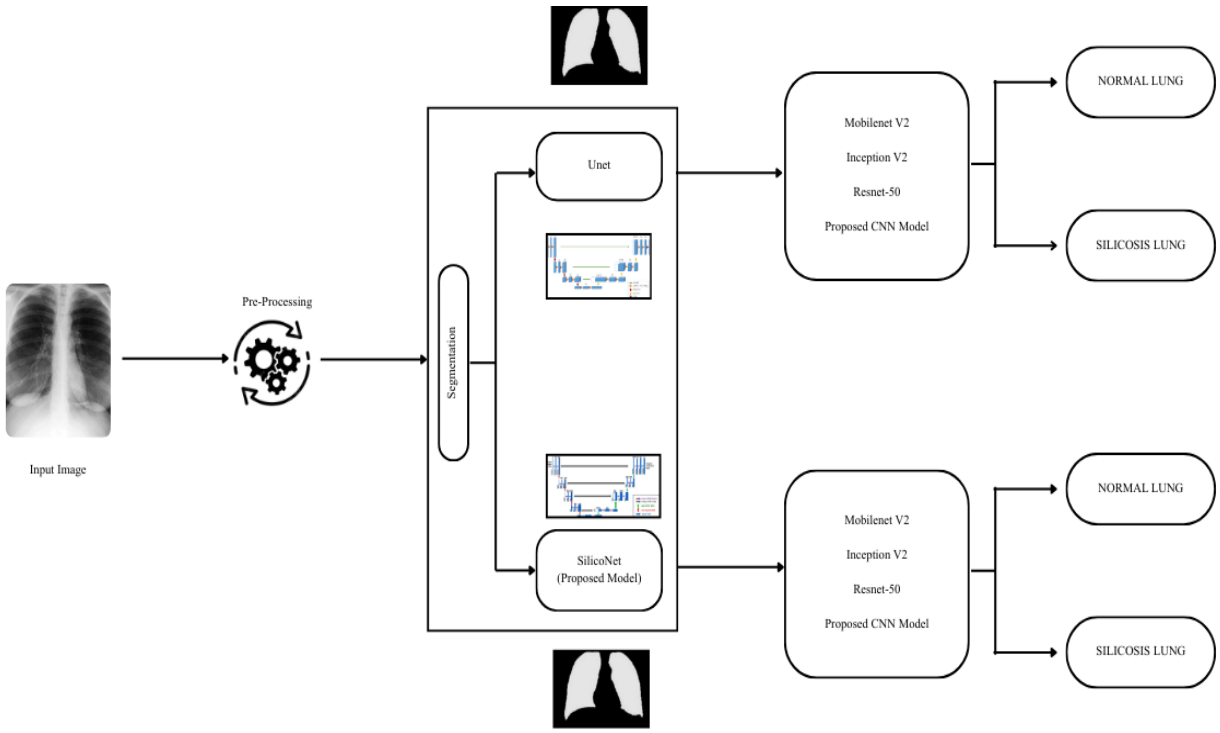


Figure 1. Block diagram of proposed methodology.

3.1. Pre-Processing

It is an essential process in the processing of medical images. Initially, input chest X-ray images are subjected to preprocessing operations to enhance image quality and standardize the data. Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to each radiograph to enhance local contrast and emphasize subtle fibrotic patterns, while simultaneously suppressing background artifacts. Image augmentation techniques were applied to the silicosis class in order to

rectify the class imbalance and guarantee that the classification model is not biased towards one class. Variations in brightness, flipping in both directions, rotation, and mild zooming were some of these enhancements. Almost 2500 images have been gathered from the different archives. Following preprocessing, we removed 1000 noisy images and then used Python's Augmentor package to enhance the data. The output size for the augmented data is set to 1100. 510 silicosis images and 1300 normal chest X-ray images were collected for this study from various public databases. We had 1310 silicosis images after the augmentation. After that, 20% and 80% of the dataset were divided for testing and training, respectively. All chest X-ray images are resized to a fixed dimension of 224×224 pixels, denoted as in Equation (1),

$$I_{resized} = R(I_{original}\ 224{,}224) \quad (1)$$

Where, $I_{resized}$ original X-ray image, $R()$=resizing function, Output $I_{resized}$ ensures compatibility with deep learning model input requirements. Every image is subjected to *CLAHE* to improve edge visibility and local contrast. This method brightens areas of interest without amplifying noise excessively. Let I_c be the contrast-enhanced image shown in Equation (2).

$$I_{c(x,y)} = CLAHE(I_{resized(x,y)}) \quad (2)$$

Where: x, y=pixel co-ordinates, $I_{c(x,\ y)}$=output pixel intensity after adaptive histogram enhancement. Then, Pixel values are normalized to fall within the range [0,1] to stabilize training, shown in Equation (3).

$$I_{c(x,y)} = CLAHE\left(I_{resized(x,y)}\right) I_{norm(x,y)} = max(I_c)\, min(I_c)\, I_{(x,y)(I_c)} \quad (3)$$

This transformation ensures that all intensity values are on a comparable scale, preventing gradient instability during backpropagation. To ensure consistency and enhance visual features in chest X-ray images, a structured preprocessing pipeline is applied. Initially, all radiographs are resized to 224×224 pixels using bilinear interpolation to maintain spatial resolution. To enhance local contrast and suppress background noise, CLAHE is employed, which adaptively redistributes pixel intensities within localized regions while restricting amplification via a predefined clip limit. Following contrast enhancement, pixel intensities are normalized to the range [0,1] using min-max scaling to stabilize training and ensure uniform distribution across all samples. The final output image is a pre-processed tensor of size R224×224×1, which is subsequently passed to the segmentation and feature extraction stages. This preprocessing strategy significantly improves the visibility of fibrotic patterns, enabling more accurate downstream learning. In our work, we used standard augmentation methods such as random rotations, horizontal and vertical flipping, zooming, and brightness adjustments. These transformations simulate real-world variations in patient positioning and imaging conditions, helping the model learn more robust features.

3.2. Segmentation and Feature Extraction

The suggested segmentation model, which draws inspiration from U-Net in segmentation, is based on a modified U-Net architecture that is optimized for the development of binary lung masks from chest X-ray images shown in Figure 2. The network has symmetric skip connections that aid in maintaining spatial context during reconstruction, and it has a standard encoder-decoder (contracting-expanding) topology. The network receives the input image, which is a greyscale chest X-ray, and gradually increases the feature channels while maintaining the spatial dimensions.

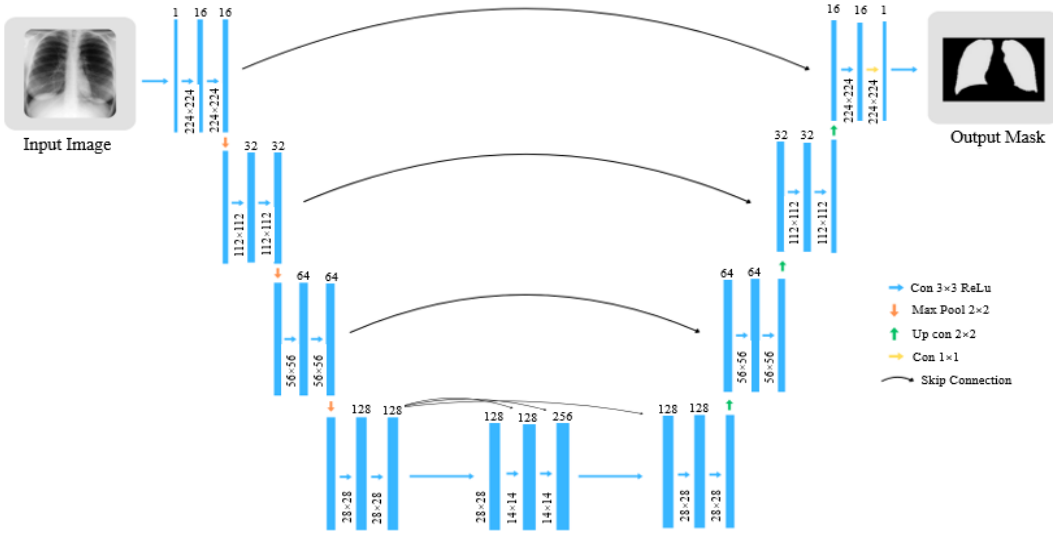


Figure 2. Proposed SilicoNet model in segmentation.

With skip connections that span corresponding layers to maintain fine-grained spatial information, the network is symmetric and consists of an expansive (decoder) path for spatial reconstruction and a contracting (encoder) channel for feature extraction. Standardized input image dimensions of 224x224x1

allow for grayscale medical image processing. To extract progressively abstract feature representations while decreasing the spatial resolution, the encoder path is composed of four consecutive blocks. A 2×2 max-pooling layer with a stride of 2 down-sampled the feature maps after two 3×3 convolutional layers with same padding and ReLU activation functions are present in each block. Four subsequent blocks make up the encoder path, which lowers the spatial resolution while extracting more abstract feature representations. Two 3x3 convolutional layers with same padding and ReLU activation functions are present in each block. The feature maps are then down-sampled using a 2x2 max-pooling layer with a stride of 2. The network receives the 224x224x1 input image in Block 1. A feature map of size 224×224×16 is produced by applying two convolutional layers, each with 16 filters. The output is then reduced to 112×112×16 by max-pooling, which is shown in below Equations (4), (5), and (6) below. Where x is input image, $f^{(l)}$ is output of the l^{th} convolution layer, $W^{(l)}$ is weights of the l^{th} convolution layer, $b^{(l)}$ is bias term of the l^{th} convolution, * denotes convolution operator, $\sigma^{(.)}$ represents Activation function, $\oplus$ denotes concatenation operator, $D^{(.)}$ is max pooling operation, $p^{(l)}$ is previous l^{th} layer input feature map, $c^{(l)}$ is concatenated feature map of l^{th} layer, $u^{(.)}$ denotes up sampling, $sigmoid^{(.)}$ is final activation for segmentation mask output.

$$f^{(1)} = \sigma\left(W^{(1)} * x + b^{(1)}\right) \quad (4)$$

$$f^{(2)} = \sigma\left(W^{(2)} * f + b^{(2)}\right) \quad (5)$$

$$p^{(1)} = D\left(f^{(2)}\right) (Size: 112 \times 112 \times 16) \quad (6)$$

The input in Block 2 is 112×112×16. Two 3×3 convolutions with 32 filters are used to process the feature maps, producing 112×112×32. This is then down sampled to 56×56×32 after pooling, as shown in Equations (7), (8), and (9).

$$f^{(3)} = \sigma(W^{(3)} * p^{(1)} + b^{(3)}) \quad (7)$$

$$f^{(4)} = \sigma(W^{(4)} * f^{(3)} + b^{(4)}) \quad (8)$$

$$p^{(2)} = D\left(f^{(4)}\right) (Size: 56 \times 56 \times 32) \quad (9)$$

Block 3 down-samples the 56×56×32 input to 28×28×64 after passing it through two convolutional layers with 64 filters, producing an output of 56×56×64, as given in Equations (10), (11), and (12).

$$f^{(5)} = \sigma(W^{(5)} * p^{(2)} + b^{(5)}) \quad (10)$$

$$f^{(6)} = \sigma(W^{(6)} * f^{(5)} + b^{(6)}) \quad (11)$$

$$p^{(2)} = D\left(f^{(4)}\right) (Size: 28 \times 28 \times 64) \quad (12)$$

In Block 4, two 3×3 convolutional layers with 128 filters are used to turn the 28×28×64 feature maps into 28×28×128, which is then down-sampled to 14×14×128 using max pooling, as shown in Equations (13), (14), and (15).

$$f^{(7)} = \sigma(W^{(7)} * p^{(3)} + b^{(7)}) \quad (13)$$

$$f^{(8)} = \sigma(W^{8} * f^{(7)} + b^{(8)}) \quad (14)$$

$$p^{(4)} = D\left(f^{(8)}\right) (Size: 14 \times 14 \times 128) \quad (15)$$

The Bridge or Bottleneck block, which acts as the interface between the encoder and decoder, is located at the core of the UNet. Here, two 3×3 convolutional layers with 256 filters are applied to the 14×14×128 feature map, producing deep semantic representations that are 14×14×256 in size, as per the Equations (16) and (17) shown.

$$f^{(9)} = \sigma\left(W^{(9)} * p^{(4)} + b^{(9)}\right) \quad (16)$$

$$f^{(10)} = \sigma\left(W^{(10)} * f^{(9)} + b^{(10)}\right) (Size: 14 \times 14 \times 256) \quad (17)$$

By gradually raising the feature map's spatial resolution, the decoder path recreates the segmented image. Before concatenating with the output of the associated encoder block via skip connections, each decoder block begins with a 2×2 transposed convolutions to double the spatial resolution. To further enhance the feature maps, two 3x3 convolutional layers with ReLU activation are added. The output of the 14×14×256 bottleneck is up-sampled to 28×28×128 in decoder Block 1, concatenated with the output of the encoder's Block 4 (28×28×128), and then goes through two convolutional layers with 128 filters to produce 28×28×128, which is shown in Equations (18), (19), (20), and (21).

$$u^{(1)} = u\left(f^{(10)}\right) (Upsample\ to\ 28 \times 28 \times 128) \quad (18)$$

$$c^{(1)} = u^{(1)} \oplus f^{(8)} \quad (19)$$

$$f^{(11)} = \sigma(W^{(11)} * c^{(1)} + b^{(11)}) \quad (20)$$

$$f^{(12)} = \sigma(W^{(12)} * f^{(11)} + b^{(12)}) \quad (21)$$

The encoder's Block 3 output is concatenated with the feature map after it has been up-sampled to 56×56×64 in decoder Block 2. After applying two 3×3 convolutions with 64 filters, 56×56×64 is produced, which is given in Equations (22), (23), (24), and (25).

$$u^{(2)} = u\left(f^{(10)}\right) (56 \times 56 \times 64) \quad (22)$$

$$c^{(2)} = u^{(2)} \oplus f^{(6)} \quad (23)$$

$$f^{(13)} = \sigma(W^{(13)} * c^{2} + b^{(13)}) \quad (24)$$

$$f^{(14)} = \sigma\left(W^{(14)} * f^{(13)} + b^{(14)}\right) \quad (25)$$

The feature map in decoder Block 3 is up-sampled to 112×112×32 and combined with the encoder output from Block 2. The outcome of two convolutions with 32 filters is 112×112×32, as given in Equations (26), (27), (28), and (29).

$$u^{(3)} = u\left(f^{(14)}\right) (112 \times 112 \times 32) \quad (26)$$

$$c^{(3)} = u^{(3)} \oplus f^{(4)} \quad (27)$$

$$f^{(15)} = \sigma(W^{(15)} * c^{(3)} + b^{(15)}) \quad (28)$$

$$f^{(16)} = \sigma(W^{(16)} * f^{(15)} + b^{(16)}) \quad (29)$$

Decoder Block 4 refines the feature map using two 3×3 convolutions with 16 filters, up samples it to

224×224×16, and concatenates it with Block 1's output, as shown in Equations (30), (31), (32), and (33).

$$u^{(4)} = u\left(f^{(16)}\right)(224 \times 224 \times 16) \quad (30)$$

$$c^{(4)} = u^{(4)} \oplus f^{(2)} \quad (31)$$

$$f^{(17)} = \sigma(W^{(17)} * c^{(4)} + b^{(17)}) \quad (32)$$

$$f^{(18)} = \sigma(W^{(18)} * f^{(17)} + b^{(18)}) \quad (33)$$

To create the final segmentation mask of size 224×224×1, a 1×1 convolutional layer with a single filter and sigmoid activation is utilized. Each pixel value in this mask represents the likelihood that silicosis is present. The segmentation mask, represented by the final output Y shown in Equation (34), has a probability value between 0 and 1 for each pixel, signifying the possibility of silicosis occurrence.

$$y = sigmoid\left(W^{(19)} * f^{(18)} + b^{(19)}\right)(Size: 224 \times 224 \times 1) \quad (34)$$

The following is a list of the evaluation metric's equations:

keywords are, True Positive (*Tru. P*), True Negative (*Tru. N*), False Positive (*Fal. P*), and False Negative (*Fal. N*) to indicate, respectively, silicosis related images that were correctly identified as such, silicosis related images that were mistakenly identified as normal lung tissue, silicosis-related images that were mistakenly identified as normal lung tissue, and silicosis-related images that were mistakenly identified as normal lung tissue.

$$Accuracy = \frac{(Tru.p + Tru.N)}{(Tru.P + Tru.N + Fal.P + Fal.N)} \quad (35)$$

Equation (35) evaluates the percentage of true predictions (both positive and negative) across all predictions to determine the model's overall accuracy.

$$Precision = \frac{(Tru.p)}{(Tru.P + Fal.P)} \quad (36)$$

Precision in Equation (36) highlights the model's capacity to prevent false positives by displaying the percentage of accurately predicted positive events among all predicted positives.

$$Recall = \frac{(Tru.p)}{(Tru.P + Fal.N)} \quad (37)$$

Recall in Equation (37) demonstrates how the model may reduce false negatives by identifying all true positive cases.

$$F1\ Score = 2 * \left[\frac{Pr * Rc}{Pr + Rc}\right] \quad (38)$$

F1-score given in Equation (38), is the balanced metric provided by the harmonic mean of precision (*Pr*) and recall (*Rc*) is particularly helpful in cases where the distribution of classes is unbalanced.

$$NVP = \frac{(Tru.N)}{(Tru.N + Fal.N)} \quad (39)$$

Equation (39) represents a Negative Predicted Value (NVP), which reduces the number of disease cases that are missed by calculating the percentage of actual negatives among all predicted negatives.

$$Jaccard\ Index = \frac{(Tru.p)}{(Tru.P + Fal.N)} \quad (40)$$

The Jaccard Index given in the Equation (40) evaluates how predicted and real regions overlap over the union. This is especially important when evaluating segmentation.

$$Dice = \frac{(2 * Tru.p)}{(2 * Tru.P + 2 * Fal.P + Fal.N)} \quad (41)$$

In segmentation tasks, a spatial overlap index measures how similar the ground truth and predicted masks overlap is a *Dice* coefficient metric represented in Equation (41). The probability that true non-hit variations will be correctly identified as non-hits is known as the Negative Predicted Value (NPV). This value can be set at any threshold. Matthews Correlation Coefficient (MCC)

$$MCC = \frac{((Tru.p + Tru.N) - (Fal.P \times Fal.N))}{\sqrt{(Tru.P + Fal.P)(Tru.P + Fal.n)(Tru.N + Fal.P)(Tru.N + Fal.N))}} \quad (42)$$

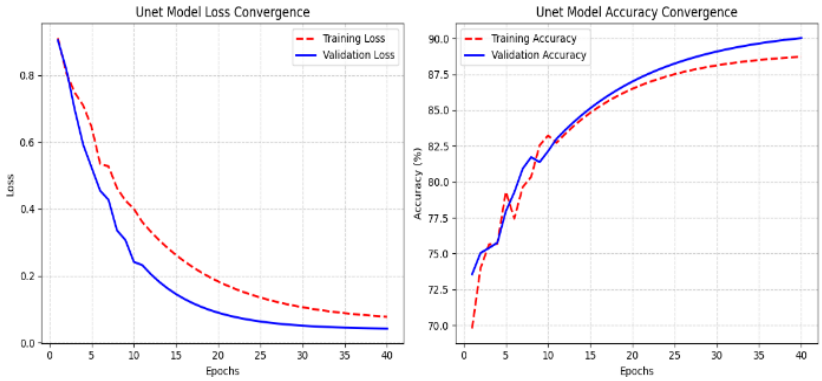
The formula for the Matthews correlation coefficient is shown in Equation (42). The MCC is a useful statistical metric for evaluating the performance of classification models, especially in binary classification situations. Because it considers each element of a confusion matrix, it is a balanced measure even in cases where datasets are unbalanced.

4. Results and Discussion

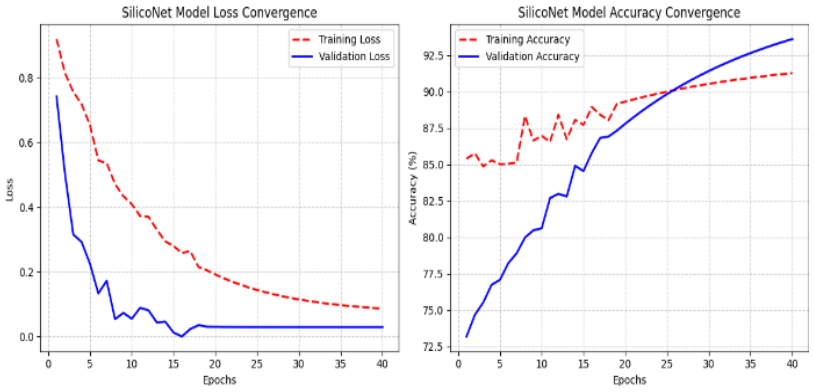
A comparison of the segmentation performance of the suggested SilicoNet architecture and the baseline UNet model during the training and validation stages is shown in Table 2. Loss, accuracy, dice coefficient, Jaccard index, and Mean Intersection over union (mean IoU) are among the evaluation metrics; these are important markers of the consistency and quality of segmentation. The proposed SilicoNet model performed better than the others. It obtained 92.40% accuracy, 0.967 Dice, 89.41 Jaccard Index, and 0.891 mean IoU during training. Even more promising were its validation results, which showed a mean IoU of 0.913, a dice coefficient of 0.9701, a Jaccard index of 90.21, and an accuracy of 96.98%. Particularly in images of X-rays affected by silicosis, the steadily improved metrics validate the SilicoNet architecture's robustness in achieving precise and accurate lung region segmentation.

Table 2. Segmentation model evaluation metrics comparison.

Model	Loss	Accuracy	Dice	Jaccard index	Mean IoU
UNet model training	0.0612	89.07	0.9313	69.35	0.784
UNet model validation	0.0399	90.79	0.9555	78.23	0.865
Proposed SilicoNet training	0.0698	92.40	0.967	89.41	0.891
Proposed SilicoNet validation	0.0291	96.98	0.9701	90.21	0.913



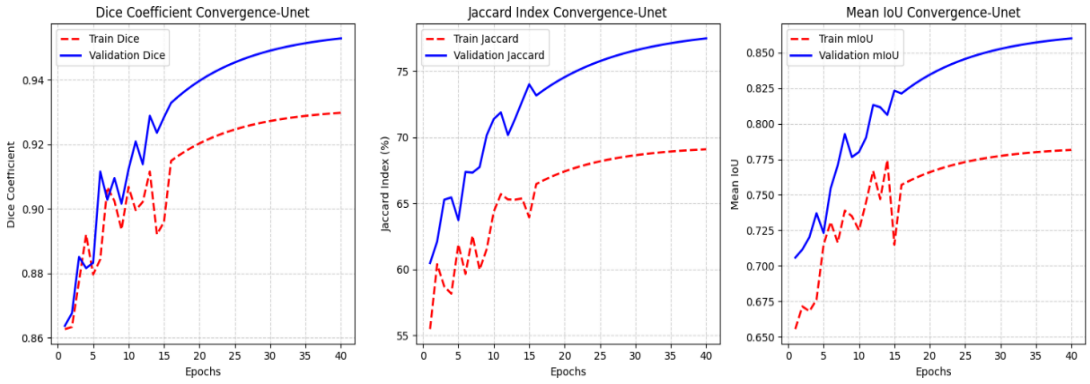
a) UNet model loss. b) UNet model accuracy.



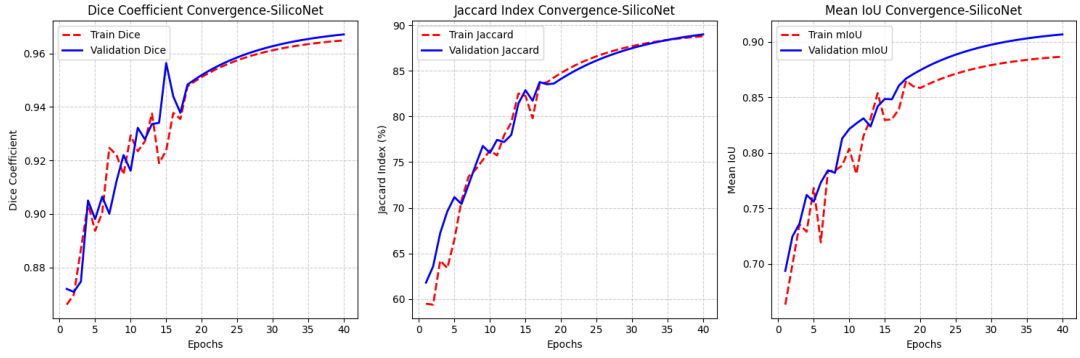
c) SilicoNet model loss. d) SilicoNet model accuracy.

Figure 3. A graphical comparison between the UNet model's (a, b) and siliconet model's (c, d) loss and accuracy convergence.

The UNet and SilicoNet segmentation model's training and validation results are shown graphically in Figure 3. In contrast, the SilicoNet model shows better segmentation capability with a faster, more stable convergence, reduced loss, and higher accuracy.



a) UNet Dice. b) UNet Jaccard index. c) UNet mean IoU.



d) SilicoNet dice. e) SilicoNet Jaccard index. f) SilicoNet mean IoU.

Figure 4. A graphical comparison between the UNet model's (a, b, c) and siliconet model's (d, e, f) dice, Jaccard index and mean IOU.

The segmentation performance metrics, Dice coefficient, Jaccard index, and mean IoU of the UNet and SilicoNet models are graphically presented in Figure 4. In terms of the SilicoNet model continuously

performs better than the UNet, indicating greater lung segmentation accuracy and region overlap.

Table 3. Segmentation model evaluation metrics comparison.

Model	Learning rate	Optimizer and epochs
UNet model training	lr: 1.0000e-04	Adam and 40
UNet model validation		
Proposed siliconet model training	lr: 1.5625e-05	Adam and 40
Proposed siliconet model validation		

The training parameters used for both the suggested SilicoNet architecture and the baseline UNet model are listed in Table 3. The Adam optimizer, widely employed because of its adaptability and effectiveness with sparse gradients, was used to train both models for 40 epochs. A learning rate of 1.0000e-04, a common beginning point for reliable convergence in encoder-decoder networks, was used for the UNet model. The SilicoNet model, on the other hand, used a lower learning rate of 1.5625e-05, enabling more precise updates during the training procedure.

4.1. Classification

Using segmented and non-segmented X-ray images for the classification of silicosis lung and normal lung, four different CNN models and our suggested custom model were trained, validated, and evaluated separately to see whether segmentation may improve classification accuracy. The comparative performance of all the models and our proposed classification custom model is displayed in the tables. Additionally, excellent evaluation metrics like precision, recall, F1-score, specificity, recognition time, training time, loss, MCC, and NPV. Following the conclusion of the training and validation phases, for the testing dataset, the performance of multiple CNNs was evaluated and contrasted. Figure 5 depicts the proposed custom CNN model block diagram used for classification. To extract low-level characteristics like edges and textures, the model starts with a Conv2D layer that applies 32 filters of a specific kernel size 3×3 to the input image. Following each convolutional process, max pooling layers are positioned intelligently to minimize the feature map's spatial dimensions. A flatten layer is incorporated into the model after the convolutional and pooling processes, transforming the acquired features into a one-dimensional vector. Two dense layers are applied to this vector: the first has 32 neurons, and the last output layer has two neurons.

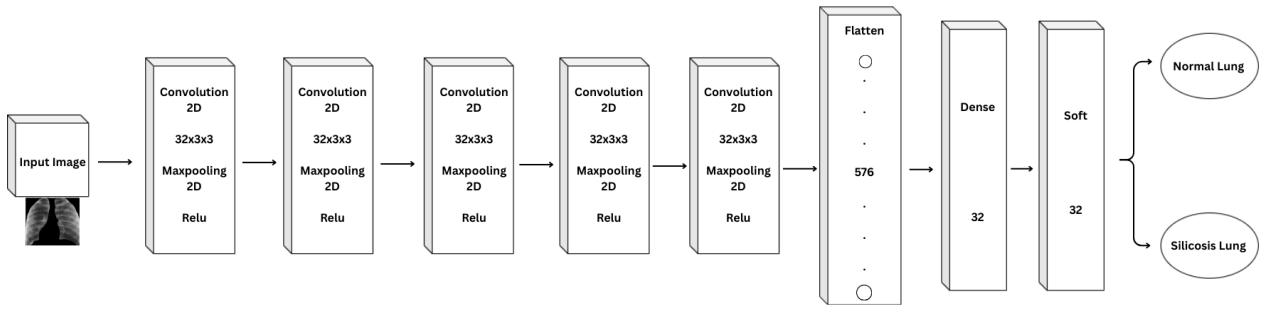


Figure 5. Proposed custom CNN model for classification.

Table 4. Results of accuracy and MCC of the models.

Segmentation model	Accuracy				Matthew correlation coefficient			
	MobileNet V2	Inception V2	Resnet 50	Proposed CNN	MobileNet V2	Inception V2	Resnet 50	Proposed CNN
UNet	90.14	92.14	90.51	91.01	0.8930	0.891	0.90	0.9075
Proposed SilicoNet model	95.6	95.5	94.5	96.40	0.9030	0.891	0.90	0.9105

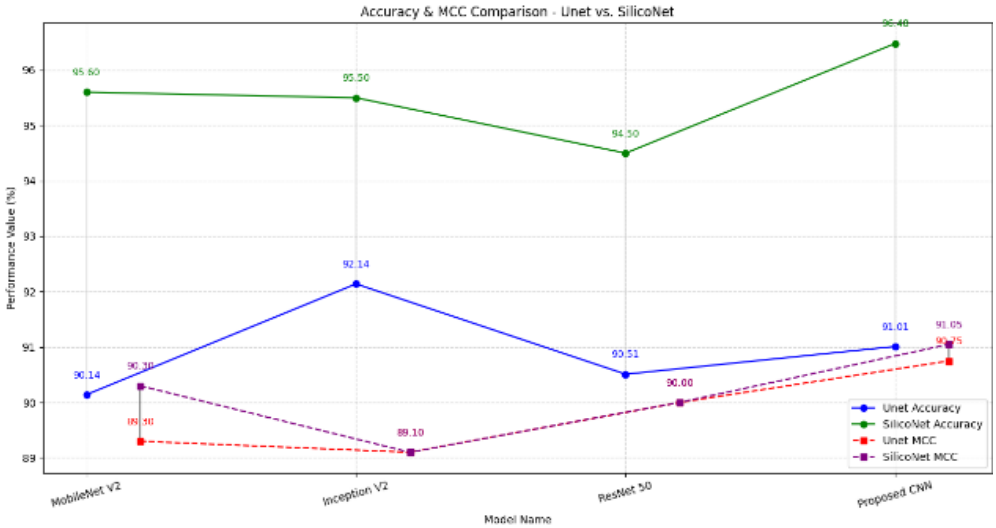
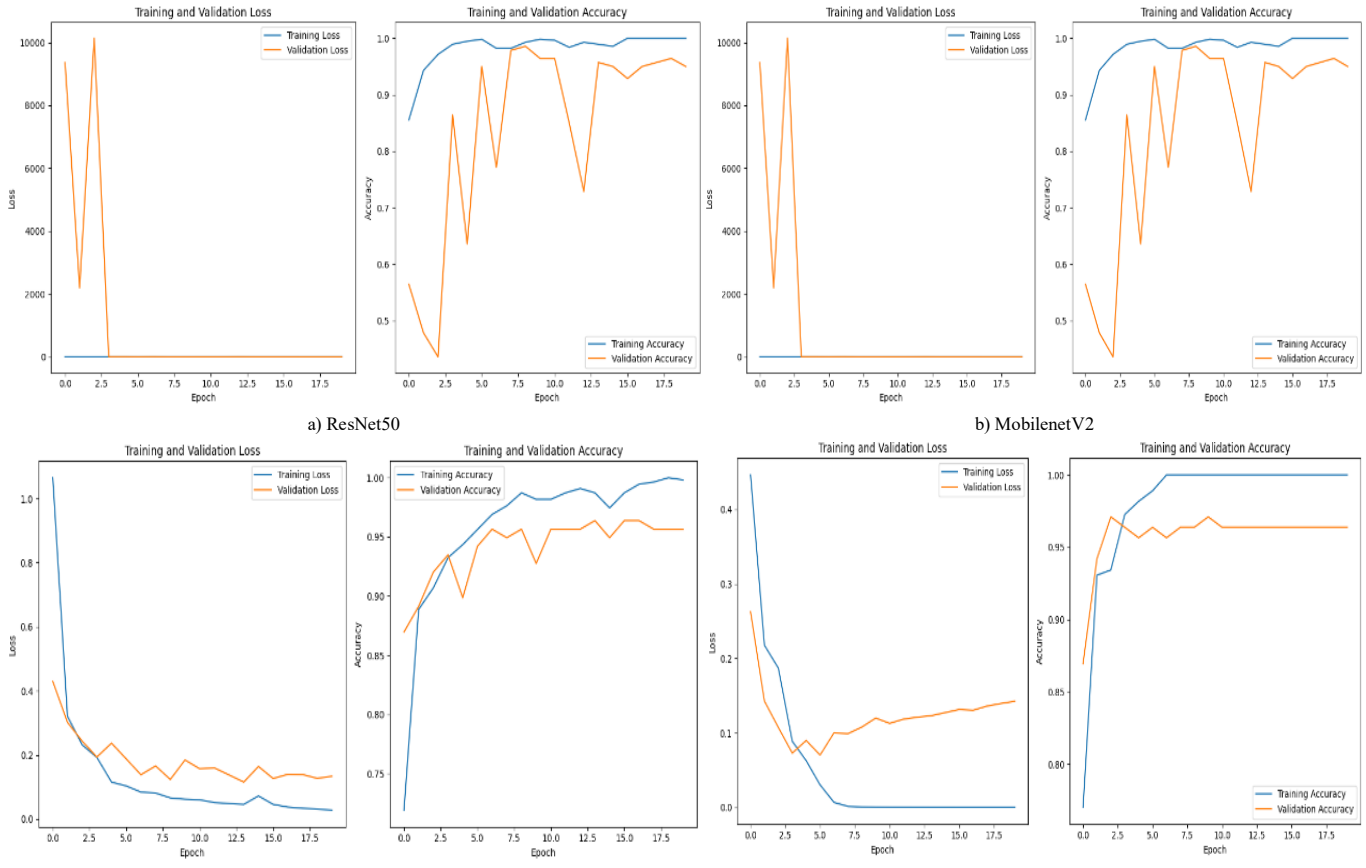


Figure 6. A graphical comparison of the four model's accuracy and MCC between the UNet and SilicoNet models.

Four feature extractors, MobileNetV2, InceptionV2, ResNet50, and the proposed CNN model, as well as two segmentation models, UNet and the proposed SilicoNet model, are used to compare classification performance in Table 4. Accuracy and MCC, two key indicators that are essential for interpreting the classification system's reliability and capacity for generalisation, form the basis of the evaluation. Across all feature extractors, the results clearly show that the SilicoNet-based classification methodology performs better than the UNet-based approach. Significantly, the suggested CNN achieves the highest accuracy of 96.40% when combined with SilicoNet, as compared to 91.01% when combined with UNet. Similar improvements are shown in other models: when collaborating with SilicoNet, MobileNet V2 improves from 90.14% to 95.6%, Inception V2 from 92.14% to 95.5%, and ResNet-50 from 90.51% to 94.5%. Graphical comparison is shown in Figure 6.

This performance sequence is further reinforced by the MCC scores. With the greatest MCC of 0.9105, the proposed CNN with SilicoNet demonstrated a higher correlation between the predicted and actual classes. The MCC of MobileNetV2 increases from 0.8930 to 0.9030, while ResNet-50 and Inception V2 stay largely unchanged but retain a little from SilicoNet's better segmentation. Regardless of the feature extractor selection, these data clearly show that SilicoNet boosts downstream classification performance and feature discrimination. When combined with SilicoNet, the proposed CNN classifier continuously provides better results on all measures. With the proposed custom CNN model, Figure 7 shows the training loss and accuracy curves of three common CNN designs. The customized model exhibits a higher classification accuracy, a smaller final loss, and a faster rate of convergence.



a) ResNet50 b) MobilenetV2

c) InceptionV2 d) SilicoNet

Figure 7. Loss and accuracy curve of three distinct CNN models and the proposed custom CNN classification model.

Table 5. UNet and the proposed SilicoNet model's overall performance.

Parameters	MobileNetV2		InceptionV2		Resnet50		Proposed CNN model	
	UNet	Siliconet	UNet	Siliconet	UNet	Siliconet	UNet	Siliconet
Precision	82.23	89.29	73.23	80.70	88.20	90.41	91.34	94.34
Recall	83.24	87.29	85.34	89.82	74.35	89.63	89.53	93.53
F1-score	86.12	89.29	78.12	82.20	84.20	89.16	89.93	92.93

Table 5 and Figure 8 display a comparative analysis of segmentation model performance based on several classification measures. The custom CNN+SilicoNet model outperforms UNet models, with the highest metrics. It shows how SilicoNet increases feature extraction and classification across baseline models. The metrics include Precision, Recall, F1-score, Specificity, and NPV for both the UNet and the

proposed SilicoNet model. Across all models, the SilicoNet model continuously outperformed UNet for precision the Proposed CNN produced the highest value, 94.34%. In the same way, SilicoNet recall scores outperformed all UNet configurations, topping at 93.53% (Proposed CNN). With a maximum of 92.93% utilizing the Proposed CNN, the F1-score further validated SilicoNet robustness by displaying superior harmonic mean values across all classifiers.

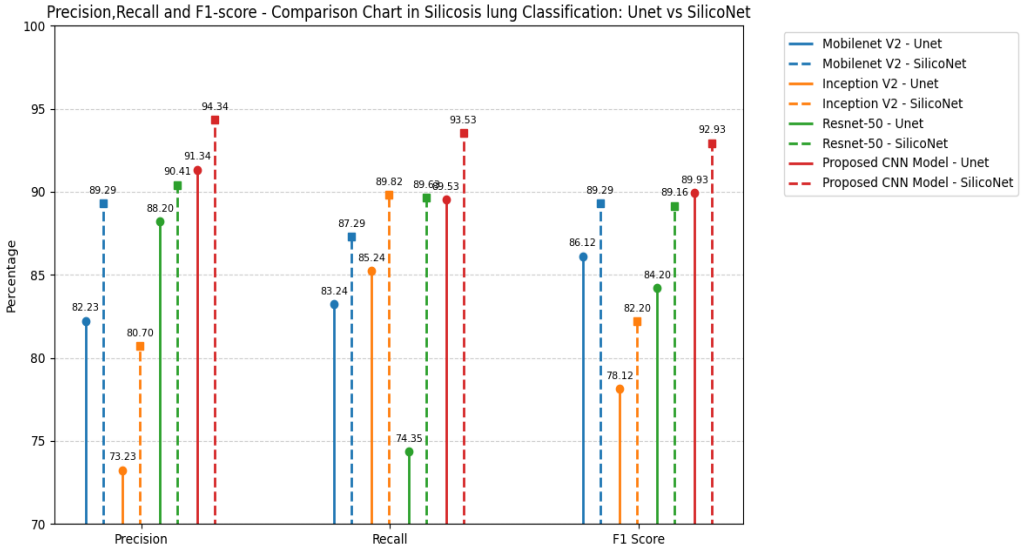


Figure 8. A graphical comparison of the four model's precision, recall, and F1-score between UNet and SilicoNet.

Table 6. UNet and the proposed SilicoNet model's overall performance.

Parameters	MobileNet V2		InceptionV2		Resnet50		Proposed CNN Model	
	UNet	SilicoNet	UNet	SilicoNet	UNet	SilicoNet	UNet	SilicoNet
Specificity	0.786	0.897	0.754	0.892	0.792	0.892	0.897	0.907
NPV	0.75	0.8226	0.791	0.848	0.82	0.898	0.88	0.914

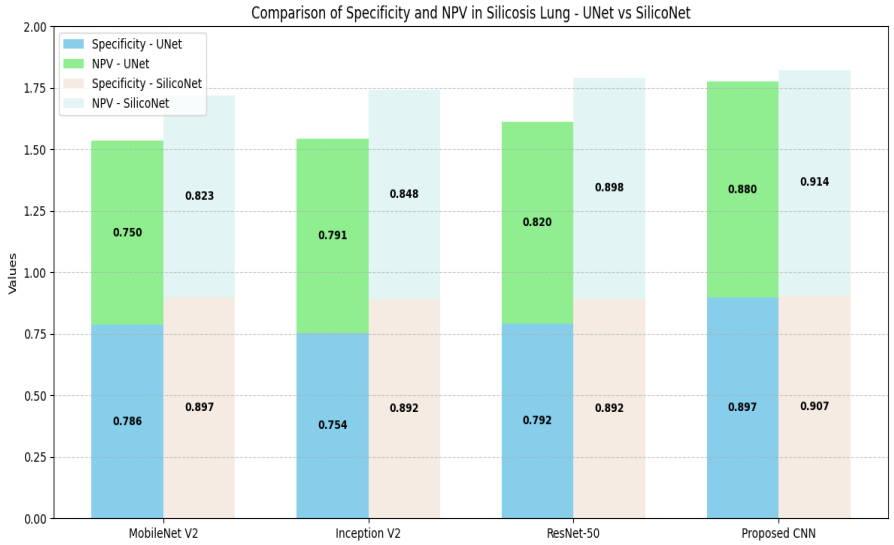
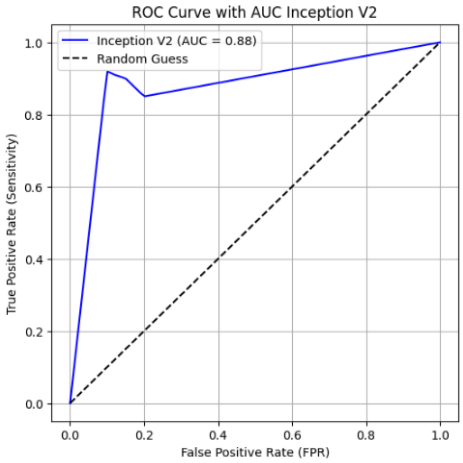


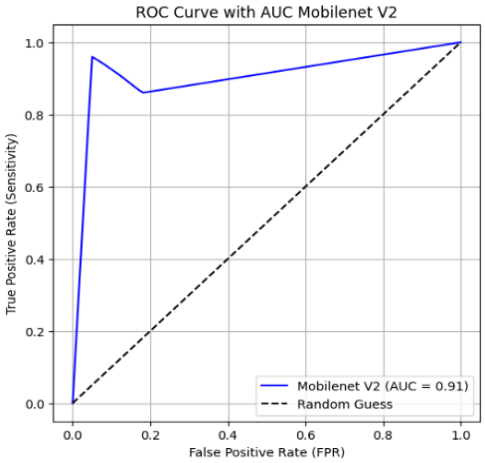
Figure 9. A graphical comparison of the four models' specificity and NPV between UNet and SilicoNet.

Using both UNet and SilicoNet segmentation outputs. Table 6 and Figure 9 show the Specificity and negative predicted values of the UNet and SilicoNet models. Higher confidence in real negative predictions was indicated by the NPV, which similarly showed significant improvements in SilicoNet, increasing to 0.914.

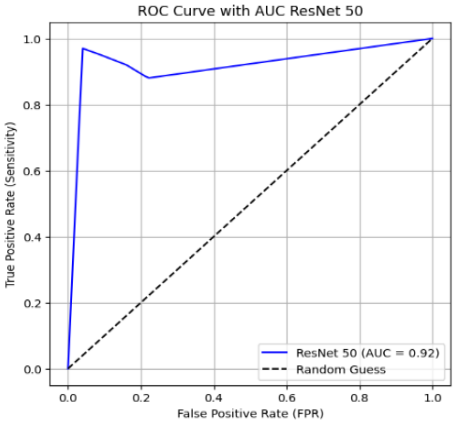
False Positive Rate or FPR (1-specificity) trade-off is graphically represented by the ROC curve across a range of threshold values. How successfully our model differentiates between the healthy (negative) and diseased (positive) lung areas is assessed by the curve. Figure 10 shows the ROC curves for Silicosis and normal lung classification using CNN-based models and the proposed model. Where InceptionV2 Model achieves Area Under the Curve (AUC) as 0.88, MobilenetV2 got 0.91, Resnet 50 model AUC is 0.92, and the proposed model AUC is 0.93, which is slightly superior to other CNN-based models.



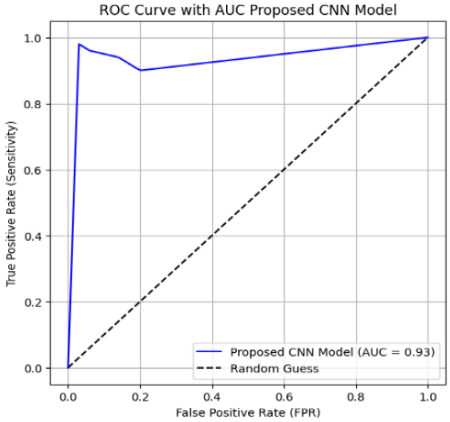
a) Inception V2 ROC.



b) Mobilenet V2 ROC.



c) ResNet 50 ROC.



d) Proposed Model ROC.

Figure 10. The ROC curves for Silicosis classification using CNN-based models (a, b, c) and the proposed model (d).

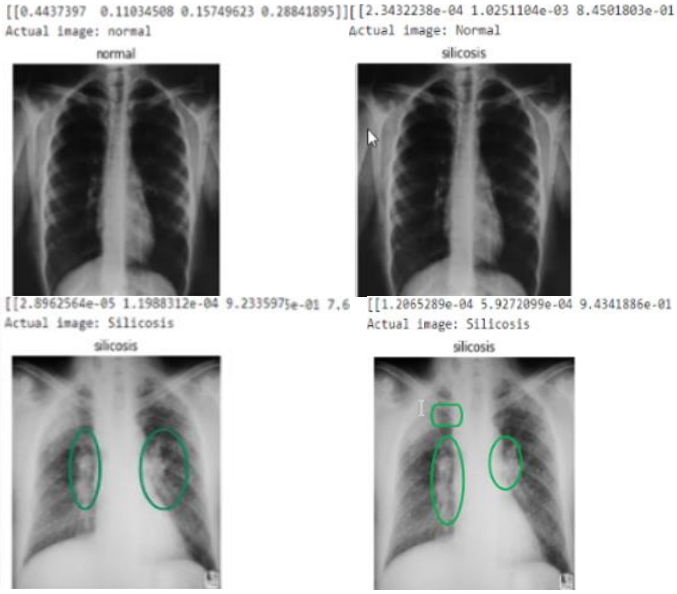


Figure 11. Sample predicted results.

Table 7. Comparison of the proposed work with existing techniques.

Author	Images and methodology	Accuracy
Sharma *et al.* [23]	X-ray images, transfer learning model.	81.66%
Ehrlich *et al.* [11]	X-ray images, CAD.	90%
Aksoy *et al.* [2]	CT images, Support Vector Machine (SVM).	97.29%
Eravcı *et al.* [12]	Clinical parameters, adaptive neuro-fuzzy Inference system.	94%
Proposed SilicoNet model	X-ray images with segmentation.	F1-score-92.33% Sensitivity-93.53% Accuracy-96.40%

Figure 11 shows the prediction of the Suggested model results in the Normal lung and the Silicosis lung. Where the green border is a correct prediction, and the red border is a misclassified prediction. In Table 7, the suggested study is compared to models outlined in the literature. Only Few studies address the silicosis detection. The proposed approach is more useful for real-world applications since it can pinpoint the impacted areas.

5. Conclusions

This study presents SilicoNet, an improved U-Net based segmentation model integrated with a custom CNN classifier for the automated detection and classification of silicosis using chest radiographs. Experimental results indicate that SilicoNet consistently improves the performance of the standard UNet and other baseline CNN architectures, including MobileNetV2, InceptionV2, and ResNet50. The model demonstrated a maximum accuracy of 96.40% and exhibited enhanced precision, recall, F1-score, and MCC, indicating its robustness in identifying silicosis at various stages. The findings demonstrate that SilicoNet is a dependable and effective tool that aids radiologists in clinical decision-making and facilitates scalable AI-driven screening in occupational healthcare. Although these results are encouraging, this study has specific limitations. The dataset size, while increased, remained relatively small in comparison to extensive clinical repositories, potentially impacting generalisability. Future work should incorporate advanced interpretability techniques, such as Grad-CAM and SHAP visual explanations, to enhance trust and clinical usability. Lastly, integrating SilicoNet with additional multi-modal learning or metaheuristic optimization-based feature selection techniques may enhance accuracy, robustness, and practicality. The Nvidia GeForce RTX 3050 Intel i7 platform served as the power source for all of the experiments.

References

[1] Ahmad H., Milne M., Buchlak Q., Ektas N., and et al., "Machine Learning Augmented Interpretation of Chest X-Rays: A Systematic Review," *Diagnostics*, vol. 13, no. 4, pp. 1-31, 2023. https://doi.org/10.3390/diagnostics13040743

[2] Aksoy H., Atila U., and Arslan S., "Deep Learning Based Detection of Silicosis from Computed Tomography Images," *Computer Methods and Programs in Biomedicine Update*, vol. 6, pp. 1-18, 2024. https://doi.org/10.1016/j.cmpbup.2024.100166

[3] Ali O., Ali H., Shah S., and Shahzad A., "Implementation of a Modified U-Net for Medical Image Segmentation on Edge Devices," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 11, pp. 4593-4597, 2022. https://doi.org/10.1109/TCSII.2022.3181132

[4] Alotaibi A., "Ensemble Deep Learning Approaches in Health Care: A Review," *Computers, Materials and Continua*, vol. 82, no. 3, pp. 3741-3771, 2025. https://doi.org/10.32604/cmc.2025.061998

[5] Arvind S., Tembhurne J., Diwan T., and Sahare P., "Improvised Light Weight Deep CNN Based U-Net for the Semantic Segmentation of Lungs from Chest X-Rays," *Results in Engineering*, vol. 17, pp. 100929, 2023. https://doi.org/10.1016/j.rineng.2023.100929

[6] Behera N., Gurunathan K., Kaliappan S., Rai V., and Manikandan V., "Enhanced CNN-Based Decision Support Framework for Gastrointestinal Neuroendocrine Tumors Using Inception-Resnet-V2," *in Proceedings of the IEEE International Conference on Big Data and Machine Learning*, Bhopal, pp. 145-150, 2024. https://doi.org/10.1109/ICBDML60909.2024.10577307

[7] Chen W., Yang F., Zhang X., Xu X., and Qiao X., "MAU-Net: Multiple Attention 3D U-Net for Lung Cancer Segmentation on CT Images," *Procedia Computer Science*, vol. 192, pp. 543-552, 2021. https://doi.org/10.1016/j.procs.2021.08.056

[8] Chhabra M. and Kumar R., "An Efficient ResNet-50 Based Intelligent Deep Learning Model to Predict Pneumonia from Medical Images," *in Proceedings of the International Conference on Sustainable Computing and Data Communication Systems*, Erode, pp. 1714-1721, 2022. https://doi.org/10.1109/ICSCDS53736.2022.9760995

[9] Devnath L., Luo S., Summons P., and Wang D., "An Accurate Black Lung Detection Using Transfer Learning Based on Deep Neural Networks," *in Proceedings of the International Conference on Image and Vision Computing New Zealand*, Dunedin, pp. 1-6, 2019. https://doi.org/10.1109/IVCNZ48456.2019.8960961

[10] Devnath L., Luo S., Summons P., Wang D., and et al., "Deep Ensemble Learning for the Automatic Detection of Pneumoconiosis in Coal Worker's Chest X-Ray Radiography," *Journal of Clinical Medicine*, vol. 11, no. 18, pp. 1-23, 2022. https://doi.org/10.3390/jcm11185342

[11] Ehrlich R., Barker S., Naude J., Rees D., and et al., "Accuracy of Computer-Aided Detection of Occupational Lung Disease: Silicosis and Pulmonary Tuberculosis in Ex-Miners from the South African Gold Mines," *International Journal of Environmental Research and Public Health*,

vol. 19, no. 19, pp. 1-14, 2022. https://doi.org/10.3390/ijerph191912402

[12] Eravci D. and Erdebilli B., "Diagnosing Silicosis Using Adaptive Neuro-Fuzzy Inference System," *Research Square*, vol. 1, pp. 1-18, 2022. https://doi.org/10.21203/rs.3.rs-450276/v1

[13] Gite S., Mishra A., and Kotecha K., "Enhanced Lung Image Segmentation Using Deep Learning," *Neural Computing and Applications*, vol. 35, no. 31, pp. 22839-22853, 2023. https://doi.org/10.1007/s00521-021-06719-8

[14] Gunasekaran S. and Vivekasaran S., "Disease Prognosis of Fetal Heart's Four-Chamber and Blood Vessels in Ultrasound Images Using CNN Incorporated VGG 16 and Enhanced DRNN," *The International Arab Journal of Information Technology*, vol. 21, no. 6, pp. 1111-1127, 2024. https://doi.org/10.34028/iajit/21/6/13

[15] Hossain M., "Microc Alcification Segmentation Using Modified U-Net Segmentation Network from Mammogram Images," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 2, pp. 86-94, 2022. https://doi.org/10.1016/j.jksuci.2019.10.014

[16] Hussain S., Naz T., Shakeel M., Akram F., and et al., "Advanced Imaging Technique-based Brain Tumor Segmentation Using ResNET-50 CNN," *in Proceedings of the 5th International Conference on Sustainable Innovation in Engineering and* Technology, Kuala Lumpur, pp. 20156, 2024. https://doi.org/10.1063/5.0229438

[17] Islam R., Rahman M., Ali S., Nafi A., and et al., "Enhancing Breast Cancer Segmentation and Classification: An Ensemble Deep Convolutional Neural Network and U-Net Approach on Ultrasound Images," *Machine Learning with Applications*, vol. 16, pp. 100555, 2024. https://doi.org/10.1016/j.mlwa.2024.100555

[18] Koul A., Bawa R., and Kumar Y., "Artificial Intelligence Techniques to Predict the Airway Disorders Illness: A Systematic Review," *Archives of Computational Methods in Engineering*, vol. 30, no. 2, pp. 831-864, 2023. https://doi.org/10.1007/s11831-022-09818-4

[19] Kumarasinghe K., Kolonne S., Fernando K., and Meedeniya D., "U-Net Based Chest X-Ray Segmentation with Ensemble Classification for Covid-19 and Pneumonia," *International Journal of Online and Biomedical Engineering*, vol. 18, no. 7, pp. 161-174, 2022. https://doi.org/10.3991/ijoe.v18i07.30807

[20] Lozano M., Vera L., Reyes A., Arteaga D., and et al., "Deep Learning for Glaucoma Detection: R-Cnn Resnet-50 and Image Segmentation," *Journal of Advances in Information Technology*, vol. 14, no. 6, pp. 1186-1197, 2023. DOI: 10.12720/jait.14.6.1186-1197

[21] Munawar F., Azmat S., Iqbal T., Gronlund C., and Ali H., "Segmentation of Lungs in Chest X-Ray Image Using Generative Adversarial Networks," *IEEE Access*, vol. 8, pp. 153535-153545, 2020. https://doi.org/10.1109/ACCESS.2020.3017915

[22] Saifullah S., Yudhana A., and Suryotomo A., "Automatic Brain Tumor Segmentation: Advancing U-Net with ResNet50 Encoder for Precise Medical Image Analysis," *IEEE Access*, vol. 13, pp. 43473-43489, 2025. https://doi.org/10.1109/ACCESS.2025.3547430

[23] Sharma G., Harjule P., Agarwal B., and Kumar R., "Silicosis Detection Using Extended Transfer Learning Model," *in Proceedings of the International Conference on Recent Trends in Image Processing and Pattern Recognition*, Derby, pp. 111-126, 2023. https://doi.org/10.1007/978-3-031-53085-2_10

[24] Shivaanivarsha N. and Kavipriya P., "Identification of Silicosis using Deep Learning Model on Computed Tomography Images," *in Proceedings of the IEEE 3rd Mysore Sub Section International Conference*, Hassan, pp. 1-6, 2023. https://doi.org/10.1109/MysuruCon59703.2023.10396879

[25] Shivaanivarsha N., Kavipriya P., and Vigita S., "A Deep Transfer Learning Approach for the Detection of Silicosis Using Chest Radiography," *in Proceedings of the National Conference on Control Instrumentation System Conference*, vol. 1236, pp. 575-589, 2023. https://doi.org/10.1007/978-981-97-5866-1_41

[26] Singh R., Gupta S., Bharany S., Almogren A., and et al., "Ensemble Deep Learning Models for Enhanced Brain Tumor Classification by Leveraging ResNet50 and EfficientNet-B7 on High-Resolution MRI Images," *IEEE Access*, vol. 12, pp. 178623-178641, 2024. https://doi.org/10.1109/ACCESS.2024.3494232

[27] Tarhini H., Mohamad R., Rammal A., and Ayache M., "Lung Segmentation Followed by Machine Learning and Deep Learning Techniques for COVID-19 Detection in Lung CT Images," *in Proceedings of the 6th International Conference on Advances in Biomedical* Engineering, Werdanyeh, pp. 222-227, 2021. https://www.researchgate.net/publication/355166634_Lung_Segmentation_followed_by_Machine_Learning_Deep_Learning_Techniques_for_COVID-19_Detection_in_lung_CT_Images

[28] Xu W., Fu Y., and Zhu D., "ResNet and its Application to Medical Image Processing: Research Progress and Challenges," *Computer Methods and Programs in Biomedicine*, vol. 240, pp. 107660, 2023. https://doi.org/10.1016/j.cmpb.2023.107660

[29] Yang X., Qu S., Wang Z., Li L., and et al., "The Study on Ultrasound Image Classification Using a Dual-Branch Model Based on Resnet50 Guided

by U-Net Segmentation Results," *BMC Medical Imaging*, vol. 24, no. 1, pp. 1-16, 2024. https://doi.org/10.1186/s12880-024-01486-z

[30] Younis A., Li Q., Afzal Z., Adamu M., and et al., "Abnormal Brain Tumors Classification Using Resnet50 and its Comprehensive Evaluation," *IEEE Access*, vol. 12, pp. 78843-78853, 2024. https://doi.org/10.1109/ACCESS.2024.3403902

[31] Yu L., Shaheema S., Sunil J., Govindan V., and et al., "Breast Cancer Segmentation Using a Hybrid AttendSeg Architecture Combined with a Gravitational Clustering Optimization Algorithm Using Mathematical Modelling," *Open Physics*, vol. 21, no. 1, pp. 1-13, 2023. file:///C:/Users/acit2k/Downloads/10.1515_phys-2023-0105.pdf

[32] Zaw K. and Mon A., "Enhanced Multi-Class Skin Lesion Classification of Dermoscopic Images Using an Ensemble of Deep Learning Models," *Journal of Computing Theories and Applications*, vol. 2, no. 2, pp. 256-267, 2024. https://doi.org/10.62411/jcta.11530

[33] Zhang L., Rong R., Li Q., Yang D., and et al., "Deep Learning-Based Model for Screening and Staging Pneumoconiosis," *Scientific Reports*, vol. 11, no. 1, pp. 1-7, 2021. https://doi.org/10.1038/s41598-020-77924-z

[34] Zhang Y., "Computer-Aided Diagnosis for Pneumoconiosis Staging Based on Multi-Scale Feature Mapping," *International Journal of Computational Intelligence Systems*, vol. 14, no. 1, pp. 1-11, 2021. https://doi.org/10.1007/s44196-021-00046-5

[35] Zhu L., Zheng R., Jin H., Zhang Q., and Zhang W., "Automatic Detection and Recognition of Silicosis in Chest Radiograph," *Bio-Medical Materials and Engineering*, vol. 24, no. 6, pp. 3389-3395, 2014. https://doi.org/10.3233/bme-141162

Shivaanivarsha Natarajan received a B.E. degree in ECE from Madras University in 1998. She received her Master's in Applied Electronics from Sathyabama University, Chennai, in 2005. She has more than two decades of experience in the academic field. She is currently an Associate Professor at Sri Sai Ram Engineering College, Chennai, Tamil Nadu, India. She is pursuing her Ph.D. Degree in Sathyabama Institute of Science and Technology, Chennai in the area of Image Processing.

Kavipriya Periyasamy obtained her Bachelor's in Engineering (B.E.) in the year 2000 from AVC College of Engineering, Bharathidasan University. She received her Master's degree in Applied Electronics, first class with distinction, from Sathyabama University, Chennai in 2005 and also obtained her doctorate in the year of 2015. Her current areas of interest are Image Processing, Wireless Networks. She has published more research articles in reputed journals and conference proceedings.