INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

# Multi-Scale Multi-Class Generative Adversarial Network for Improving Deep Learner based Lung Disease Detection

**T. Kousiga**[1]*, **P. Nithya**[2]

**1** Research Scholar, Department of Computer Science, PSG college of Arts and Science, Coimbatore – 641014, Tamil Nadu, India
**2** Associate Professor and Head, Department of Networking and Mobile Application, PSG College of Arts and Science, Coimbatore - 641014, Tamil Nadu, India

OPEN ACCESS

**Received:** 14/01/2025

**Accepted:** 10/03/2025

**Published:** 23/04/2025

**Citation:** Kousiga T, Nithya P (2025) Multi-Scale Multi-Class Generative Adversarial Network for Improving Deep Learner based Lung Disease Detection. Indian Journal of Science and Technology 18(14): 1105-1117. https://doi.org/10.17485/IJST/v18i14.93

* **Corresponding author**.

kousiga2318phd@gmail.com

**Funding:** None

**Competing Interests:** None



Published By Indian Society for Education and Environment (iSee)

**ISSN**
Print: 0974-6846
Electronic: 0974-5645

## Abstract

**Objective:** The prediction results of Deep Learning (DL) model rely heavily on labeled training images and the quality of images in the datasets. The most of datasets consists of blurry and low quality images due to reconstruction and compressions processes. The main objective of this research is to generate quality and sufficient images for improving lung disease prediction of CovLscan. **Methods:** Radiopaedia-Chest X-Ray and Kaggle-COVID-19 Lung CT datasets are used for lung disease prediction. Multi-Scale Multi-Class Generative Adversarial Network (MsMcGAN) model is proposed for generating quality images to improve lung disease prediction accuracy of CovLscan. MsMcGAN uses Multi-Level Dilated Residual Blocks (ML-DRBs) to generate lung images that preserves visual structure and contextual disease information. A Patch- based Multi-scale Discriminator (PMsD) is created for fine-tuning the generator through distinguishing the patches representing varying scales of lesions of generated images from real images. Perceptual content (PC) loss is added to map between produced and reference modalities at varying levels of semantic abstraction. MsMcGAN model enhances lung disease prediction of CovLscan by increasing labeled lung disease images through generating high-quality target images. **Findings:** The test outcomes reveal that the MsMcGAN - CovLscan model accomplishes an overall accuracy of 97.5% and 97.6% on the collected CXR and CT images that is contrasted with the classical CNN models. **Novelty:** This study uniquely integrates ML-DRBs with a PMsD to enhance image synthesis and lung disease detection, combining multi-scale contextual feature extraction and adversarial training stability, which has not been explored in previous studies.

**Keywords:** Lung disease Classification; Generative Adversarial Network; Chest X-Ray; Computed Tomography; Deep Learning

## 1 Introduction

In recent times, Deep learning (DL) is an emerging field that is applied to diagnose a variety of lung diseases and provides beneficial hand for healthcare providers/doctors in determining accurate medical decisions[1]. According to this network structure, DL approaches are often classed as Convolutional Neural Network (CNN), Deep Neural Network (DNN), Long Short Term Memory (LSTM), Recurrent Neural Network (RNN) and so on[2]. All these techniques are organized using a multi-layered neural network, where numerous training samples pass through several intermediate layers positioned between the input and output layers. The network learns intricate features, enhancing its ability to predict or classify images for accurate results[3].

By utilizing the idea of DL models, various literature has been suggested for the automated identification of pulmonary diseases by utilizing various imaging modalities[4]. For instances, Kim et al.[5] developed DL based EfficientNet v2-M utilized for the classification of lung diseases in CXR images. The model uses pre-trained weights from ImageNet to enhance diagnostic performance of CADs. Data is augmented to increase sample numbers and variance diversities, then inputted into the model to extract meaningful features for lung disease classification. Shamrat et al.[6] created LungNet22, a model for classifying and predicting lung diseases using raw CXR data from multiple sources. The model uses the EnsNet method to remove annotations, CLAHE enhancement methods to improve image quality, and Green Fire Blue filtering technique to enhance image properties.

Junayed et al.[7] developed an End-to-End DNN (E2E-DNN) for recognizing and classifying interstitial lung disease using lung CT image patches. The architecture included convolutional, batch normalization, max pooling, dense and dropout layers. Siddiqui et al.[8] presented an Ensemble Learning (EL) through TL using CXR and CT for lung diseases identification. Initially, the images underwent pre-processing as well as normalization and then fed into pre-trained CNN models employed with TL model to extract the features. Finally, EL model was devised for the identification and classification of lung diseases. Alshmrani et al.[9] constructed a multi-class pulmonary disease categorization model based on DL model using CXR images. Initially, the collected images were pre-processed and normalized. Then, pre-trained model VGG19 model and three blocks of CNN was adopted as the feature extraction. Podder et al.[10] developed an optimized DenseNet201 layer for diagnosing infectious lung diseases called LDDNet using CXR and CT images. The model was enhanced with global average pooling, batch normalization, dense layer, and dropout layer, using optimizers like Adam, nadam, and SGD.

Heidari et al.[11] constructed a lung cancer prediction technique using chest CT images based on Federated Learning (FL) and blockchain systems. In this method, the FL trained the model using blockchain technology ensuring the organization's anonymity while authenticating the data. Rajasekar et al.[12] constructed a DL model for lung cancer detection using CT images. In this method, multiple models based on CNN and VGGNet were used for the disease classification. Farhan et al.[13] introduced the Robust Multi-modal DL-based Automatic Lung Disease Classification (RMDL-ALDC) model, which utilizes both CXR and CT images. This modified CNN architecture incorporates L1 normalization and convolutional layers for efficient low-dimensional feature extraction. Al-Sheikh et al.[14] constructed a Multi-Class CNN (MC-CNN) structure to classify the lung diseases from CXR and CT images. In this approach, a k-symbol Lerch transcendent model was used to create an image enrichment model during pre-processing

Salam et al.[15] created a hybrid model by combining CNN and various Machine Learning (ML) algorithms. CNN extracted features are then classified using ML algorithms like Random Forest classifiers. The hybrid model increases the accuracy while reducing training complexity. Aharonu et al.[16] proposed a Deep Convolutional Generative Adversarial Network (DCGAN) that incorporates Enhanced Prism Refraction Search (EPS). To eliminate extraneous pixels and noise, an adaptive wavelet denoising was employed. Afterwards, an augmented attention U-Net was used for accurate picture segmentation. The EPS algorithm enhances the neural network's loss parameters to boost detection accuracy. Table 1 enlists all these models with their advantages and disadvantages.

On the other hand, DL models used to diagnose lung illnesses are significantly hindered by this model's epistemic ambiguity. Koushika et al[17] developed Convolutional Lung disease scan (CovLscan) eliminates the problem of epistemic uncertainty in the fast diagnosis of different types of lung illnesses using CXR and CT images. The U-Net model was used to segment the images. Additionally, features from CXR and CT images were retrieved using the pre-trained CNN models: ResNet50, DenseNet121, InceptionResNetV2, and XceptionV3. Ensemble-Convolutional Long Short Term Memory with Extreme Machine Learning (E-conLSTM-ELM) processed the classifcation of diseases.

### 1.1 Research Gap

From the Table 1, it is observed that the limitations of exiting model including potential overfitting indicated by differences in validation and testing accuracies, the data imbalance issues even in ensemble learning and multiple CNN architecture make challenges with scalability, and difficulty of handling class imbalance and noisy images, which may affect prediction accuracy. In contrast, all these models rely heavily on labeled training data with limited data annotations especially for new diseases.

**Table 1.** Summary of the Existing Lung Disease detection and Classification Models

| Author [Ref No] | Methods | Merits | Demerits |
|---|---|---|---|
| Kim et al. [5] | EfficientNet v2-M | The model benefits from pre-trained weights on large datasets, enhancing feature extraction capabilities and improving classification performance, especially when labeled medical data is limited. | This model loses a substantial amount of useful data, resulting in overfitting difficulties. |
| Shamrat et al. [6] | LungNet22 | The model can efficiently identify cases requiring immediate attention, optimizing the allocation of medical resources. | This model results in high time complexity issues. |
| Junayed et al. [7] | E2E-DNN | This approach enhances the model's learning efficiency, allowing for better generalization and more predictions that are accurate. | This model requires a significant amount of time to train. |
| Siddiqui et al. [8] | EL (Inceptionv3, Xception, InceptionResNetv2 and DenseNet12) | Demonstrates the potential for automated, fast, and accurate COVID-19 detection, which is crucial for timely interventions. | The model was trained with limited number of samples and results with high dimensionality issues |
| Alshmrani et al. [9] | CNN, VGG19, FC | Utilizes a multi-class approach, allowing the classification of six lung diseases, enhancing diagnostic capabilities. | The weight layers were not optimized which increases the computational complexity. |
| Podder et al. [10] | LDDNet | LDDNet features fewer trainable parameters, improving efficiency | Improper adjustment of the learning rate during early training stages may lead to overfitting issues |
| Heidari et al. [11] | FL, blockchain | Federated learning ensures that patient data remains on local devices, preserving privacy and confidentiality | The validated accuracy was slightly reduced while processing with full version of dataset |
| Rajasekar et al. [12] | CNN, VGGNet | The model provides enhanced accuracy with higher recognition rate | It needs testing to distinguish every class of pulmonary disorders as it was only suitable for a smaller dataset |
| Farhan et al. [13] | RMDL-ALDC | Suitable for a range of lung-related diseases, enhancing its clinical applicability. | Due to the increased number of learning factors, there was a higher level of sophistication |
| Al-Sheikh et al. [14] | Multi-Class CNN | This approach could yield more realistic results for image enhancement compared to linear methods | The considered database was limited, and the restrictions of hardware conditions were not solved |
| Salam et al. [15] | Hybrid CNN-Random Forest | In a variety of testing conditions, the hybrid model demonstrated strong performance. | The obtained accuracy is low for the given dataset |
| Aharonu et al [16] | DCGAN- EPS | The sufficient images were generated to improve the recognition accuracy of lung diseases | The parameter setting is challenging for applying various size of input dataset |
| Koushika et al [17] | CovLscan | Provided best results for balanced datasets | The augmentation models affected with blurry images due to compressions |

### 1.2 Major Contributions

CovLscan framework[17] achieved best results for balanced datasets. The datasets were balanced using simple augmentation models. The simple augmented models perform poor while creating images since the input images are frequently blurry and of lesser quality because of compression. This study proposes an MsMcGAN - CovLscan model to address the aforementioned challenges and enable effective lung disease identification and categorization. This approach use ML-DRBs to enhance lung region of generated lung disease images by MsMcGA by preserving visual structure and contextual lesion information. The ML-DRB enlarges the image while maintaining low computational overhead and captures contextual details from CXR and CT scan images. It also maintains disease structure, shape, and connectivity, resulting in more detailed feature maps. This method improves adversarial training stability and promotes the generator to mimic lung disease image distribution PMsD is used to fine-tune the image generation of generator by discriminating the generated images with real lung disease images at different scales. The integration of numerous scales increases the generator's effectiveness. PC loss is used to map the produced and target modalities at various semantic scales. MsMcGA enables large-scale patches to produce images from tiny disease-affected areas, while small-scale patches generate images from whole images. Thus, the MsMcGAN improves CovLscan training by increasing the number of labeled lung disease images, creating high-quality target images from source modal images and allowing for more efficient lung illness predictions.

## 2 Methodology

The figure 1 depicts the block structure of the proposed model.

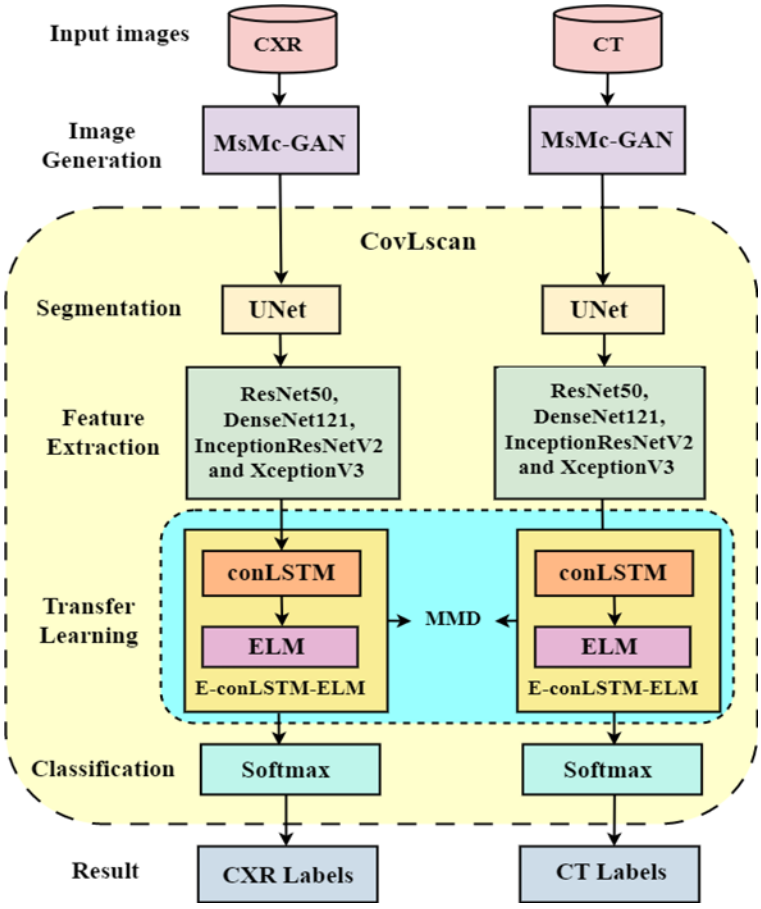


**Fig 1.** Pipeline of the proposed model

Initially, the CXR and CT images is fed into the MsMcGAN model to produce the reference models from input sources. Then, the generated images will be fed into the CovLscan framework for the detection of lung diseases. In CovLscan, the

generated images will be then inputted to UNet model for segmentation. Then, the pre-trained models derive the attributes from segmented CXR and CT images. The classifcation of diseases is processed by E-conLSTM-ELM is applied in CovLscan to exchange the knowledge between features and classes relation among CXR and CT images. Deep domain adaptation reduces domain shift using Maximum Mean Discrepancy (MMD). Finally, the softmax layer of conLSTM was employed for the classification of different lung disease types.

## 2.1 Dataset Description mention the dataset availability URL and the date of access

In this study, two datasets are used which are created from various benchmark databases. The construction of dataset and definition are as follows:

**CXR data**: The CXR dataset [18] contains 112,120 frontal-view X-ray images from 30,805 patients, covering fourteen common thoracic pathologies, including Pneumothorax, Pneumonia, Fibrosis, Effusion, Emphysema, Atelectasis, Edema, Infiltration, Mass, Cardiomegaly, Nodule, Consolidation, Pleural thickness, and Hernia. It also includes data on both COVID-19 and non-COVID CXR cases, which may involve other conditions such as bacterial or viral infections or chronic obstructive pulmonary disease [19]. For the purposes of this experiment, only five pathologies like Infiltration, Atelectasis, Normal, Pneumonia and COVID-19 are considered.

**CT data**: The same five disease categories utilized in CXR dataset is used for the CT images, with data sourced from various open public repositories. Lung Atelectasis images were sourced from [20], COVID-19 and Non-COVID (Normal) images from [21], viral Pneumonia images from [22] and Infiltration images from [23].

## 2.2 Multi-Scale Multi-Class Generative Adversarial Network for Image Generation

The proposed MsMcGAN is developed for multiclass (CXR and CT) image generation. The traditional CycleGAN is enhanced by substituting the standard residual block with a DRB in the generator and upgrading the Patch-based Single-scale Discriminator (PSsD) to a dual-scale patch-based discriminator. Additionally, the perceptual loss is incorporated to broaden the context loss. When the input image is an $a$, an image $\hat{b}$ is determined by the $G_{a\to b}$. Then, the image $\hat{b}$ is inputted to $G_{b\to a}$ to obtain the image $\hat{a}$, where the image $\hat{b}$ is a necessitated target-class image and the image $\hat{a}$ is considered as cyclically generated image. Given an input image $b$, the output $\hat{a}$ represents the desired target class image, while the generated $\hat{b}$ corresponds to the cyclically generated version of image $b$.

### *2.2.1. Generator-driven Dilated Residual Block*

The DRB is adopted as the generator of MsMcGAN. In addition to the standard residual blocks, the DRBs incorporate a dilated convolution with a dilation rate of two, positioned prior to the Layer of Occurrence Normalization (LON) and is statistically illustrated in **Eq. (1)**.

$$b_{suvq} = \frac{a_{suvq} - \delta_{su}}{\sqrt{\mu_{su}^2} + \beta} \tag{1}$$

$$\delta_{su} = \frac{1}{hw} \sum_{k=1}^{w} \sum_{z=1}^{h} a_{sukz} \tag{2}$$

$$\mu_{su}^2 = \frac{1}{hw} \sum_{k=1}^{w} \sum_{z=1}^{h} (a_{sukz} - zx_{su})^2 \tag{3}$$

Consider $a \in R^{t*c*h*w}$ be an input tensor constitutes a set of $t$ images. $suvq$ defines the input image in which $v$ and $q$ are the spatial dimensions. $u$ and $k$ denotes feature channel and image index in the batch. The derived lung region from CXR and CT images identifies the dependency among the remote areas that play a key role in texture synthesis. The MsMcGAN consists of nine DRBs that can significantly expand lung regions from $3 \times 3$ to $5 \times 5$ while maintaining computational efficiency. The DRB is found to effectively retain contextual and structural information from CXR and CT images. The Figure 2 depicts the entire structure of MsMcGAN Model.

### *2.2.2. Patch based Multi-Scale Discriminator (PMsD)*

The PMsD is introduced to differentiate the CXR and CT images using two patches of various lung region in MsMcGAN. The discriminator with dimensions of $4 \times 4$, is used to create a feature map with dimensions of $4 \times 4$. The discriminator with a $4 \times 4$ patch size has an additional convolutional layer with stride 2 than the $8 \times 8$ discriminator with the rest remaining unchanged. The final layer of the standard discriminator produces a scalar value, representing the overall weighted value of the image, which fails to fetch the local attributes of the image. The PMsD generates a matrix $M$ with $n \times n$ elements that depict the local areas in an image. It detects each $n \times n$ patch as genuine or false whuich extracting and characterizing local image features characteristics and generating high-accuracy images.

The PSsD is unable to effectively focus on both local and global regions of CXR and CT images. This model utilizes the dual PMsD of various scales. The $4 \times 4$ feature map patches each have an extended lung region processing with large-scale global information on CXR and CT images. Each $8 \times 8$ feature map patch has a small lung region for processing local information for CXR and CT images. Integrating the patch-based discriminators results in coherent global and local images for the prediction of large and small patches of CXR and CT images.

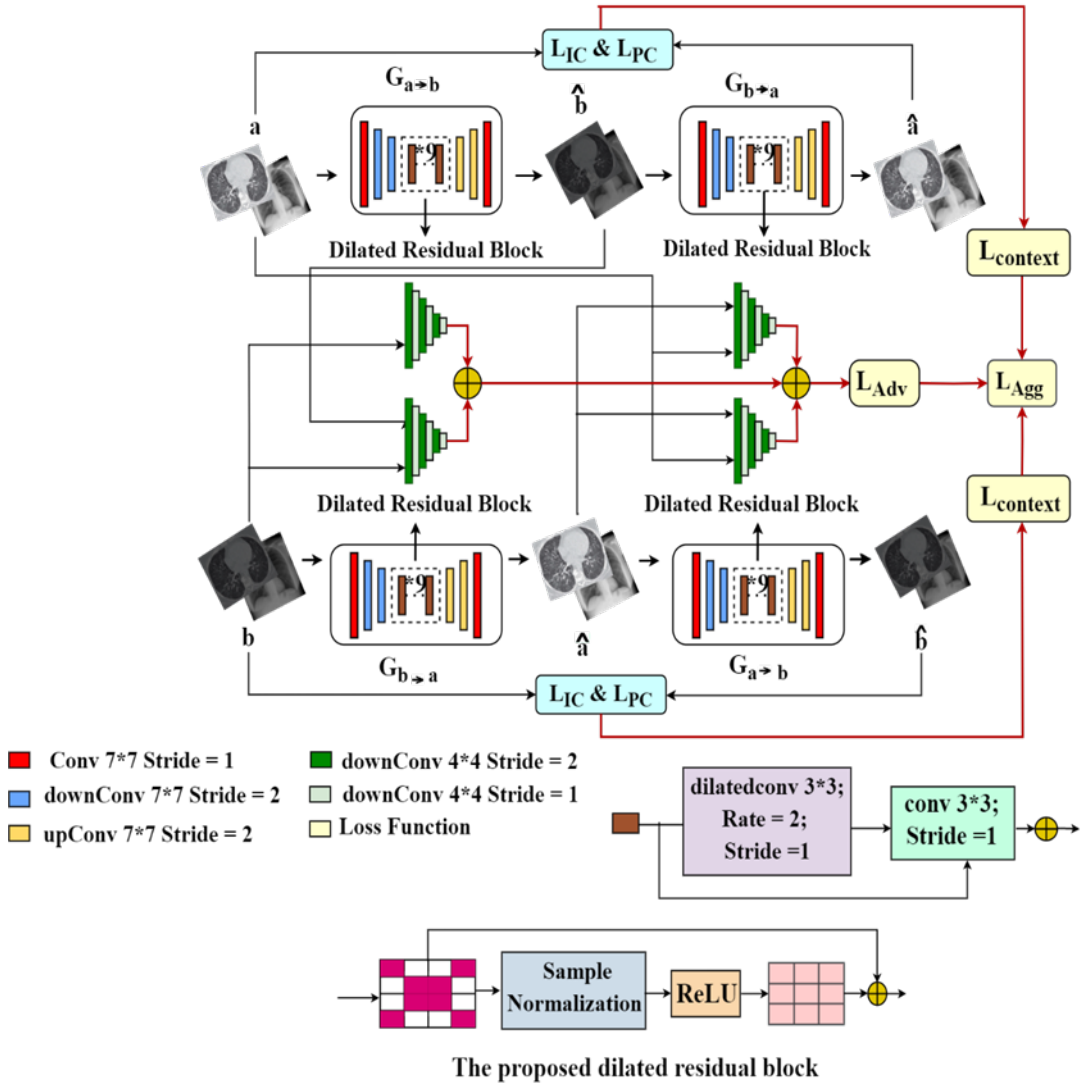


**Fig 2.** MsMcGAN Architecture

### 2.2.3. Loss Operation

Context Loss: The generator must regularize the loss function by adjusting the contextual loss function, which is described as the disparity between the source and synthesized images. Here, the contextual loss of MsMcGAN composed of the Iterative Context (IC) and PC loss. The IC loss is integrated in single phase GAN that employs the $L_1$ norm to compute the one-pixel average error among the actual and generated images. The IC loss is statistically illustrated by **Eq. (4)**

$$L_{ICL}\left(G_{a\to b}, G_{b\to a}\right) = E_{aJ_{data}}(a)¿ \tag{4}$$

In **Eq. (4)**, $a$ and $b$ represents the actual images of the input. $G_{b\to a}\left(G_{a\to b}(a)\right)$ and $G_{a\to b}\left(G_{b\to a}(b)\right)$ are the generated images. $\|.\|_1$ illustrates the $L_1$ norm. The $L_1$ norm alone will not suffice to reconstruct all texture features due to the missing high-frequency details in the actual as well as generated image. The PC loss is implied to improve the standard of synthesized CXR and CT images. The model's approach focuses on conceptual information and rebuild intricate texture details, resulting in a more accurate image compared to the real images. The actual and generated image are derived from pre-trained CNN models[17]. The PC loss is provided in **Eq. (5)**

$$\begin{aligned} &L_{PC}\left(G_{a\to b}, G_{b\to a}\right) = \\ &E_{aJ_{data}}(a)¿ \\ &E_{bJ_{data}}(b)¿ \end{aligned} \tag{5}$$

In **Eq. (5)**, $\Phi$ represents the features retrieved from pre-trained CNN models[17]. The consistency loss is the sum of IC and PC loss as given in **Eq. (6)**

$$L_{context}\left(G_{a\to b}, G_{b\to a}\right) = \gamma L_{occurence}\left(G_{a\to b}, G_{b\to a}\right) + \alpha L_{PC}\left(G_{a\to b}, G_{b\to a}\right) \tag{6}$$

In **Eq. (6)**, $\gamma$ and $\alpha$ regulates the comparative prominence of the $L_{occurence}$ and $L_{PC}$ correspondingly.

Adversarial loss: The adversarial loss trains the generators and discriminators in a competitive manner, enabling them to generate real images. This adversarial loss employs the cross-entropy loss and is provided in **Eq. (7)**.

$$L_{classical} = E_{aJ_{data(a)}}¿ \tag{7}$$

In **Eq. (7)**, $f$ is the input data, $a$ is the original information. It is determined that the loss operation employed in the training task might cause vanishing gradients. So, the minimal square loss approach is adopted for more efficient discriminator training. The generated and target image is labelled as $x$ and $y$ in which the $x = 0$ and $y = 1$. The adversarial loss is computed in **Eq. (8)**

$$L_{Adv}\left(G_{a\to b}, G_{b\to a}, d_a^{4\times 4}, d_b^{4\times 4}, d_a^{8\times 8}, d_b^{8\times 8}\right) = E_{aJ_{data(a)}}¿ \tag{8}$$

In Eq. (8), $d_a^{4\times 4}$ and $d_b^{4\times 4}$ are discriminators with $4 * 4$ patch dimension. $d_a^{8\times 8}$ and $d_b^{8\times 8}$ are discriminators with $8 \times 8$.

Aggregative Loss: The cumulative loss is determined by integrating the PC loss and adversarial loss, provided in **Eq. (9)**

$$L_{Agg}\left(G_{a\to b}, G_{b\to a}, d_a^{4*4}, d_b^{4\times 4}, d_a^{8\times 8}, d_b^{8\times 8}\right) = L_{context}\left(G_{a\to b}, G_{b\to a}\right) + L_{Adv}\left(G_{a\to b}, G_{b\to a}, d_a^{4\times 4}, d_b^{4\times 4}, d_a^{8\times 8}, d_b^{8\times 8}\right) \tag{9}$$

The optimal generators $G_{a\to b}^{¿} G_{b\to a}^{¿}$ are determined by lessening $G_{a\to b}^{¿} G_{b\to a}^{¿}$ and maximizing $d_a^{4\times 4}, d_b^{4\times 4}, d_a^{8\times 8}, d_b^{8\times 8}$. The objective of the total loss is illustrated as **Eq. (10)**.

$$\begin{aligned} &G_{a\to b} G_{b\to a} = \arg \min_{G_{a\to b}, G_{b\to a}} \max_{d_a^{4\times 4}, d_b^{4\times 4}, d_a^{8\times 8}, d_b^{8\times 8}} \\ &L_{Agg}\left(G_{a\to b}, G_{b\to a}, d_a^{4\times 4}, d_b^{4\times 4}, d_a^{8\times 8}, d_b^{8\times 8}\right) \end{aligned} \tag{10}$$

The generated CXR and CT images will be fed into the CovLscan for the identification and categorization of lung diseases. This approach improves training stability and lung disease image distribution that improves the CovLscan training by increasing

labelled lung disease images, generating high-quality target images from source modal images and facilitating efficient lung disease prediction. The algorithm for the proposed MsMcGAN - CovLscan model is illustrated below.

***Algorithm: MsMcGAN-CovLscan***

***Input:*** CXR and CT images

***Output:*** Generating high-quality target images for lung disease prediction

**Step 1:** The collected CXR and CT images are generated using ML-DRBs to improve lung region.

**Step 2:** Multi-scale discriminator is developed to adjust the generator for distinguishing the patches representing varying sizes of lesions.

**Step 3:** PC loss is added to map between generated and target modalities at different semantic levels.

**Step 4:** PMsD optimizes the residual block of GAN enables to generate the large-scale patches images to small disease-affected regions and small-scale patches from whole images.

**Step 5:** ML-DRB enlarges the image to extract contextual details, preserving the disease's form, layout, and connections to improve feature maps at higher resolutions.

**Step 6:** The adversarial training stability helps to improvise the generator to replicate disease image distribution.

**Step 7:** The generated CXR and CT images from MsMcGAN are fed intro the CovLscan model.

**Step 8:** In CovLscan, U-Net model is utilized to segment the images.

**Step 9:** Pre-trained CNN models are used to select the features from the generated images

**Step 10:** E-conLSTM-ELM exchanges the knowledge between features and classes relation among CXR and CT images using Maximum Mean Discrepancy (MMD).

**Step 12:** Finally, the softmax layers of CovLscan provides the high-quality target images for facilitating efficient lung disease prediction.

## 3 Results and Discussion

This section analyses the efficiency of the proposed MsMcGAN - CovLscan technique via executing it in MATLAB 2017b with the help of CXR and CT images (provided in section 3.1). A comparative study is presented between MsMcGAN – CovLscan with existing pulmonary diseases models like LDDNet[10], RMDL-ALDC[13], MC-CNN[14], CovLscan[17] and. For the experimental purposes, the collected dataset is divided into 70% (training) and 30% (testing). The total number of CXR and CT images used for performance evaluation before and after augmentation is presented in Table 2.

**Table 2.** Dataset Details

| Classes | Count of CXR images before augmentation | Count of CXR images after augmentation | Count of CXR images before augmentation | Count of CT images before augmentation |
|---|---|---|---|---|
| COVID-19 | 1345 | 1562 | 1002 | 1510 |
| Pneumonia | 1443 | 1598 | 1762 | 1762 |
| Normal | 1345 | 1562 | 984 | 1555 |
| Infiltrate | 270 | 1501 | 260 | 1521 |
| Atelectasis | 290 | 1513 | 310 | 1532 |

The suggested and existing models are assessed based on performance metrics below.

- **Accuracy (Acc)**: It represents the proportion of suitably categorized instances for each category of lung diseases relative to whole examples considered.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{11}$$

In above **Eq. (11)**, TP (True Positive) refers to the scenario where the model correctly identifies the type of lung disease, such as classifying Infiltrate as Infiltrate. TN (True Negative) denotes when the model precisely identifies Pneumonia as Pneumonia.

FP (False Positive) denotes the wrongly categorized model as lung disease (e.g., Covid-19, Pneumonia, Normal, Infiltrate and Atelectasis) as Atelectasis. FN (False Negative) refers to the case where the model wrongly classifies Pneumonia as Normal.

- **Precision (Pre):** It represents the percentage of correctly classified lung disease instances in relation to both TP and FP rates.

$$Precision = \frac{TP}{TP + FP} \tag{12}$$

- **Recall(Re):** Proportion of TP and FN rates for cases of precisely identified lung illnesses.

$$Recall = \frac{TP}{TP + FN} \tag{13}$$

- **F1-Score (F1):** It is the weighted mean of precision and recall.

$$F1 - score = 2 \times \frac{Precision \bullet Recall}{Precision + Recall} \tag{14}$$

Table 3 shows the MsMcGAN-CovLscan model performs better for every category of lung diseases when compared to previous classification models. Thanks to its enhanced labelled data and capacity to produce high-quality target images for each illness category from CXR and CT images, it achieves outstanding performance. In the case of infiltrate categorization utilizing CXR pictures, for example, MsMcGAN-CovLscan outperforms LDDNet, RMDL-ALDC, MC-CNN, and CovLscan by 14.09%, 11.08%, 7.50%, and 3.02% for CXR images; 12.73%, 10.06%, 6.01%, and 2.31% for CT images. In the case of COVID-19 classification, the precision of MsMcGAN – CovLscan is 14.58% and 14.8% higher than LDDNet, 9.95% and 9.9% greater than RMDL-ALDC, 6.24% and 5.34% greater than MC-CNN, 1.88% and 1.04% higher than CovLscan models for CXR and CT images respectively. The another instance, in the case of normal classification, the recall of MsMcGAN - CovLscan is 15.94%, 11.85%, 7.56% and 2.24% (for CXR images); 20.42%, 15.31%, 11.18%, 7.91% and 1.83% (for CT images) is higher than other methods images respectively. It is clear from the assessments that the suggested approach consistently outperforms the current models.

Figure 3 illustrate the overall accuracy (in %) achieved by MsMcGAN-CovLscan model which outperforms other classification models for both CXR and CT images. For CXR images, MsMcGAN-CovLscan achieves 97.59% accuracy, which is 13.21%, 9.43%, 5.76%, and 2.13% higher than LDDNet, RMDL-ALDC, MC-CNN and CovLscan models for CXR images; MsMcGAN-CovLscan achieves 98.27%, which is also 12.54%, 9.76%, 6.61%, and 2.44% higher than other models for CT images.

Figure 4 represents the overall precision (in %) obtained by different models for predicting various lung diseases from CT and CXR images. The precision of MsMcGAN - CovLscan is superior to other classification models for both CXR and CT images. MsMcGAN – CovLscan is 12.57% and 13.08% higher than LDDNet, 9.39% and 8.91% greater than RMDL-ALDC, 5.78% and 5.10% greater than MC-CNN, 2.17% and 1.81% higher than CovLscan models for CXR and CT images respectively. From the analysis of Figure 5, the overall recall obtained by proposed models is almost higher than all existing models for CT (96.37%) and CXR (96.19%) images. The recall of MsMcGAN - CovLscan is 14.08%, 10.58%, 6.74%, and 2.85% (for CXR images); 12.29%, 9.32%, 5.58%, and 1.74% (for CT images) is higher than LDDNet, RMDL-ALDC, MC-CNN and CovLscan for CXR and CT images respectively.

Figure 6 depicted the overall F1 score (in %). MsMcGAN - CovLscan is achieved 12.71%, 9.41%, 6.41%, and 2.42% (for CXR images); 12.02%, 8.91%, 5.65%, and 1.83% (for CT images) is higher than LDDNet, RMDL-ALDC, MC-CNN and CovLscan for CXR and CT images respectively.

The reason for achieving higher result of MsMcGAN – CovLscan is introducing a multi-level dilated residual block in generator. This block captures disease patches, which are, differ in size and shape and their precise location is unknown. The patch-based multi-scale discriminator proposed in this study uses large-scale patches, which are created images from the similar tiny diseased regions, and small-scale patches, which are, create images from the whole images. Thus, the decimator can verify the lesions regions in the images irrespective of sizes. That is why MsMcGAN is beneficial to the generator's performance on various scales. The generated images by MsMcGAN are better in quality, which improve the training accuracy of CovLscan. This is the reason MsMcGAN – CovLscan obtained best results than any other model.

**Table 3.** Class wise Performance metric analysis

| Dataset | | | COVID-19 | Pneumonia | Normal | Infiltrate | Atelectasis |
|---|---|---|---|---|---|---|---|
| CXR Image dataset | LDDNet | Acc | 85.19 | 86.47 | 83.61 | 83.47 | 83.13 |
| | | Pre | 83.83 | 81.35 | 83.85 | 84.65 | 84.42 |
| | | Re | 81.25 | 83.59 | 83.21 | 82.17 | 81.63 |
| | | F1 | 83.22 | 83.31 | 83.91 | 84.63 | 83.38 |
| | RMDL-ALDC | Acc | 88.88 | 89.91 | 87.76 | 86.59 | 87.65 |
| | | Pre | 87.36 | 86.29 | 86.7 | 86.22 | 87.41 |
| | | Re | 85.87 | 85.76 | 86.25 | 86.64 | 84.82 |
| | | F1 | 88.78 | 86.88 | 87.04 | 87.12 | 85.13 |
| | MC-CNN | Acc | 92.09 | 93.38 | 91.85 | 90.28 | 91.53 |
| | | Pre | 90.41 | 89.45 | 91.85 | 89.98 | 90.36 |
| | | Re | 90.47 | 89.82 | 89.69 | 89.74 | 88.85 |
| | | F1 | 90.37 | 89.45 | 90.17 | 91.42 | 88.56 |
| | CovLscan | Acc | 95.44 | 96.13 | 95.66 | 94.54 | 95.55 |
| | | Pre | 94.28 | 92.72 | 95.72 | 93.97 | 93.4 |
| | | Re | 94.28 | 92.72 | 94.36 | 93.17 | 93.45 |
| | | F1 | 94.16 | 93.74 | 94.87 | 94 | 93.15 |
| | MsMc-GAN-CovLscan | Acc | 97.65 | 97.71 | 97.44 | 97.56 | 97.58 |
| | | Pre | 96.05 | 96.05 | 96.05 | 96.05 | 96.05 |
| | | Re | 96.53 | 96.61 | 96.47 | 96.38 | 96.24 |
| | | F1 | 96.29 | 96.14 | 96.38 | 96.66 | 96.51 |
| CT Image dataset | LDDNet | Acc | 85.46 | 86.82 | 85.43 | 86.37 | 86.19 |
| | | Pre | 83.64 | 83.12 | 84.81 | 81.54 | 83.36 |
| | | Re | 85.99 | 84.58 | 83.68 | 82.68 | 83.45 |
| | | F1 | 85.19 | 84.53 | 84.41 | 83.39 | 84.68 |
| | RMDL-ALDC | Acc | 88.15 | 89.91 | 89.29 | 88.47 | 88.33 |
| | | Pre | 87.37 | 86.98 | 88.94 | 86.69 | 87.31 |
| | | Re | 88.95 | 87.64 | 86.79 | 85.18 | 86.71 |
| | | F1 | 88.27 | 87.94 | 87.38 | 86.94 | 87.21 |
| | MC-CNN | Acc | 91.79 | 92.26 | 92.37 | 91.85 | 91.61 |
| | | Pre | 91.15 | 90.38 | 92.48 | 90.41 | 91.94 |
| | | Re | 91.28 | 92.76 | 89.42 | 89.84 | 90.66 |
| | | F1 | 91.73 | 90.37 | 91.11 | 90.38 | 90.46 |
| | CovLscan | Acc | 95.76 | 96.97 | 96.89 | 95.17 | 95.97 |
| | | Pre | 95.03 | 93.63 | 95.87 | 94.02 | 94.23 |
| | | Re | 95.11 | 95.09 | 94.76 | 93.97 | 94.21 |
| | | F1 | 94.87 | 94.06 | 95.09 | 94.73 | 94.38 |
| | MsMc-GAN-CovLscan | Acc | 97.79 | 97.73 | 97.43 | 97.37 | 97.61 |
| | | Pre | 96.02 | 96.54 | 96.49 | 96.53 | 96.28 |
| | | Re | 96.02 | 96.54 | 96.49 | 96.53 | 96.28 |
| | | F1 | 96.67 | 96.54 | 96.39 | 96.27 | 96.43 |

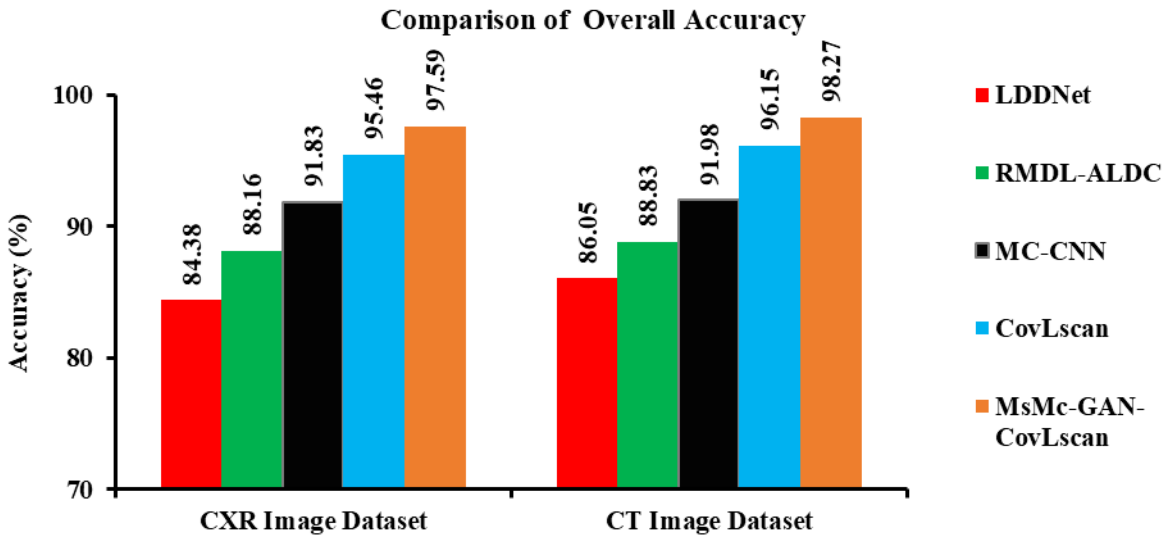


**Fig 3.** Overall Accuracy Comparison for Lung Disease Classification

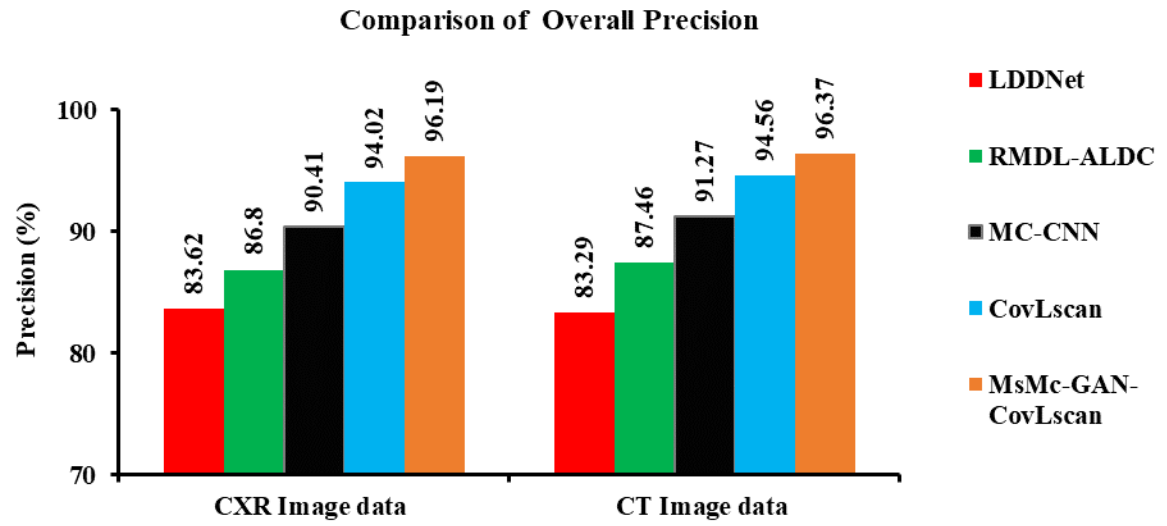


**Fig 4.** Overall Precision Comparison for Lung Disease Classification

## 4 Conclusion

This study has introduced MsMcGAN – CovLscan for efficient lung disease detection and classification. MsMcGAN uses ML–DRBs to generate lung region through preserving visual structure and contextual disease information while generating images.

**Important Achievements and Originality**:

A multi-scale discriminator distinguishes patches representing varying sizes of lesions. PC loss is added to map between synthesized and reference modalities at varying levels of semantic abstraction for generating good quality images. This approach allows extensive patches to synthesize images from smaller regions affected by disease; this approach enhances adversarial training stability and encourages the generator to replicate lung disease from the range of image distribution.

**Strengths and Weaknesses:**

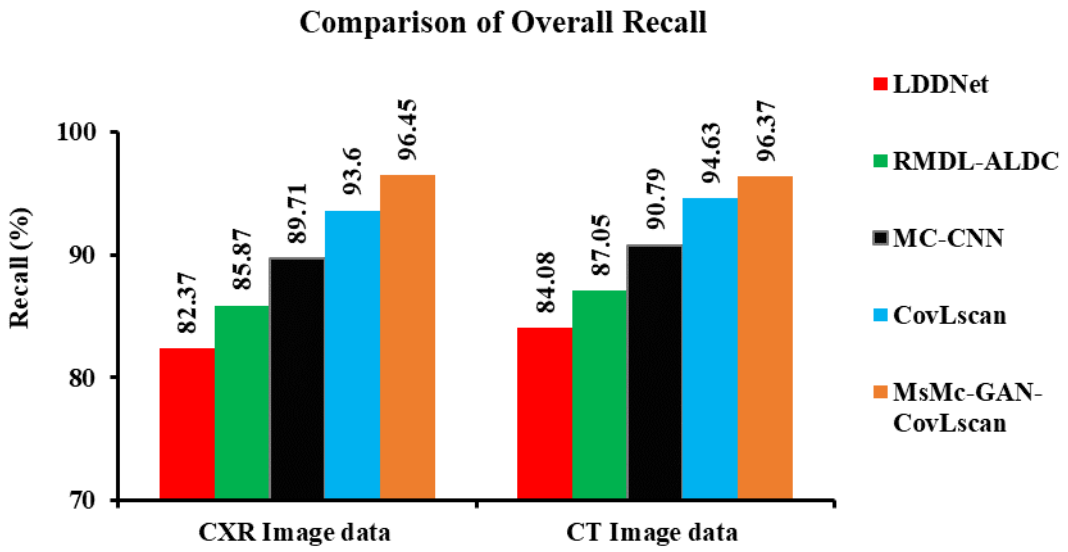


**Fig 5.** Overall Recall Comparison for Lung Disease Classification

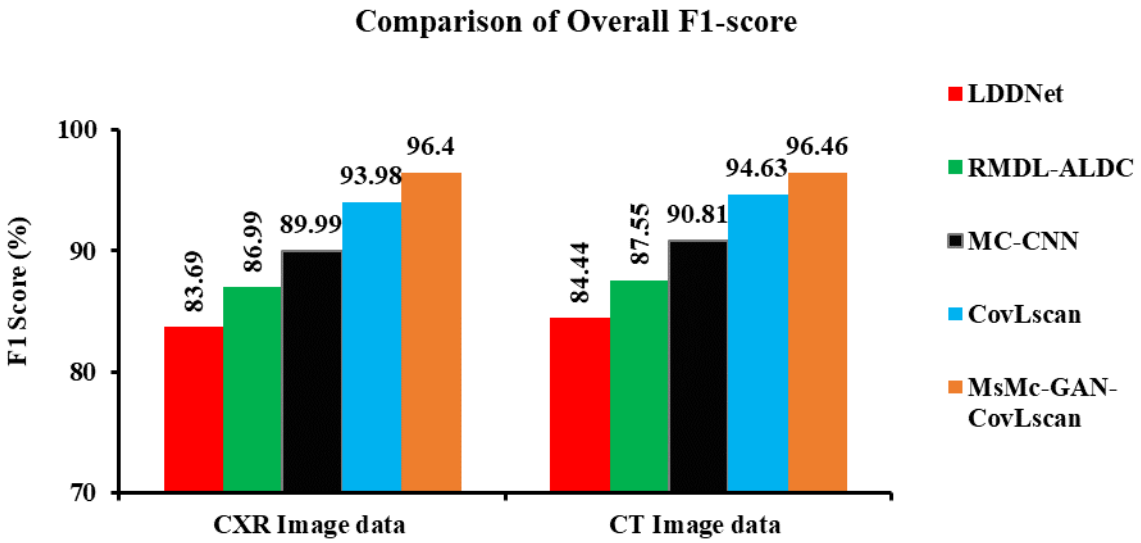


**Fig 6.** Overall F1-score Comparison for Lung Disease Classification

Although proposed method perform very well, success of results is manly based on hyper-parameters setting of proposed model, the values of which affect the final error rate. However, finding suitable hyper-parameter values is difficult; it is typically done manually and requires a lot of effort.

**Prospects and Recommendation:**

In future, it can be solved by utilizing optimization algorithms. In addition, the resemblance of disease features can be examined comprehensively by the incorporation of Siamese neural networks, which would facilitate the understanding of the relationships and similarities among the characteristics of analogous diseases.

## References

1) Yu Z, Wang K, Wan Z, Xie S, Lv Z. Popular deep learning algorithms for disease prediction: a review. *Cluster Computing*. 2023;26(2):1231–1251. https://doi.org/10.1007/s10586-022-03707-y.

2) Kieu STH, Bade A, Hijazi MHA, Kolivand H. A survey of deep learning for lung disease detection on medical images: state-of-the-art, taxonomy, issues and future directions. *Journal of Imaging*. 2020;6(12):131. https://doi.org/10.3390/jimaging6120131.
3) Sutha P, Shenbagavalli R. A survey of machine learning techniques for lung disease detection on medical images. *AIP Conference Proceedings*. 2023;2618(1):1. https://doi.org/10.1063/5.014978.
4) Al-qaness MA, Zhu J, AL-Alimi D, Dahou A, Alsamhi SH, Elaziz MA, et al. Chest X-ray Images for Lung Disease Detection Using Deep Learning Techniques: A Comprehensive Survey. *Archives of Computational Methods in Engineering*. 2024;31:3267–3301. https://doi.org/10.1007/s11831-024-10081-y.
5) Kim S, Rim B, Choi S, Lee A, Min S, Hong M. Deep learning in multi-class lung diseases' classification on chest X-ray images. *Diagnostics*. 2022;12(4):915. https://doi.org/10.3390/diagnostics12040915.
6) Shamrat FJM, Azam S, Karim A, Islam R, Tasnim Z, Ghosh P, et al. LungNet22: A Fine-Tuned Model for Multiclass Classification and Prediction of Lung Disease Using X-ray Images. *Journal of Personalized Medicine*. 2022;12(5):680. https://doi.org/10.3390/jpm12050680.
7) Junayed MS, Jeny AA, Islam MB, Ahmed I, Shah AS. An efficient end-to-end deep neural network for interstitial lung disease recognition and classification. *Turkish Journal of Electrical Engineering and Computer Sciences*. 2022;30(4):1235–1250. https://doi.org/10.48550/arXiv.2204.09909.
8) Siddiqui SA, Fatima N, Ahmad A. Chest X-ray and CT scan classification using ensemble learning through transfer learning. *EAI Endorsed Transactions on Scalable Information Systems*. 2022;9(6):e8–e8. https://doi.org/10.4108/eetsis.vi.382.
9) Alshmrani GMM, Ni Q, Jiang R, Pervaiz H, Elshennawy NM. A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images. *Alexandria Engineering Journal*. 2023;64:923–935. https://doi.org/10.1016/j.aej.2022.10.053.
10) Podder P, Das SR, Mondal MRH, Bharati S, Maliha A, Hasan MJ, et al. LDDNet: A Deep Learning Framework for the Diagnosis of Infectious Lung Diseases. *Sensors*. 2023;23(1):480. https://doi.org/10.3390/s23010480.
11) Heidari A, Javaheri D, Toumaj S, Navimipour NJ, Rezaei M, Unal M. A new lung cancer detection method based on the chest CT images using Federated Learning and blockchain systems. *Artificial Intelligence in Medicine*. 2023;141:102572. https://doi.org/10.1016/j.artmed.2023.102572.
12) Rajasekar V, Vaishnnave MP, Premkumar S, Sarveshwaran V, Rangaraaj V. Lung cancer disease prediction with CT scan and histopathological images feature analysis using deep learning techniques. *Results in Engineering*. 2023;18:101111. https://doi.org/10.1016/j.rineng.2023.101111.
13) Farhan AM, Yang S, Al-Malahi A. An Automatic Classification Framework using Multi-Modal Deep Learning for Lung Diseases. 2023;p. 28–32. https://doi.org/10.1109/icnwc60771.2024.10537498.
14) Al-Sheikh MH, Dandan OA, Al-Shamayleh AS, Jalab HA, Ibrahim RW. Multi-class deep learning architecture for classifying lung diseases from chest X-ray and CT images. *Scientific Reports*. 2023;13(1):19373. https://doi.org/10.1038/s41598-023-46147-3.
15) Salam MA, Abdellatif A, Abdallah M, Salam NA. A Hybrid Deep Learning and Machine Learning Model for Multi-Class Lung Disease Detection in Medical Imaging. *International Journal of Intelligent Engineering & Systems*. 2025;18(1). https://doi.org/10.22266/ijies2025.0229.58.
16) Aharonu M, Ramasamy LK. An intelligent generative adversarial network multistage lung cancer detection and subtypes classification. *Int J Mach Learn & Cyber*. 2024. https://doi.org/10.1007/s13042-024-02484-x.
17) Koushika T, Nithya P. Domain Adapted Transfer Learning and Ensemble Method in Deep Learning for Pulmonary Disease Detection and Classification. *International Journal of Intelligent Engineering and Systems*. 2024;17(5):294–305. https://doi.org/10.22266/ijies2024.1031.24.
18) ChestXray8 dataset. Accessed 10 January 2025. Available from: https://nihcc.app.box.com/v/ChestXray-NIHCC.
19) Kaggle Thymooney ChestXray Pneumonia Dataset. Accessed 15 January 2025. Available from: https://www.kaggle.com/paultimo/thymooney/chestxray-pneumonia.
20) Radiopaedia. Lung Atelectasis. Available from: https://radiopaedia.org/articles/lung-atelectasis?lang=us.
21) Kaggle. COVID-19 Lung CT Scans Dataset. Accessed 20 January 2025. Available from: https://www.kaggle.com/datasets/mehradaria/covid19-lung-ct-scans.
22) Radiopaedia. Viral Respiratory Tract Infection. Accessed 20 January 2025. Available from: https://radiopaedia.org/articles/viral-respiratory-tract-infection.
23) Radiopaedia. Chest X-Ray Playlist. Accessed 20 January 2025. Available from: https://radiopaedia.org/playlists/41156?lang=us.