

Research Article

An Enhancing Lung Disease Detection with Scalable UNet-Mamba Segmentation

Sivasakthi S¹, Radha V²

^{1,2}Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu, India.

¹Corresponding Author : 21phcsp002@avinuty.ac.in

Received: 24 March 2026

Revised: 13 August 2026

Accepted: 26 August 2026

Published: 30 September 2026

Abstract - Lung diseases are conditions that damage the lung tissues, airways or blood vessels of the lungs, causing breathing difficulties and decreased oxygen supply to the body. Computed Tomography (CT) plays a crucial role in lung disease prediction because it provides detailed cross-sectional images for detecting tumors at initial stages. However, CT scans have drawbacks such as radiation exposure, high cost, large data size, and the possibility of false positives. Recently, many Deep Learning (DL) techniques have been used to automatically analyze CT images for early and accurate detection of lung diseases, improving diagnostic efficiency. However, existing models may struggle to capture fine spatial details, leading to inaccurate boundary delineation and poor segmentation of small or low-contrast regions. Some models are also sensitive to imaging artifacts, which can reduce performance in real-world conditions. In this paper, a lung disease prediction model is developed using ResNet and Mamba Network (LRMNet) to solve the aforementioned issues for lung disease prediction. The first step is improving the model's ability to conduct pre-processing and enrich the data using a Deep Convolutional Generative Adversarial Network (DCGAN). Then, the Scalable Vision UNet-Mamba (SVU-M) model is developed to efficiently capture fine spatial details along with long-range contextual information. SVU-M leverages Scalable Vision State Space (SVSS) blocks to efficiently model the global dependencies while maintaining linear computational complexity. The integration of SVSS within a U-shaped framework enhances the boundary delineation and improves segmentation of small or low-contrast regions, which increases the efficiency against imaging artifacts for more reliable performance. The decoder feature maps of SVU-M are further processed using Global Average Pooling (GAP) to determine image-level feature representations for classification. Simultaneously, the input image is directly fed into ResNet-50 for capturing the whole image's semantic features. A Multi-Layer Perceptron (MLP) classifier is trained with the extracted features to forecast lung diseases. A total accuracy of 94.15% on the customized dataset and 93.29% on the SARS-CoV-2 CT dataset was achieved by the suggested model, according to the experimental results, exceeding the performance of other models.

Keywords - Computed Tomography, Deep Learning, Lung Disease, ResNet-50, U-Mamba.

1. Introduction

Lung diseases are medical conditions that affect the lungs, which are the vital organs responsible for breathing and gas exchange in the human body [1]. The lungs deliver oxygen to the blood and eliminate carbon dioxide, and any structural or functional impairment can result in serious health problems. Lung diseases are a major global health concern and are caused by factors such as smoking, air pollution, infections, occupational exposure and genetic predisposition [2]. Pneumonia, tuberculosis and Chronic Obstructive Pulmonary Disease (COPD) are examples of lung diseases [3]. Since most lung diseases do not initially present with visible symptoms, early prediction and diagnosis are crucial [4]. Proper treatment, slowed disease progression, increased life expectancy, and improved quality of life are all outcomes of early diagnosis.

Lung disease can be identified using a range of medical imaging techniques, including CT scans, MRIs, ultrasounds and Positron Emission Tomography (PET) [5]. Because it produces high-quality segmented images of the lung structure, CT imaging is the most appropriate of these modalities for a better understanding of the lung's anatomy [6]. CT scans are a modern diagnostic tool that produce high-quality three-dimensional and cross-sectional images of the body by combining computer reconstruction with X-ray beams that spin around the patient [7]. The manual analysis of CT images is possible only for trained physicians. The shortage of physicians may delay the analysis of diseases. Furthermore, it lacks specificity and could produce a false positive result due to the fact that viral or benign inflammatory lesions can mimic cancer. Machine Learning (ML) based models require more preprocessing work for analysing CT images. Also, ML-



based models cannot process large-scale images. Lately, DL has become more popular for predicting lung diseases with high accuracy [8]. DL allows models with a multi-layered neural network design to identify complicated patterns from a large quantity of lung diseased images [9]. Convolutional Neural Networks (CNNs), a type of DL, were developed to accurately classify diseases by extracting spatial features from medical images, including CT scans [10]. Vision Transformers (ViTs), Autoencoders, Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNNs) are another type of DL models that capture features from various types of diseased and normal images. The captured features are then represented compactly to learn global relationships among diseases and image features. They also learn long-range dependent features for specific diseases through various attention mechanisms [11, 12].

Early identification of lung diseases, classification accuracy and clinical decision support have all been improved by CT-based lung disease prediction systems by using DL based models [13–16]. However, modern approaches often fall short in accurately depicting the fine-spatial characteristics of the lungs' CTs, which can cause problems with segmentation of small or low-contrast lesions and incorrect boundary delineation. The complex anatomy and subtle changes in contrast and brightness in CT images further add to the difficulty of segmentation. Additionally, some models are more susceptible to having feature representations destroyed by acquisition-related artifacts and image noise. Because of these limitations, the generalization of the model is affected by images captured under various clinical settings. Consequently, improving contextual modelling while preserving spatial precision remains a critical challenge in lung disease prediction systems.

In this paper, the LRMNet model is developed to overcome the above-mentioned limitations for an effective prediction of lung diseases. In order to enhance its ability for generalizing, pre-processing and data enrichment approaches based on DCGAN are employed. Subsequently, an SVU-M model is built to capture the fine spatial detail along with the long-distance contextual information. SVU-M adds blocks of the model of SVSS to model global dependencies efficiently while guaranteeing that the increase in computational cost will be proportional to the input size.

The fusion of SVSS within SVU-M increases the delineation of the boundaries and the segmentation of small or low-contrast areas, which leads to a higher robustness against imaging artifacts. The decoder feature maps are further processed using GAP to get compact image-level feature representations. In parallel, the original CT image is fed into the ResNet 50 network for extracting comprehensive semantic features. At last, an MLP classifier is trained with all the features to reliably predict Lung illness. The remaining document is organised as follows: Existing DL-based

algorithms for predicting lung diseases are reviewed in Section 2. The suggested methodology is thoroughly described in Section 3. The outcomes are shown and discussed in Section 4. Finally, the study is summarized in Section 5.

2. Literature Survey

A DL framework named LDDNet was developed for the diagnosis of lung diseases [17]. To enhance the model's capability during classification, GAP, dense layers, batch normalization, dropout and softmax layers were included. However, high computational complexity and limited generalizability if trained on specific or restricted datasets. Al-Sheikh et al. [18] proposed a multi-class DL algorithm for lung disease categorization using CXR and CT images. First, the images are improved using a pixel-based k-symbol Lerch transcendence function; then, a custom CNN and pre-trained techniques (AlexNet and VGG16) were used for classification. However, the model achieved lower accuracy as it was trained on limited datasets.

An Intelligent framework was developed for lung cancer detection using CT images [19]. The images were first processed for noise removal using a Butterworth smooth filter. Afterwards, the Chaotic Crow Search Algorithm and Random Forest (CCSA-RF) method was used to gather features for bi-level feature identification. Finally, Probabilistic Neural Network (PNN) with Sparse CNN (SCNN) approach performed lung cancer prediction. However, the hyperparameters were not optimized properly, which leads to lower accuracy and increased computational complexity. A U-Net-based Feature extraction model cascaded with a Graph Convolutional Network (FGCN) was developed for lung disease detection [20]. In this method, the U-Net architecture was developed for image segmentation. Then, the GCN model was applied for effective feature extraction. By integrating the U-Net with GCN, pixel-level connectivity information with neighbouring patterns was retrieved to improve the predictions of diseases. However, this model results in limited scalability and may struggle to capture complex feature variations across different lung disease patterns effectively.

A hybrid DL network model was presented for the segmentation and classification of Interstitial Lung Diseases (ILD) [21]. In this method, a multi-level enhanced U-Net++ was applied for lung segmentation. The Refined Attention Pyramid Network (RAPNet) were employed for feature extraction, while MobileUNetV3 was used for ILD classification. Due to a high-size integrated feature set, progressed computational complexity was determined. A DL multi-label classification approach was developed for lung diseases [22]. In this method, geometric image augmentation was performed for data diversity. A transfer-learning approach integrating MobileNetV1 with a three-layer deep neural network for multilabel categorisation of pulmonary diseases. However, this model struggles to capture more complex disease patterns and has lower scalability.

By utilizing a convolutional spatial transformer network, an estimative model for predicting the prognosis of pulmonary fibrosis was constructed (FibroRegNet) [23]. The model uses a regression function to map multimodal inputs, like CT scans and demographic data, to predict Forced Vital Capacity (FVC) through quadratic polynomial ridge regression. FibroRegNet calculates pulmonary volume from designated CT segments, incorporates demographic data, and integrates this data with multilayer characteristics derived via spatial transformer units in synchronised channels. However, heavily reliable on accurate CT slice selection and demographic data quality, which may affect prediction reliability if the input data is inconsistent. ReciproCAM-enhanced Lightweight two-dimensional CNN (LW-2DCNN) for efficient lung disease identification [24]. In this model, ReciproCAM spatially masks the desired layer's feature maps to separate every spatial position. Then, LW-2DCNN was performed to extract the high-level features. Softmax was used to categorize several pulmonary disorders from chest CT scans. However, restricted generalizability due to a single population and limits differential diagnosis ability.

By using a verifiable CNN augmented with the Nutcracker Optimizer method, a model for lung disease detection was constructed using EfficientNetV2 (ENV2-LDD-VCNN) [25]. Feature extraction and prediction were carried out using ENV2-VCNN, while picture scaling and pixel normalization were handled by High Accuracy Distributed Kalman Filter (HADKF) and as a last step, optimized the VCNN parameters using the Nutcracker Optimizer Algorithm (NOA). However, the computational cost and accessibility in clinical applications are both increased due to the integration of numerous sophisticated components. Classification of lung diseases using an explainable fusion-based transfer learning method was introduced [26]. To begin with, CXR and CT scan images were classified using three CNN architectures: ResNet50, ResNet101 and EfficientNetB0. Also, two hybrid designs were identified: feature and decision-level fusion. The first group combines feature representations from ResNet50 and EfficientNetB0, while the latter one integrates the prediction outputs from ResNet50 and ResNet101-however, the lower accuracy due to inefficient hyperparameter optimization.

3. Proposed Methodology

A comprehensive explanation of the planned LRMNet for the accurate prediction of lung diseases is provided in this section. To see the proposed model's structure, look at Figure 1.

3.1. Pre-Processing and Augmentation

In this stage, CT images are pre-processed on a collected image dataset to improve the image resolution and preserve the consistency of image characteristics. This is important for proposed DL models to capture quality features for classification. At first, intensity normalization is done to

minimize deviations between scanners and acquisition parameters. Noise reduction techniques such as Gaussian filtering are used to suppress imaging noise but preserve crucial imaging details with regard to their structures. Contrast enhancement techniques, such as histogram equalization, are applied to enhance the visibility of the lung tissues and lesions.

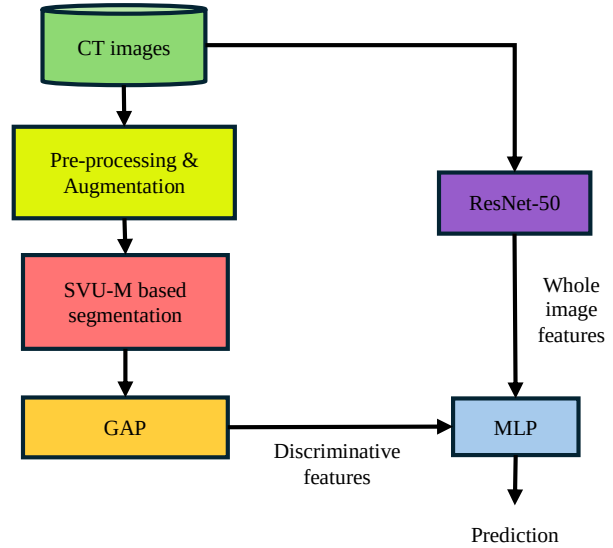


Fig. 1 Outline of the proposed model

Generative DL models like DCGAN [27] learn the distribution of the training data to generate synthetic images that appear realistic. It is mostly constituted of primary parts, and there are the Generator and the Discriminator, both of which are trained in an adversarial manner. The generator is taught to generate synthesized CT images from arbitrary noise, and the discriminator attempts to differentiate between created and real CT images. The generator gradually improves its ability to generate actual and high-quality lung CT samples using a continuous learning task. DCGAN for data augmentation. DCGAN allows for greater deviation of the data, lessens overfitting issues and enhances the model's adaptability capabilities by generating realistic and diverse synthetic lung CT images. The framework that is proposed overcomes the situation of class imbalance and limited dataset size by employing the DCGAN for data augmentation. DCGAN allows for greater representation obtained from the ResNet50 and EfficientNetB0, and decision-level fusion, which integrates the prediction results of ResNet50 and ResNet101. However, the lower accuracy due to inefficient hyperparameter optimization. This is particularly useful for medical imaging, where it can be a challenge to get large, well-defined datasets.

3.2. SVU-M based Image Segmentation

The SVU-M model segments the lungs into several regions by executing SVSS blocks into a U-shaped encoder-decoder structure. This design effectively captures both fine spatial features and long-range contextual information,

accurately defining the boundaries and enhanced separate small or low-contrast lesions.

3.2.1. Preliminaries

Recent Structured State Space Models (SSMs) include S4, i.e., Structured State Space for Sequence Modelling for Mamba, which are derived from the principles of a traditional time continuous time adaptive system. In these models, a one-dimensional input signal as $a(t) \in \mathcal{R}$ is processed through an intermediate latent state $h(t) \in \mathcal{R}^N$ which then generates an output signal $y(t) \in \mathcal{R}$. The association among the input, hidden state and output is described using a linear Ordinary Differential Equation (ODE) as depicted in Equation (2).

$$h'(t) = Xh(t) + Ya(t) \quad (1)$$

$$y(t) = Zh(t) \quad (2)$$

In this case, the variables X and Y , which are both variables in $\in \mathcal{R}^{N \times 1}$ They are accountable for connecting the input data to the hidden state and the hidden state to the outcome individually. S4 and Mamba discretize the continuous-time state space system suitable for DL applications. In this procedure, a time-step variable Δ is incorporated to transform the parameters X and Y into discrete counterparts. Generally, the zero-order discretization method is employed, as in Equations (3) and Equation (4).

$$\bar{X} = \exp(\Delta X) \quad (3)$$

$$\bar{Y} = (\Delta X)^{-1} (\exp(\Delta X) - I) \cdot \Delta Y \quad (4)$$

After discretization, the SSM can be computed either through linear recurrence as in Equation (5).

$$h'(t) = \bar{X}h(t) + \bar{Y}a(t) \quad (5)$$

$$y(t) = Zh(t) \quad (6)$$

Where, $Q \in \mathcal{R}^l$ represents a configured convolutional kernel, and l denotes the input length of sequence a .

$$\bar{Q} = (Z\bar{Y}, Z\bar{X}\bar{Y}, \dots, Z\bar{X}^{l-1}\bar{Y}) \quad (7)$$

$$y(t) = a * \bar{Q} \quad (8)$$

3.2.2. Design of SVU-M

Figure 2 illustrates the complete framework of the proposed SVU-M model. Figure 2 shows the components of the SVU-M architecture, which consist of a GAP module for feature extraction, a final projection layer with skip relations, an encoder module, a decoder module, etc. Unlike conventional symmetric U-shaped networks, an efficient

hierarchical design is adopted to balance contextual modelling and computational efficiency.

In order to avoid overlapping, the patch embedding layer divides the input CT image $a \in \mathbb{R}^{H \times W \times 3}$ into 4×4 dimensions. Each patch is then transformed into a feature depiction with dimensions Z , where Z is adjusted to be 96. This transformation generates embedded feature maps $a' \in \mathbb{R}^{\frac{H}{4} \times \frac{W}{4} \times Z}$, which are subsequently normalized before being passed to the encoder module. The encoder is organized into four layered phases with channel sizes $[C, 2C, 4C, 8C]$. Each phase contains two SVSS blocks $[2, 2, 2, 2]$ across the four stages. At the end of the first three stages, a patch merging function is performed, excluding the final phase to down-sample the spatial sizes while simultaneously increasing the channel capacity. High-level context representations and efficient multi-scale features can be extracted using this progressive down-sampling technique.

The decoder also includes four phases. To expand the spatial resolution (height and width) and decrease the number of feature channels, a patch expanding process is used at the start of the final three phases. Channel dimensions are $[8C, 4C, 2C, C]$ and there are four stages of SVSS blocks with these dimensions: $[2, 2, 2, 1]$. The feature map is up-sampled to the desired segmentation size using a Final Projection layer after the decoder. Specifically, when a 4x up-sampling is finished using patch expanding to recover the spatial resolution, a projection layer is employed to recover the required number of channels. The skip connections are created using simple element-wise addition; no additional parameters or modifications are utilized.

3.2.3. Architecture of the SVSS Block

The SVSS block is the core computing unit of the suggested SVU-M architecture. It is intended to capture long-range spatial interdependence rapidly while preserving fine-grained local structure information. Unlike traditional state space modelling, which is sequential and requires significant memory and computing overhead for extended sequences, the SVSS block employs a locally bidirectional state space formulation. Instead of performing separate global forward and backward scans on the entire sequence, SVSS includes a local backward scan in sub-sequences during the forward scanning operation. This architecture allows each spatial token to cluster contextual information in both directions via a single read/write operation. As a result, the computational cost increases linearly with input size, scaling effectively and efficiently with high-resolution medical images. Figure 3 depicts the SVSScan operation. As per the Figure 3, a global forward scan is conducted is performed as in Equation (9).

$$h_t^f = \bar{A}^f h_{t-1}^f + \bar{B}^f x_t \quad (9)$$

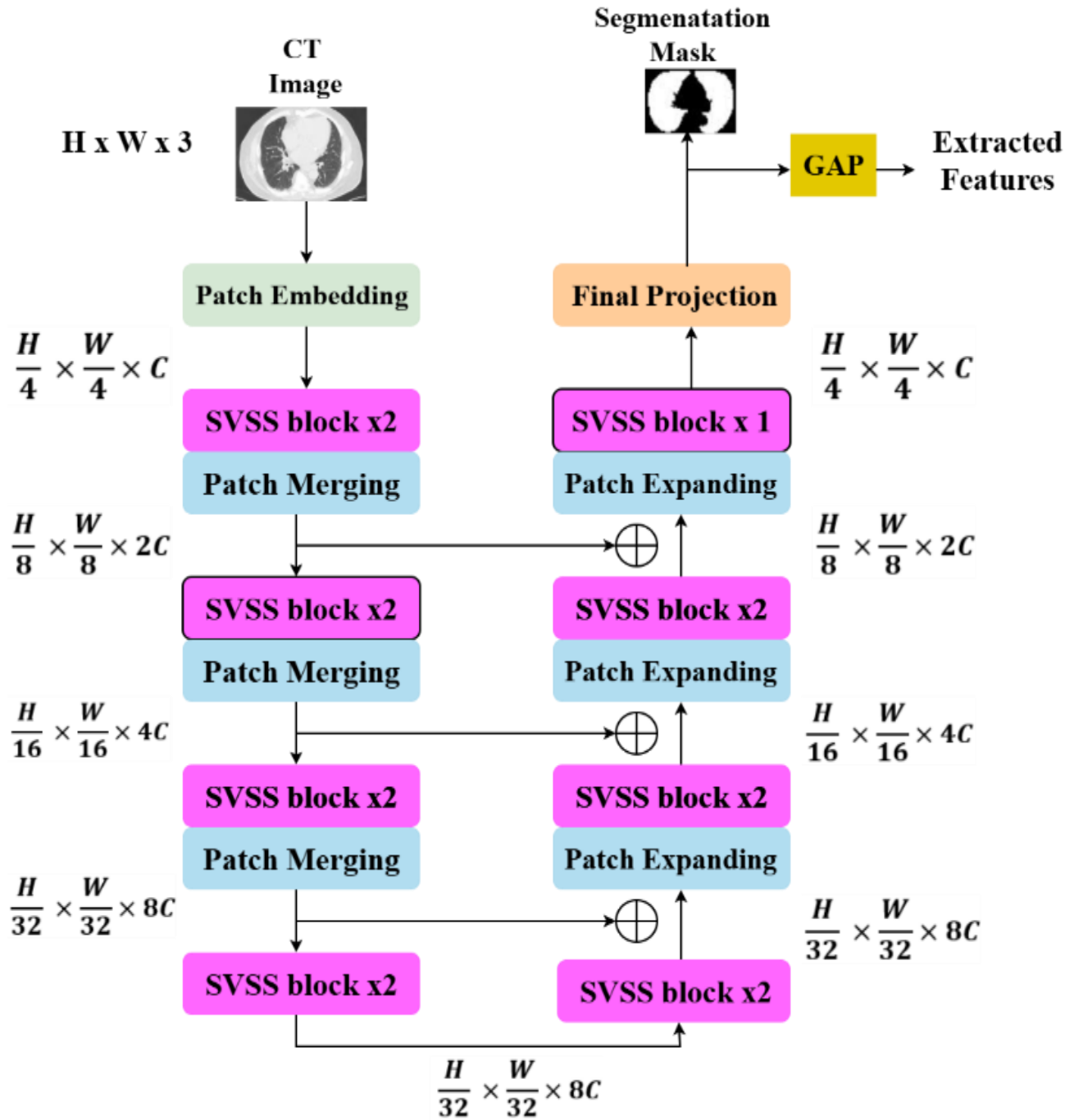


Fig. 2 Structure of SVU-M Model

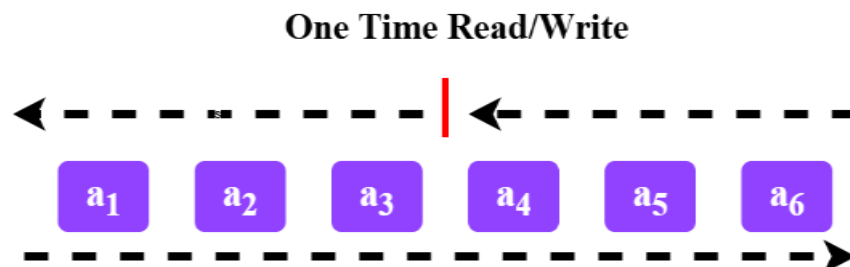


Fig. 3 SVScan operation

Where, x_t signifies the input token, h_t^f signifies the latent state at time t , $\bar{A}$ and $\bar{B}$ These are the functions that allow the SVSS to dynamically adapt to the input context. f denotes the forward scan. Then, the input sequences of length L is divide into sub-sequences which have a length of S where S is adjusted to be 3. S is set to the number of elements for one thread processes for succeeding computation. In each of these sub-sequences, a backward scan (b) as in Equation (10) is conducted.

$$h_t^b = \bar{A}^b h_{t-1}^b + \bar{B}^b x_t \quad (10)$$

Specifically, the state h_t^b obtained during the local backward scan as in Equation (11).

$$h_t^b = \begin{cases} B^f x_t & \text{if } t \% M = 0 \\ \bar{A}^f h_{t-1}^b + \bar{B}^f x_t & \text{Otherwise} \end{cases} \quad (11)$$

Since the backward scan is integrated within the forward scanning process, the same state transition parameters are reused for both directions. The backward hidden state is then combined with the forward hidden state h_t^f to obtain the final representation.

As the input projection term $B^f x_t$ appears at both forward and backward computations, it is subtracted during aggregation to avoid redundant accumulation in the summation. The final hidden state after combining the forward and backward scans is expressed as in Equation (12), and the corresponding output is computed as in Equation (13).

$$h_t = h_t^f + (h_t^b - B^f x_t) \quad (12)$$

$$y_t = C^f h_t + D^f x_t \quad (13)$$

Here, y_t is the output token, D is a parameter, C^f is the state dimension coefficient to aggregate N state dimensions into a single output.

Figure 4 illustrates the architecture of the SVSS Block in the SVU-M model. As per Figure 4, the SVSS block first applies Layer Normalization to the input feature map $a \in \mathbb{R}^{H \times W \times C}$ to stabilize feature statistics. The normalized features are projected into two parallel branches through linear projection layers.

In the first branch, a 1D convolution is applied along the flattened spatial dimension to capture local dependencies, followed by a nonlinear activation function σ . The resulting features are reshaped into a sequence as in Equation (14).

$$a_s = \{a_1, a_2, \dots, a_L\}, L = H \times W \quad (14)$$

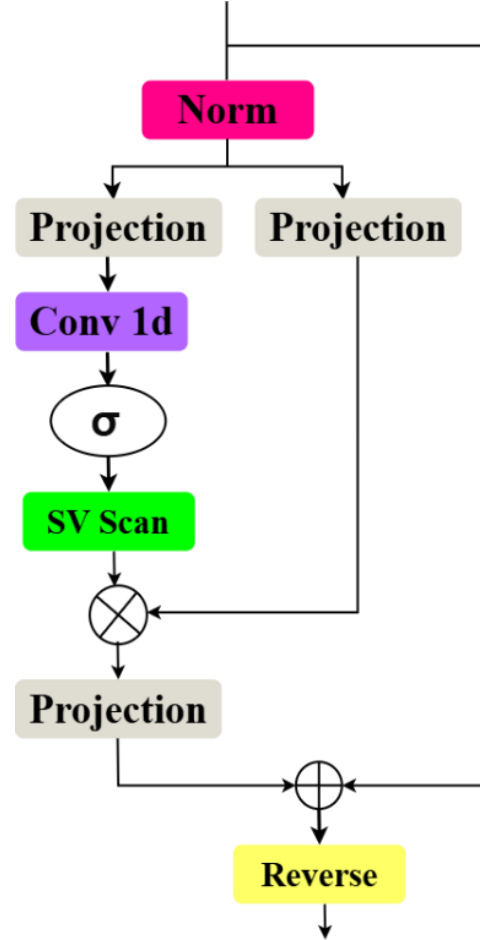


Fig. 4 Architecture of SVSS Block

Then, the SV Scan operation is performed sequentially to extend the hidden states throughout the spatial tokens. This helps every token to get relevant data from the previous positions while maintaining efficient computation.

The result obtained from the SVScan operation is then modulated through element-wise multiplication with the second projection branch, acting as a gating mechanism to enhance informative features. A final linear projection restores the channel dimension, and a reverse operation is introduced to alternate scanning direction across stages. Through these stages, the SVSS block effectively captures both fine local structures and long-range global dependencies.

3.2.4. Global Average Pooling

The SVU-M architecture produces a refined feature map $K \in \mathbb{R}^{H \times W \times C}$ that encodes segmentation-aware spatial representations of the lung CT image. These features capture multi-scale structural information, including fine boundary details, lesion localization, texture variations, intensity patterns, and long-range contextual dependencies modelled by the SVSS blocks, resulting in spatially enriched and pathology-sensitive feature representations.

These features contain multi-scale structural information extracted through hierarchical encoding and decoding, including fine boundary details, lesion localization cues, texture variations, intensity distributions, and long-range contextual dependencies modelled through the SVSS blocks. Thus, the output of SVU-M represents spatially enriched and pathology-sensitive feature embedding. To convert these spatial features into a compact image-level representation suitable for classification, GAP is applied across the spatial dimensions. GAP computes the average response of each feature channel (k) as in Equation (15).

$$K_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W k_{i,j,c} \quad (15)$$

where $k_{i,j,c}$ indicates the activation at spatial location (i, j) in the c^{th} channel. The resulting pooled feature vector is given as in Equation (16).

$$K_{gap} = [k_1, k_2, \dots, k_C] \in \mathbb{R}^C \quad (16)$$

This operation performs spatial aggregation by computing the channel-wise mean response across all spatial locations, thereby transforming the feature map $K \in \mathbb{R}^{H \times W \times C}$ into a compact vector $K_{gap} \in \mathbb{R}^C$. By averaging spatial activations, GAP retains the dominant channel-wise activation strength while eliminating spatial redundancy.

The resulting feature vector encodes globally aggregated segmentation-aware features, including structural, lesion-specific, textural, and contextual representations modelled by the SVSS blocks and serves as input to the classification module for lung disease prediction.

3.3. Whole Image Feature Extraction using ResNet-50

In this framework, ResNet-50 is utilized as a deep multi-level feature extractor network to derive comprehensive global semantic information from the complete CT images. The refined CT images with dimensions $H \times W \times 3$ are initially processed by a 7×7 convolution layer containing 64 filters and a stride of 2.

This large-kernel convolution performs spatial filtering to identify fundamental visual features like edges, texture patterns, borders and variations in intensity across the lung region.

Figure 5 depicts the structure of the ResNet 50 model for whole image feature extraction. The ReLU helps to improve the non-linearity in feature learning, and batch normalization aids in stabilizing the learnt feature distributions. In order to preserve the most instructive structural responses for subsequent processing, a subsequent max pooling minimizes spatial dimensions while maintaining dominating activations. Four hierarchical residual stages (Conv2× to Conv5×) are used to improve the extracted low-level characteristics.

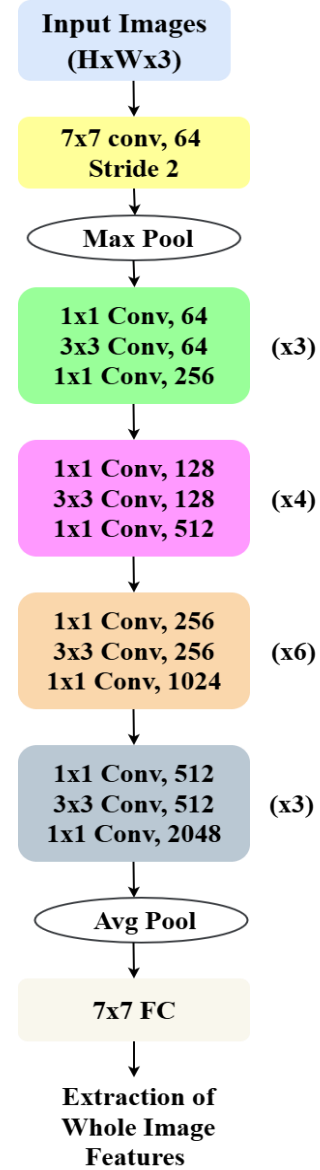


Fig. 5 Structure of ResNet50

Bottleneck blocks in Conv2× lower the feature representation from 64 channels to 256 channels by containing 1×1 , 3×3 , and 1×1 convolutions. While the final 1×1 convolution broadens the channel space to improve representational richness, the first 1×1 convolution shrinks the feature dimensions to guarantee greater computing efficiency, and the 3×3 convolution captures localized patterns of space and structure. The spatial resolution progressively drops, and the channel depth rises to 512 and 1024, respectively, as it moves to Conv3× and Conv4×. By merging data from wider receptive fields, the model is able to recover increasingly more abstract and discriminative features due to the progressive spatial compression and channel expansion. Consequently, systematic representations of anatomical configurations and spatial connections between lung regions are used to construct intermediate features from local texture descriptors.

In order to produce extremely abstract feature maps with intricate semantic information regarding the inputs from nearly the whole CT images, Conv5x further deepens the channel in the last stage, reaching 2048. The 2048-channel feature maps are then subjected to GAP, which uses spatial response aggregations to produce a compact 2048-dimensional feature vector. The integrated patterns of tissue density, structural alignment, lesion distribution, and pathologic changes are represented by the characteristics at this depth. Global anatomical structure, spatial interactions between lung regions, texture characteristics, intensity distributions, and high-level disease signs derived from every lung part throughout the entire CT scan are all included in the collective features of these full image features. This vector is a highly discriminative global descriptor that can provide a powerful representation to follow the classification of lung diseases and efficiently summarize the derived hierarchical features of the entire CT images.

3.4. MLP-based Classification

After feature extraction from SVU-M (discriminative features) and ResNe-50 (whole image features), two complementary feature vectors are obtained. The SVU-M branch produces a segmentation-aware feature vector as in Equation (17).

$$f_{SVU-M} = R^{C_1} \quad (17)$$

Simultaneously, the ResNet-50 branch extracts a global semantic feature vector (whole image features) as in Equation (18).

$$f_{ResNet-50} = R^{C_2} \quad (18)$$

These two feature vectors are fused through concatenation to form a unified representation as in Equation (19).

$$f_{fusion} = [f_{SVU-M} || f_{ResNet-50}] \in R^{C_1 + C_2} \quad (19)$$

Where || denotes channel-wise concatenation. As a next step, an MLP classifier with nonlinear activation functions is fed the combined feature vector. It is shown in Equation (20) that the MLP carries out a nonlinear transformation.

$$z = \phi(W_1 f_{fusion} + B_1) \quad (20)$$

$$y = W_2 z + B_2 \quad (21)$$

Where, W_1, W_2 and B_1, B_2 imply learnable weights and biases, $\phi(\cdot)$ represents a complex activation operation like ReLU. A softmax function is then used to calculate the class probabilities, as shown in Equation (22).

$$\hat{y} = Softmax(y) \quad (22)$$

The predicted class corresponds to the maximum probability in $\hat{y}$ enabling accurate lung disease classification.

Algorithm: LRMNet-based lung disease prediction

Input: CT image dataset $I = \{I_1, I_2, \dots, I_n\}$

Output: Predicted lung disease class $\hat{y}$

- 1) Acquire the lung disease-based CT images
 \\Pre-processing and Data Augmentations
- 2) Normalize image intensities.
- 3) Apply noise reduction using Gaussian filtering.
- 4) Perform contrast enhancement by Histogram Equalization.
- 5) Train DCGAN on pre-processed CT images.
- 6) Generate synthetic CT images to balance the dataset.
- 7) Combine real and generated images to form an augmented dataset.
 \\SVU-M Based Segmentation
- 8) Input the augmented CT image into the Patch Embedding layer.
- 9) Extract hierarchical features using an encoder with SVSS blocks.
- 10) Apply patch merging for down-sampling.
- 11) Decode features using patch expanding.
- 12) Generate segmentation-aware feature map.
- 13) Apply GAP to obtain discriminative features
 \\ResNet-50 Based Whole Image Feature Extraction
- 14) Input the original CT image into ResNet-50
- 15) Pass the image through the initial 7×7 convolution layer, max-pooling layer, and subsequent residual bottleneck blocks (Conv2x to Conv5x) to progressively extract hierarchical features.
- 16) Reduce the spatial dimensions while increasing channel depth from 64 to 2048 through stacked residual learning.
- 17) Apply Average Pooling on the final 2048-channel feature map to obtain a compact global semantic feature vector
 \\MLP Classification
- 18) Concatenate extracted features from both SVU-M of GAP (discriminative features) and ResNet 50 (whole image features)
- 19) Feed the fused feature vector into a fully dense layer to perform feature transformation.
- 20) Pass the transformed features into the final output layer
- 21) Class probabilities can be obtained by using the Softmax function.
- 22) Return predicted lung disease class. $\hat{y}$.

4. Results and Discussion

4.1. Dataset Description

Two datasets are used for experimentation.

1. Customized Dataset: The CT scans identified five different diseases, including pneumonia, atelectasis, infiltration, COVID-19 and non-COVID. In this study, we evaluate a number of open-source portals that can be

used to collect CT data for the purpose of the experiment. Images of lung atelectasis are obtained from [28], Covid and Non-Covid (Normal) images are collected from [29], viral pneumonia is received from [30], and infiltration is gained from [31]. Table 1 lists all of the CT scans that were taken from the dataset provided for the performance assessment. The customized CT dataset consists of five lung disease categories, including COVID-19, Pneumonia, Normal, Infiltrate, and Atelectasis. The first step was the collection of 4,318 CT images from various public sources. To address class imbalance and increase dataset size, data augmentation was performed using DCGAN, resulting in 8,810 images (1762 per class) as shown in Table 1. There were 6,167 training images and 2,645 testing images overall after the augmented dataset was split into 70% for training (1,233 images per class) and 30% for testing (529 images per class).

2. SARS-CoV-2 CT Dataset: In this dataset, 2482 CT scans are included; 1252 of these scans show evidence of SARS-CoV-2 infection (COVID-19), while 1230 scans show no infection at all [32]. Patients in hospitals in São Paulo, Brazil, provided the data used to compile these statistics. With this dataset, we aimed to promote the investigation and advancement of AI methodologies that can use CT scans to diagnose SARS-CoV-2 infection.

Table 1. Customized dataset

Lung Disease Categorizes	Number of CT images considered	Images after augmentation
Covid-19	1002	1762
Pneumonia	1762	1762
Normal	984	1762
Infiltrate	260	1762
Atelectasis	310	1762

Table 2. SARS-CoV-2 CT dataset

Lung Disease Categorizes	No. of. CT images Considered
Covid-19	1252
Non-Covid-19	1230

Table 2 lists all of the CT images that were taken from the dataset provided for the performance evaluation. Data augmentation was not required for the SARS-CoV-2 CT dataset because it is already well-balanced with 1,252 COVID-19 and 1,230 Non-COVID images (2,482 total CT images). There is no need for synthetic data expansion because the dataset is balanced and adequate for binary classification, reducing the risk of overfitting. The SARS-CoV-2 dataset was split into two parts: one for COVID-19 cases, which included 876 training images, and another for non-COVID cases, which included 861 training images and 369 testing images. A visual representation of the performance measures used to assess the suggested LRMNet model can be found below.

4.2. Experimental Configuration and Evaluation Metrics

The current and suggested models are administered on a Windows 11 system using a 64-bit Windows 10 laptop with

an Intel® Core™ i5-4210 CPU @ 3 GHz, 4 GB RAM, and a 1 TB HDD. PyTorch 2.2 and TorchVision 0.17 were used as the main DL frameworks in the Python 3.10 software environment. Table 3 lists the suggested model's parameters.

1. Accuracy: It is the ratio of the number of instances properly classified into each category of lung illness to the overall number of instances tested.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (23)$$

In Equation (23), the result that the model correctly identifies the categories of lung diseases as itself is known as a True Positive (TP); for example, Infiltrate is categorized as Infiltrate. When the model classifier correctly identifies pneumonia as pneumonia, the result is known as a true negative (TN). When a lung disease (Covid-19\Pneumonia\Normal\Infiltrate\Atelectasis) is incorrectly classified by the model as Atelectasis, it is referred to as a False Positive (FP). The result that occurs when the model incorrectly characterizes the pneumonia as normal is known as a False Negative (FN).

2. Precision: It is the proportion of correctly classified lung disease examples at TP and FP rates.

$$Precision = \frac{TP}{TP+FP} \quad (24)$$

3. Recall: It is the ratio of exactly classified categories of lung disease cases at TP and FN rates.

Table 3. Parameter Settings of the LRMNet Framework

Module	Parameter	Value
Pre-processing	Input Image Size	256 × 256
	Normalization	Min-Max scaling [0,1]
	Contrast Enhancement	CLAHE (Clip Limit = 2.0)
	Noise Reduction	Gaussian Filter (3 × 3)
SVU-M	Patch Size	4 × 4
	Embedding Dimension (C)	96
	Encoder Stages	4
	SVSS Blocks (Encoder)	[2, 2, 2, 2]
	Encoder Channels	[96, 192, 384, 768]
	Decoder SVSS Blocks	[2, 2, 2, 1]
	Decoder Channels	[768, 384, 192, 96]
	Activation Function	GELU
	Normalization	Layer Normalization
	SVSS-State Dimension (N)	16 / 32
	SVSS-Sub-sequence Length (S)	3
	Output Feature Size	2048

ResNet-50	Fine-tuning	Last 2 Residual Blocks
MLP Classifier	Hidden Layer Size	512
	Activation	ReLU
	Final Activation	Softmax
Training Settings	Learning Rate	1e-4
	Optimizer	AdamW
	Dropout	0.5
	Weight Decay	1e-5
	Batch Size	16
	Epochs	100
	Early Stopping	Patience = 10
	Loss Function	Categorical Cross-Entropy

$$Recall = \frac{TP}{TP + FN} \quad (25)$$

4. F1-Score: It is the average of recall and precision, with 1 being the best and 0 the worst.

$$F1 - score = 2 \times \frac{Precision \cdot Recall}{Precision + Recall} \quad (26)$$

4.3. Analysis of Proposed and Existing Models on Customized Dataset

Here, compare the proposed LRMNet model against other models, using the customized dataset from Section 4.1, and evaluate their performance using various metrics. These models include LDDNet [17], FGCN [19], LW-2DCNN [24], and EN-LDD-VCNN [25]. Figure 6 shows the confusion matrix on the customized dataset. Figure 7 displays the accuracy (in %) achieved by the LRMNet model and existing models LDDNet, FGCN, LW-2DCNN and EN-LDD-VCNN for diagnosing various lung disease categories such as Covid-19, pneumonia, normal, infiltrate and atelectasis using CT images. According to the results of the previous and proposed models, LRMNet outperforms other classification models in terms of accuracy for each category of lung diseases. For example, in the categorization of Covid-19, the accuracy of LRMNet is 14.16%, 10.25%, 7.56%, and 3.92%, higher than that of existing models, respectively.

		Customized Dataset				
Actual Label	Covid-19	496	14	6	8	5
	Pneumonia	2	502	6	10	9
	Normal	6	6	501	9	7
	Infiltrate	5	6	9	494	15
	Atelectasis	7	9	9	7	497
		Predicted Label				
		Covid-19	Pneumonia	Normal	Infiltrate	Atelectasis

Fig. 6 Confusion matrix for the proposed model on the customized dataset

Figure 8 displays the precision (in %) achieved by the LRMNet model and existing models like LDDNet, FGCN, LW-2DCNN and EN-LDD-VCNN for diagnosing various lung disease categories such as pneumonia, Covid-19, normal, atelectasis and infiltrate using CT images. Based on the results, LRMNet is the most accurate classification model for all types of lung diseases. For instance, LRMNet outperforms current models in precision for pneumonia classification by 18.62%, 13.80%, 9.94%, and 6.55%, respectively.

Figure 9 displays the recall (in %) achieved by the LRMNet model and existing models like LDDNet, FGCN, LW-2DCNN and EN-LDD-VCNN for diagnosing various lung disease categories such as pneumonia, Covid-19, normal, atelectasis and infiltrate using CT images. Based on the results of the previous and proposed models in Figure 9, LRMNet emerges as the best classification model for lung diseases. For example, LRMNet improves recall by 16.25%, 12.70%, 9.59%, and 6.96% over the existing models LDDNet, FGCN, LW-2DCNN and EN-LDD-VCNN in the Normal category, respectively.

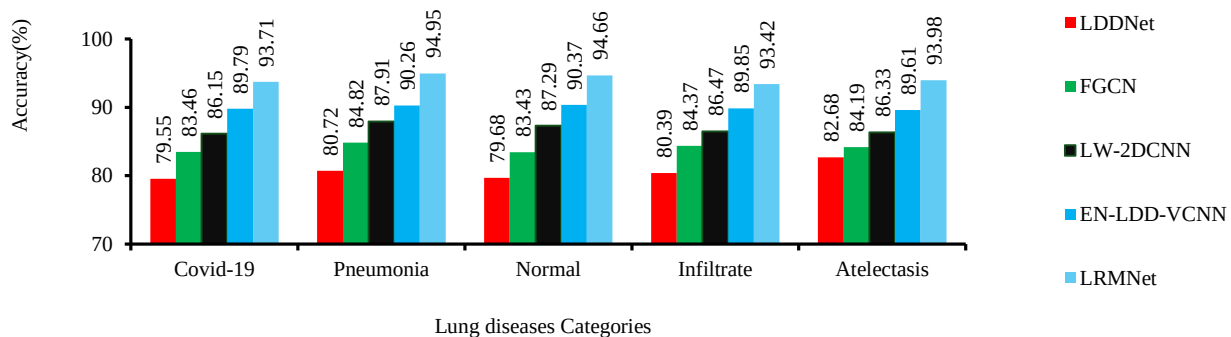


Fig. 7 Accuracy comparison for the lung disease category on the customized dataset

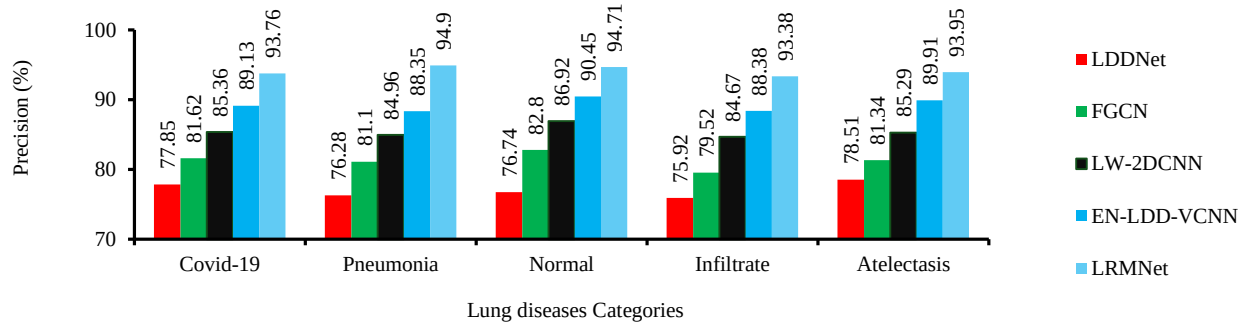


Fig. 8 Precision comparison for the lung disease category on the customized dataset

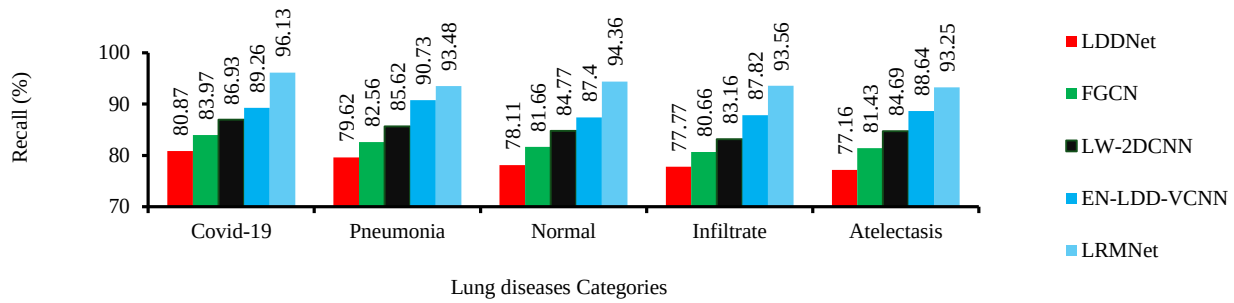


Fig. 9 Recall comparison for the lung disease category on the customized dataset

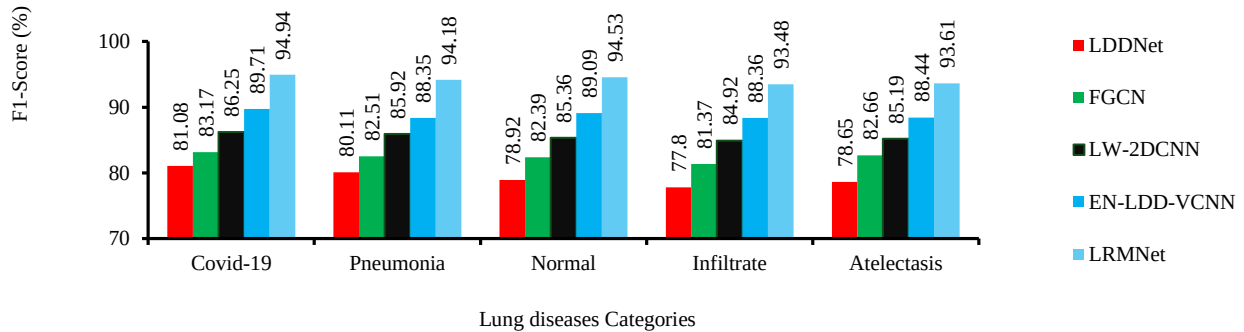


Fig. 10 F1-Score comparison for lung disease category on customized dataset

Table 4. Overall result of the customized dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
LDDNet	80.60	77.06	78.70	79.31
FGCN	84.05	81.27	82.05	82.42
LW-2DCNN	86.83	85.44	85.03	85.52
EN-LDD-VCNN	89.9	89.24	88.77	88.79
LRMNet	94.15	94.14	94.16	94.15

Figure 10 displays the F1-score achieved by the LRMNet model for Infiltrate is 15.68%, 12.11%, 8.56% and 5.12% higher than the existing models LDDNet, FGCN, LW-2DCNN and EN-LDD-VCNN, respectively.

The overall results of diagnosing various lung disease categories, such as pneumonia, Covid-19, normal, atelectasis and infiltrate using CT images, are listed in Table 4 for the customized dataset.

4.4. Analysis of Proposed and Existing Models on SARS-CoV-2 CT Dataset

Here, compared the proposed LRMNet model to other models using the SARS-CoV-2 CT dataset given in Section 4.1 and evaluated using several performance measures. These models include LDDNet [17], FGCN [19], LW-2DCNN [24], and EN-LDD-VCNN [25]. Figure 11 shows the confusion matrix on the SARS-CoV-2 CT Dataset. Figures 12 and 13 display the performances achieved by the LRMNet model and existing models, LDDNet, FGCN, LW-2DCNN and EN-LDD-VCNN on the SARS-CoV-2 CT Dataset using CT images, respectively. The results show that LRMNet outperforms existing classification models for both COVID-19 and non-COVID-19 images. For Covid-19, LRMNet outperforms previous models, yielding improvements of 12.43%, 9.66%, 6.78% and 3.47% in accuracy; 11.91%, 9.17%, 6.28% and 3.22% in precision; 11.70%, 8.76%, 5.84% and 2.67% in recall; and 11.89%, 9.02%, 6.06% and 2.9% in F1-score, respectively. For non-Covid-19 category, LRMNet outperforms competing models, yielding improvements of 11.04%, 8.31%, 5.57% and 2.49% in accuracy; 11.55%, 8.67%, 5.73% and 2.62% in precision; 11.7%, 8.81%, 5.99%

and 2.75% in recall; and 11.75%, 8.87%, 5.99%, and 2.86% in F1-score, respectively. Table 5 presents the overall results on SARS-CoV-2 CT datasets.

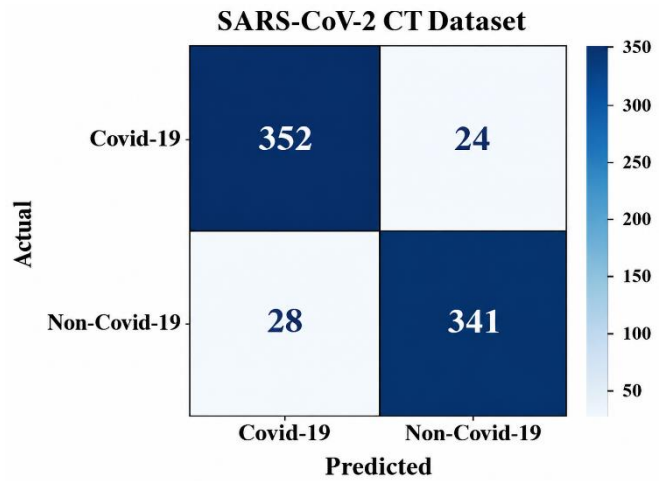


Fig. 11 Confusion matrix for the proposed model on the SARS-CoV-2 CT dataset

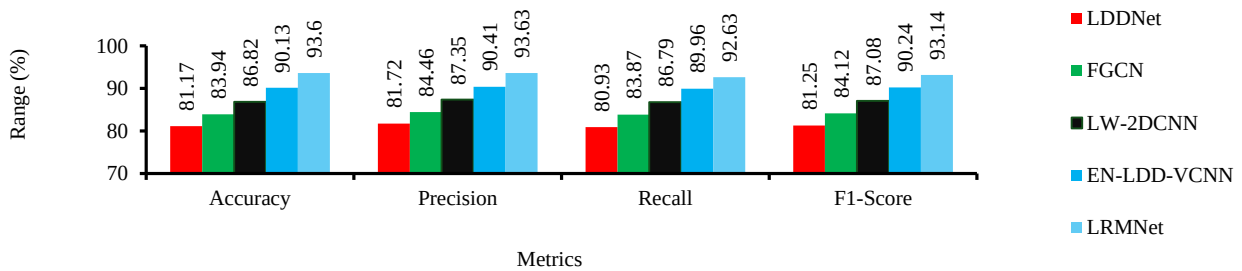


Fig. 12 Comparison of the SARS-CoV-2 CT dataset for COVID-19 images

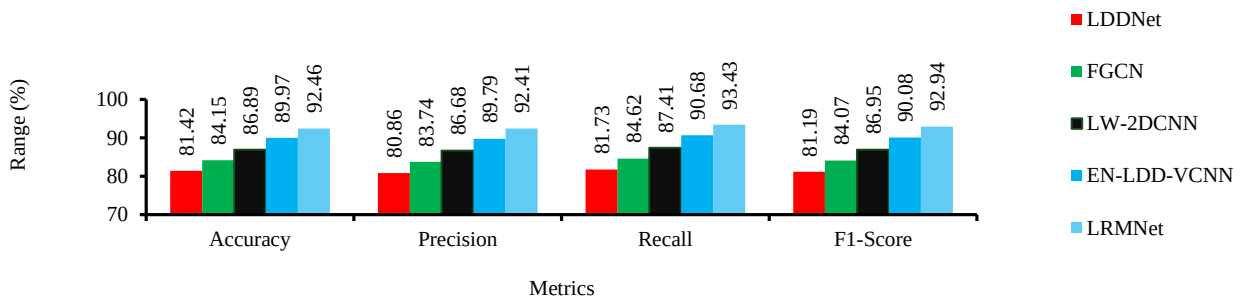


Fig. 13 Comparison on SARS-CoV-2 CT dataset for Non-Covid-19 Images

Table 5. Overall result of the SARS-CoV-2 CT dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
LDDNet	81.29	82.29	81.33	81.22
FGCN	84.04	84.10	84.24	84.09
LW-2DCNN	86.85	87.01	87.10	87.01
EN-LDD-VCNN	90.05	90.10	90.32	90.16
LRMNet	93.03	93.02	93.03	93.04

Table 6. Ablation Study for the proposed model

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Customized Dataset				
ResNet-50 Only	88.46	87.12	87.34	87.23
SVU-M Only	90.18	88.97	89.14	89.05
SVU-M + ResNet-50	91.89	90.76	90.91	90.83
SVU-M + MLP	92.43	91.25	91.38	91.31
ResNet-50 + MLP	90.72	89.56	89.71	89.63
SVU-M + ResNet-50 + MLP (LRMNet)	94.15	92.71	92.60	92.61
SARS-CoV-2 CT Dataset				
ResNet-50 Only	89.17	88.94	89.31	89.12
SVU-M Only	90.48	90.26	90.61	90.43
SVU-M + ResNet-50	91.86	91.64	92.03	91.83
SVU-M + MLP	92.34	92.11	92.50	92.30
ResNet-50 + MLP	91.09	90.85	91.26	91.05
SVU-M + ResNet-50 + MLP (LRMNet)	93.29	93.28	93.29	93.28

4.5. Ablation Study

Ablation studies using the Customized and SARS-CoV-2 CT datasets are demonstrated in Table 6 for the proposed techniques. Both the bespoke dataset and the SARS-CoV-2 CT dataset show that LRMNet accomplishes the effective accuracy of the models tested. The results show that ResNet-50 and the SVU-M achieve a relatively good recognition performance, but combining both models gives a significant result in the classification accuracy.

Similarly, SVU-M + MLP and ResNet-50 + MLP achieve higher result that solely ResNet-50 and the SVU-M models. The proposed LRMNet integrated SVU-M + ResNet-50 + MLP, which achieves better performance than the individual and hybrid models, highlights the integration of pairwise segmentation-aware and global feature representations.

4.6. Discussion

The methods LDDNet [17], FGCN [19], LW-2DCNN [24], and EN-LDD-VCNN [25] discussed in the literature survey are reimplemented for the parameter Table 3. The results obtained for both datasets of these models are compared with the proposed work in sections 4.3 and 4.4. From this analysis, it is found that the proposed model performs better than other models for both datasets. This is due to the augmentation of required images using the DCGAN model. UNet-Mamba segment the images, which helps to get even boundary pixels of diseased regions effectively. The enhancements in the visual state space of Mamba, aimed at improving scalability through the SVScan operation, are primarily responsible for the effective acquisition of detailed

spatial information with extensive contextual global dependencies, all while preserving linear computing complexity.

5. Conclusion

An effective method for predicting lung diseases is presented in this research, called LRMNet. In this framework, the SVU-M model is designed to capture long-range context and fine spatial details, effectively modelling global dependencies with linear computational cost using SVSS blocks. It enhances border delineation and small region segmentation, improving performance against imaging artefacts within a U-shaped framework. GAP processes SVU-M decoder feature maps for image-level feature representation used in classification.

Simultaneously, the input image is fed directly into ResNet-50 in order to capture all semantic information. The generated features are aggregated and fed into an MLP classifier for lung disease detection. At 94.15% on the specialized dataset and 93.29% on the SARS-CoV-2 CT dataset, the suggested technique demonstrates superior performance compared to other current models.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Funding Statement

Funding Statement: The research work and publication have not been funded by any organization or agency.

References

- [1] Wei Shi, Saverio Bellusci, and David Warburton, "Lung Development and Adult Lung Diseases," *Chest Journal*, vol. 132, no. 2, pp. 651-656, 2007. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [2] Stephanie M. Levine, and Darcy D. Marciniuk, "Global Impact of Respiratory Disease: What can we do, Together, to make a Difference?," *Chest Journal*, vol. 161, no. 5, pp. 1153-1154, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [3] Ying Zhai et al., “Global, Regional, and National Burden of Chronic Respiratory Diseases, 1990–2021 and Predictions to 2035: Analysis of Data from the Global Burden of Disease Study 2021,” *Annals of Medicine*, vol. 57, no. 1, pp. 1-25, 2025. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] William Cookson et al., “A Pandemic Lesson for Global Lung Diseases: Exacerbations are Preventable,” *American Journal of Respiratory and Critical Care Medicine*, vol. 205, no. 11, pp. 1271-1280, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [5] Sunil Kumar et al., “A Methodical Exploration of Imaging Modalities from Dataset to Detection through Machine Learning Paradigms in Prominent Lung Disease Diagnosis: A Review,” *BMC Medical Imaging*, vol. 24, no. 1, pp. 1-42, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [6] Shri Ramtej Kondamuri et al., “Chest CT Image based Lung Disease Classification–A Review,” *Current Medical Imaging*, vol. 20, no. 1, pp. 1-14, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [7] Khamidov Obid Abdurakhmanovich et al., “Respiratory Computed Tomography,” *Central Asian Journal of Medical and Natural Science*, vol. 2, no. 2, pp. 1-8, 2021. [[Publisher Link](#)]
- [8] Abobaker Mohammed Qasem Farhan, Shangming Yang, and Mohammed Al-Nehari, “Automatic Classification of Lung Diseases from CT Images,” *World Academy of Science, Engineering and Technology, International Journal of Health and Medical Engineering*, vol. 17, no. 4, pp. 87-91, 2023. [[Publisher Link](#)]
- [9] Stefanus Tao Hwa Kieu et al., “A Survey of Deep Learning for Lung Disease Detection on Medical Images: State-of-the-Art, Taxonomy, Issues and Future Directions,” *Journal of Imaging*, vol. 6, no. 12, pp. 1-38, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] M. Jasmine Pemeena Priyadarsini et al., “Lung Diseases Detection using various Deep Learning Algorithms,” *Journal of Healthcare Engineering*, vol. 2023, no. 1, pp. 1-13, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Min Hong et al., “Multi-Class Classification of Lung Diseases using CNN Models,” *Applied Sciences*, vol. 11, no. 19, pp. 1-17, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Poonam Devi, Shrdha Sagar, and Harish Rohil, “Prediction of Lung Disease using Machine and Deep Learning Techniques: A Review,” *International Journal of Health Sciences*, vol. 6, no. S2, pp. 7583-7601, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] Andreas Christe et al., “Computer-Aided Diagnosis of Pulmonary Fibrosis using Deep Learning and CT Images,” *Investigative Radiology*, vol. 54, no. 10, pp. 627-632, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [14] Joaee Choe et al., “Content-based Image Retrieval by using Deep Learning for Interstitial Lung Disease Diagnosis with Chest CT,” *Radiology*, vol. 302, no. 1, pp. 187-197, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [15] Hassaan Malik et al., “Deep Learning-based Classification of Chest Diseases using X-Rays, CT Scans, and Cough Sound Images,” *Diagnostics*, vol. 13, no. 17, pp. 1-33, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [16] Shimpy Goyal, and Rajiv Singh, “Detection and Classification of Lung Diseases for Pneumonia and Covid-19 using Machine and Deep Learning Techniques,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 4, pp. 3239-3259, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [17] Prajoy Podder et al., “Lddnet: A Deep Learning Framework for the Diagnosis of Infectious Lung Diseases,” *Sensors*, vol. 23, no. 1, pp. 1-26, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [18] Mona Hmoud Al-Sheikh et al., “Multi-Class Deep Learning Architecture for Classifying Lung Diseases from Chest X-Ray and CT Images,” *Scientific Reports*, vol. 13, no. 1, pp. 1-14, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [19] Nasr Y. Gharaibeh et al., “Automated Lung Cancer Diagnosis Applying Butterworth Filtering, Bi-Level Feature Extraction, and Sparse Convolutional Neural Network to LUNA 16 CT Images,” *Journal of Imaging*, vol. 10, no. 7, pp. 1-31, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [20] Pshtiwan Qader Rashid, and Iker Türker, “Lung Disease Detection using U-Net Feature Extractor Cascaded by Graph Convolutional Network,” *Diagnostics*, vol. 14, no. 12, pp. 1-14, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [21] Surendra Reddy Vinta et al., “Segmentation and Classification of Interstitial Lung Diseases based on Hybrid Deep Learning Network Model,” *IEEE Access*, vol. 12, pp. 50444-50458, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [22] Muhammad Irtaza et al., “Multi-Label Classification of Lung Diseases using Deep Learning,” *IEEE Access*, vol. 12, pp. 124062-124080, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [23] Pardhasaradhi Mittapalli, and V. Thanikaiselvan, “FibroRegNet: A Regression Framework for the Pulmonary Fibrosis Prognosis Prediction using a Convolutional Spatial Transformer Network,” *IEEE Access*, vol. 12, pp. 131928-131941, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [24] Hoan-Suk Choi, and Jinhong Yang, “High-Performance Lung Disease Identification and Explanation using a ReciproCAM-Enhanced Lightweight Convolutional Neural Network,” *IEEE Access*, vol. 13, pp. 84954-84965, 2025. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [25] C. Shyamala Kumari, and K. Seethalakshmi, “Efficient NetV2 for Lung Disease Diagnosis and Treatment by Harnessing the Synergy of Verifiable Convolutional Neural Network optimized with Nutcracker Optimizer Algorithm,” *Biomedical Signal Processing and Control*, vol. 113, pp. 1-13, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [26] Aicha Akrou, Amira Echtioui, and Mohamed Ghorbel, "Explainable Fusion-based Transfer Learning for Lung Diseases Classification," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 15, no. 1, pp. 1-31, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [27] John Jenkins, and Kaushik Roy, "Exploring Deep Convolutional Generative Adversarial Networks (DCGAN) in Biometric Systems: A Survey Study," *Discover Artificial Intelligence*, vol. 4, no. 1, pp. 1-22, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [28] Weerakkody Y, Silverstone L, and Walizai T, "Lung Atelectasis," *Radiopaedia*, 2025. [[CrossRef](#)] [[Publisher Link](#)]
- [29] Mehrad Aria, COVID-19 Lung CT Scans Dataset, Kaggle. [Online]. Available: <https://www.kaggle.com/datasets/mehradaria/covid19-lung-ct-scans>
- [30] Weerakkody Y, Elfeky M, and Bell D, "Viral Respiratory Tract Infection," *Radiopaedia*, 2025. [[CrossRef](#)] [[Publisher Link](#)]
- [31] Udithamala Priyanthi Ratnayak, Samitha - Lung Infiltrates, Radiopaedia, 2026. [Online]. Available: <https://radiopaedia.org/playlists/41156?lang=us>
- [32] PlamenEduardo, SARS-CoV-2 CT Scan Dataset, Kaggle, 2020. [Online]. Available: <https://www.kaggle.com/datasets/plameneduardo/sarscov2-ctscan-dataset>