Pertanika J. Sci. & Technol. 34 (2): 601 - 624 (2026)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Multi-classification of Chest X-ray Images Using an Enhanced Deep Learning Approach

Revathi A[1, 2]* and Savadam Balaji[1]

[1]*Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, 500075 Hyderabad, Telangana, India*
[2]*Department of Information Technology, Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering and Technology, 500090 Hyderabad, Telangana, India*

ABSTRACT

Healthcare practitioners can rapidly recognise lung problems due to the invaluable and essential role that chest X-ray imaging plays in lung diagnosis. Recently, deep learning techniques have gained popularity and demonstrated promising outcomes in automatically interpreting medical images, especially in the specialty of chest radiology. For classifying chest X-ray images into multiple classes, we propose an enhanced deep learning approach in this research. Initially, the input images are resized, normalised, and data augmented to improve the network's generalisation ability. The local and global features are extracted using the MobileNetV2 model. An Intelligent Prairie Dog Optimisation (IPDO) algorithm is employed to select the optimal features, thereby minimising the dimensionality of the features. This algorithm finds the optimal solution with a higher convergence rate. Finally, the multi-class chest x-ray images are classified using the Enhanced GoogLeNet (EGoogLeNet) model. The model parameters are optimised using the Adam optimiser. The proposed approach is evaluated using a large-scale C19RD dataset, and the findings indicate improved performance in lung multi-class diagnosis. Extensive experiments are conducted against existing state-of-the-art approaches to assess the proposed model's performance, yielding significant improvements in F1-score, recall, precision, and accuracy. Promising findings are found in the assessment of the proposed method's performance (EGoogLeNet), which achieves a multi-class classification accuracy of 99.41%. The proposed model performed better, enabling medical professionals to diagnose and treat patients more rapidly and effectively.

Keywords: Adam optimiser, chest x-ray images, deep learning, enhanced GoogLeNet, MobileNetV2 model

ARTICLE INFO

Article history:
Received: 18 September 2024
Accepted: 09 December 2025
Published: 01 April 2026

DOI: https://doi.org/10.47836/pjst.34.2.01

E-mail addresses:
revathi_a@vnrvjiet.in (Revathi A)
balajis@klh.edu.in (Savadam Balaji)
* Corresponding author

e-ISSN: 2231-8526

INTRODUCTION

Lung disorders, including opacity, fibrosis, viral pneumonia, and COVID-19 pneumonia, as well as tuberculosis, are significant global health problems that affect many people worldwide (Singh et al., 2025). These diseases are characterised by their detrimental impact on pulmonary function, most notably reducing lung elasticity (Prasath et al., 2025). Due to this decreased flexibility, the lungs can store less air overall, thereby compromising respiratory performance (Slimi et al., 2025). Certain lung diseases can spread quickly, particularly infectious diseases like pneumonia and opacity. This highlights the vital importance of early and precise diagnosis (Sharmili et al., 2026). Early diagnosis of these conditions is crucial because it enables the prompt initiation of proper care, which is necessary to slow the disease's progression and improve patient outcomes (Slika et al., 2024). Individual patients benefit from timely and accurate lung disease diagnosis, as it enables them to receive proper treatment. Maintaining public health is essential because it prevents the transmission of contagious respiratory diseases (Bhardwaj et al., 2025).

Using cutting-edge technology has become crucial in the dynamic and constantly evolving field of healthcare diagnosis, particularly in pulmonary health (Kaushik et al., 2025). In this regard, chest X-ray imaging has become a vital diagnostic tool. It provides an effective and non-invasive method for identifying and evaluating lung abnormalities (Odeh et al., 2024). For healthcare practitioners, this imaging technology is crucial. It is essential for the timely and precise detection of various lung diseases, making it a crucial component of lung disease management and treatment (Kumar et al., 2025). Chest X-ray (CXR) imaging remains one of the most widely used and accessible imaging modalities for diagnosing lung diseases due to its low cost, speed, and ability to provide sufficient anatomical information for detecting thoracic abnormalities (Satapathy et al., 2025). Although CXR may not match the resolution of CT scans, it still enables the visualisation of critical features such as infiltrates, consolidations, and nodules that are relevant for clinical assessment (Tekin et al., 2025). Clinicians need these intricate visualisations to correctly detect, monitor, and treat lung diseases (Okolo et al., 2025).

The importance of chest radiography in diagnosing and treating lung disorders is increasing as medical science advances, underscoring the need for ongoing research and improvement in this vital field of medicine (Gonçalves et al., 2025). However, CXR images are often affected by issues such as overlapping anatomical structures, low contrast for subtle lesions, and variations in image quality, which can obscure critical diagnostic details (Chen et al., 2026). Radiologists also face challenges, including inter-observer variability and difficulty in distinguishing visually similar abnormalities, underscoring the need for advanced deep learning approaches to enhance accuracy and consistency (Yadav et al., 2024).

Deep learning methods utilise extensive medical image datasets to identify complex patterns and anomalies that may be overlooked by more conventional techniques, thereby leading to a more comprehensive understanding of pulmonary disorders

(Maniruzzaman et al., 2024). As a result, deep learning in chest radiology signifies not only a breakthrough in technology but also a substantial improvement in the capacity of healthcare providers to accurately and successfully identify and treat lung conditions (Zhao et al., 2025). Incorporating deep learning into chest imaging improves diagnostic precision and encourages the creation of individualised treatment plans (Egala et al., 2025a). These developments represent a breakthrough in managing and treating pulmonary health junctures and are essential to improving patient outcomes (Varadharajan et al., 2025).

To identify different lung-related disorders, this research proposes a robust and enhanced multi-class classification framework utilising a deep learning model. Our method utilises CNNs due to their demonstrated ability to automatically learn hierarchical and spatially invariant characteristics from medical images, despite many previous studies having used CNNs for comparable diagnostic tasks. CNNs are especially effective at gathering local and global contextual information, which is crucial for identifying subtle abnormalities in chest X-rays, such as nodules, consolidations, or opacities. Furthermore, the proposed model integrates an optimised feature selection mechanism and an improved classification module, which together enhance the discriminative power of CNN-based representations. In contrast to conventional techniques and previous CNN variations, our improved CNN-based method ensures consistent performance across diverse imaging conditions, reduces false diagnoses, and enhances classification accuracy. The framework exhibits strong generalisation after being trained and validated on a sizable, well-annotated dataset, providing physicians with a diagnostic tool that is more precise, effective, and interpretable.

The significant contributions of the proposed research are:

- A novel hybrid framework is proposed that combines an efficient feature extraction mechanism using MobileNetV2 with an advanced classification model (Enhanced GoogLeNet) tailored for accurate multi-class classification of chest X-ray images.
- A new meta-heuristic feature selection algorithm, named IPDO, has been developed to optimise the high-dimensional feature space extracted from images. By selecting the most pertinent features with the least computational overhead, the IPDO algorithm significantly reduces redundancy and improves classification performance.
- An enhanced variant of the GoogLeNet model (EGoogLeNet) is designed and integrated into the framework, providing enhanced discriminative power and deeper feature representation for fine-grained classification across multiple lung disease classes, including COVID-19, normal, lung opacity, and viral pneumonia cases.
- Using the extensive publicly accessible C19RD dataset, extensive experimental validation is carried out. The proposed model yields higher performance results, demonstrating its superiority in both scientific research and practical applications.

The remaining part of this research is structured as follows: recent research-related works on chest X-ray image-based deep learning for disease detection and classification are presented in Section 2. Section 3 describes the methodology. The experimental results and comparison details are included in Section 4. Section 5 concludes the proposed research.

Problem Statement

Chest X-ray (CXR) imaging is one of the most widely used diagnostic tools for detecting a variety of lung diseases due to its cost-effectiveness, speed, and non-invasiveness. However, there are several significant challenges to automatically interpreting CXR images. These include overlapping tissues, low contrast across anatomical structures, slight variations in visual patterns among various lung diseases, and varied image capture settings. Even experienced radiologists find it challenging to reliably distinguish between diseases such as COVID-19, lung opacities, pneumonia, and normal lung conditions due to these complications.

Standard Convolutional Neural Networks (CNNs), while successful in binary classification tasks, often struggle with multi-class classification of lung diseases from CXR images. For number of classes, their accuracy decreases due to their limited capacity to capture contextual connections and fine-grained spatial data. Furthermore, issues such as uneven data distribution, intra-class similarity, and inter-class variation, all of which are prevalent in real-world CXR datasets, may not be adequately handled by traditional CNNs.

An improved deep learning framework that can reliably and accurately diagnose several types of lung disease from CXR images is urgently needed for advancement, in addition to addressing these limitations. The proposed research introduces a novel enhanced CNN-based architecture, specifically designed to handle the complex nature of CXR images and the challenges of multi-class disease classification. The primary objective is to improve early diagnosis, treatment planning, and patient outcomes by providing a diagnostic assistance system that is more reliable and practical.

Related Prior Works

To contextualise the novelty of our proposed framework, this section thoroughly examines previous contributions in the field, with an emphasis on architectural improvements, optimisation techniques, attention mechanisms, feature fusion, and lightweight deployment.

Convolutional Neural Networks with Architectural Enhancements

Ali et al. (2024) introduced CX-RaysNet, combining standard and group convolutional layers with small filters and dropout regularisation. This design enhanced subtle pattern recognition in CXR images, achieving 97.25% multi-class accuracy. However, it focussed

primarily on low-power device deployment, lacking a mechanism for dimensionality reduction of high-level features, which limits scalability in high-dimensional feature scenarios.

Similarly, XcepCovidNet, as described by Juneja et al. (2024), utilised transfer learning with hyperparameter tuning and outperformed multiple pre-trained models. While effective for COVID-19 detection, it did not introduce a novel feature selection or optimisation mechanism, limiting model interpretability and customisation.

By combining an improved feature extraction model (MobileNetV2) with a unique IPDO technique, which not only selects the most pertinent features but also reduces computational cost, our proposed model builds upon these studies.

Hybrid Models using Wavelet Transforms and MBConv Blocks

According to Bhosale and Yadav (2024), the wavelet features improved frequency-based analysis for five-class classification. Although their model achieved 95.5% accuracy, it required substantial computational resources and lacked effective mechanisms for model interpretability.

For optimal feature dimensionality reduction, our model employs IPDO, a metaheuristic feature selector, which reduces computational load while maintaining performance.

Attention Mechanisms and Ensemble Learning

Sayeed et al. (2024) utilised attention-based ensembles of Xception, EfficientNetB7, and DenseNet201, integrated with a Residual U-Net for segmentation. Although this approach achieved high accuracy (98.21% for binary), its complexity and ensemble design increased inference time and model overhead.

Our proposed method achieves better accuracy with significantly less complexity and training time by utilising a single upgraded classification model (EGoogLeNet) with Adam optimization and feature optimisation using IPDO.

Optimisation and Explainability in Deep Models

Abhishek et al. (2025) proposed DICA-Net, using Attention U-Net for segmentation and Pigeon Local Search Optimisation (PLSO) for feature selection. With an accuracy of 99.21%, it is one of the best-performing models. However, the optimisation algorithm is tailored to a specific task and may not generalise across diverse datasets.

Inspired by the behaviour of prairie dogs, our IPDO algorithm offers greater adaptability, enabling faster convergence and wider applicability across various lung diseases.

Balaha et al. (2025) introduced LDD-VTS, a ViT-based model with SHAP for explainability. Although interpretability is improved, the model relies on transformer architectures, which can be computationally intensive and less efficient for smaller datasets.

With a CNN-based architecture and lightweight feature selection, we prioritise efficiency and performance balance to enable real-time applications without compromising accuracy in our proposed research.

Feature Enhancement and Transformer-based Models

Egala et al. (2025b) employed a Stacked Residual Coordinate Network and innovative techniques, including the Gaussian Fourier Pyramid, Termite Alate Optimisation, and a Multi-GAN Transformer. While it achieved 99% accuracy, the multi-layer and transformer-based approach makes deployment on standard hardware more difficult.

In contrast, our approach utilises a modified MobileNetV2 + EGoogLeNet architecture, which is optimised for both training efficiency and deployment scalability.

CNNs with SE Blocks, Capsule Networks, and Hybrid Pipelines

Shanthi et al. (2025) presented a model combining SENet attention and capsule networks with CBSO optimisation. The system achieves 99% accuracy on the COVIDx CXR-4 dataset, primarily optimised for COVID-related data. Similarly, Zhu et al. (2025) combined VGG16 with self-attention and SE blocks, but their model was limited to six disease categories with an accuracy of 93%.

In contrast, our model is verified on the larger C19RD dataset and achieves 99.41% accuracy, outperforming many of the research studies above. It handles multi-class classification of lung diseases.

It is clear from the review of the literature that recent developments in deep learning for the classification of chest X-ray images have investigated a range of strategies, such as attention mechanisms, hybrid frameworks (e.g., CNN-transformers, CNN-wavelet), architectural modifications (e.g., group convolutions, MBConv blocks), and optimisation techniques. Although many of these models have produced encouraging outcomes, several drawbacks remain. Notably, most methods either lack an efficient and scalable feature selection mechanism or rely heavily on computationally expensive architectures, which can hinder real-time deployment and generalisability across diverse datasets.

These findings underscore the need for a classification framework that can handle high-dimensional image features and is not only accurate but also interpretable, computationally efficient, and effective. By combining MobileNetV2 for efficient and lightweight feature extraction, an effective IPDO algorithm for feature selection and dimensionality reduction that enhances model robustness and efficiency, an EGoogLeNet model designed for high-precision multi-class classification, and performance validation on a large-scale dataset (C19RD), our proposed approach directly solves these gaps. It also shows superior accuracy and generalisation when compared to other methods.

As a result, the literature analysis highlights the importance and necessity of our proposed method, establishing it as a significant advance in automated lung disease diagnosis utilising CXR imaging from both scientific and practical perspectives.

MATERIALS AND METHODS

The overall architecture of the proposed enhanced deep learning framework is illustrated in Figure 1. The following processes are included in a traditional framework for predicting and classifying lung disease using chest X-ray images: data collection, pre-processing and augmentation, feature extraction, feature selection, and classification. The spatial and channel attention mechanisms are combined with EGoogLeNet as the primary method for classifying diseases. For experimental analysis, a large multi-class chest X-ray dataset for lung disease is used.

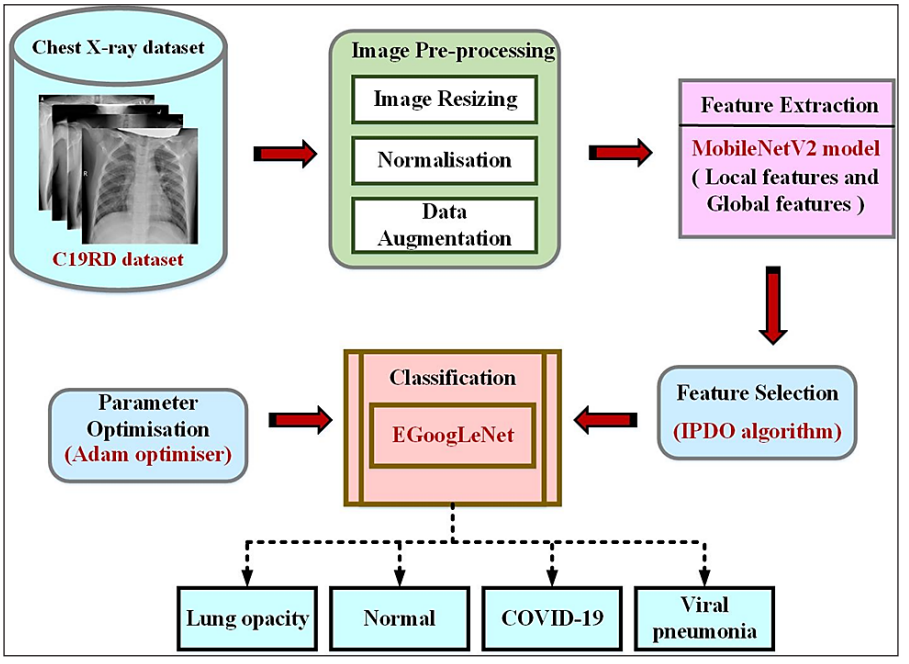

Figure 1. Overall architecture of the proposed enhanced deep learning framework

Database Description

The COVID-19 Radiography Database (C19RD) served as the main experimental collection in our research (Patel et al., 2024). This extensive collection comprises 21,165 chest X-rays, covering a variety of cases. The images of COVID-19-positive cases total 3,616, images of normal lung conditions indicate 10,192, and images of viral pneumonia

show 1,345. Representative sample chest X-ray images from the C19RD dataset used in this research are shown in Figure 2. The C19RD dataset is publicly available on Kaggle and was used in this study under its permissible open-access research license (Kaggle, 2024).

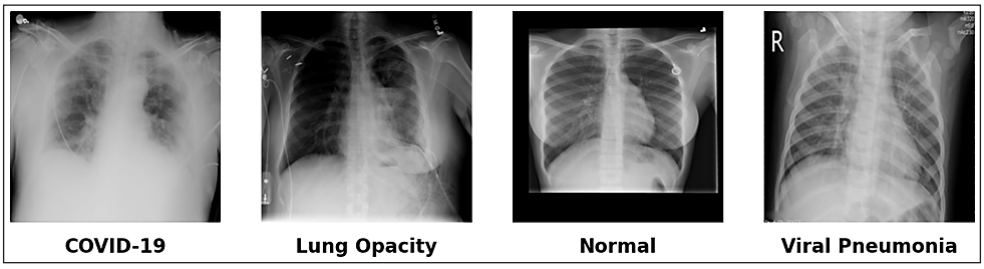

Figure 2. Representative sample images from the C19RD dataset showing (a) COVID-19; (b) lung opacity; (c) normal; and (d) viral pneumonia classes

Data Splitting

According to our research, the dataset was split into three sections: 60% was used for training, 20% for validation, and the remaining 20% for testing. Table 1 presents the distribution of the dataset before data augmentation across the training, validation, and testing partitions.

Table 1
Distribution of the dataset before data augmentation

Class Name	Training	Validation	Testing
Normal	6516	1629	2047
COVID-19	2291	572	753
Viral Pneumonia	867	216	262
Opacity	3833	958	1221
Total	13507	3375	4283

Image Pre-processing and Augmentation

To guarantee that the CNN architecture is compatible, image pre-processing is essential. Each image in the collection is initially resized to a consistent 224 × 224 pixels size. Additionally, image augmentation methods are employed to address the issue of the training set's limited number of images and enhance training effectiveness. Standard data augmentation techniques are used to rectify the training dataset's class imbalance. Even though the entire dataset contains 21,165 chest X-ray images, there is a notable disparity in the distribution of classes, particularly in the training set. For example, classifications such as "Viral Pneumonia" and "COVID-19" have far fewer samples (867 and 2,291, respectively), whereas the "Normal" class comprises 6,516 images. This imbalance can lead

to biased model learning, where the network favours the overrepresented class, reducing its generalisation ability and performance on minority classes. To artificially enhance the number of minority class samples, data augmentation was applied to the training set. This made each class roughly comparable to the "Normal" class, which consisted of about 6,500 images. Moreover, although the overall dataset size is substantial, data augmentation ensures robustness by simulating variability and preventing overfitting, especially since the proposed model is designed for high sensitivity across multiple lung disease categories.

Table 2 summarises the data augmentation techniques applied to the training dataset, along with their parameter ranges for balancing classes and improving model generalization. The augmentation, normalisation, and resizing procedures work together to maximise training efficiency and enhance the overall performance and reliability of the CNN model for lung disease diagnosis. Examples of the applied data augmentation techniques used to enhance dataset diversity are presented in Figure 3.

Table 2
Data augmentation techniques and parameter ranges used in the training dataset

Parameter	Value
Shear	0.35
Height shift	0.25
Zoom	(0.4, 0.9)
Rotation	(+30, −30)
Width shift	(0.7, 1.25)
Brightness Range	(0.5, 1.30)
Fill mode	Nearest
Horizontal flip	TRUE

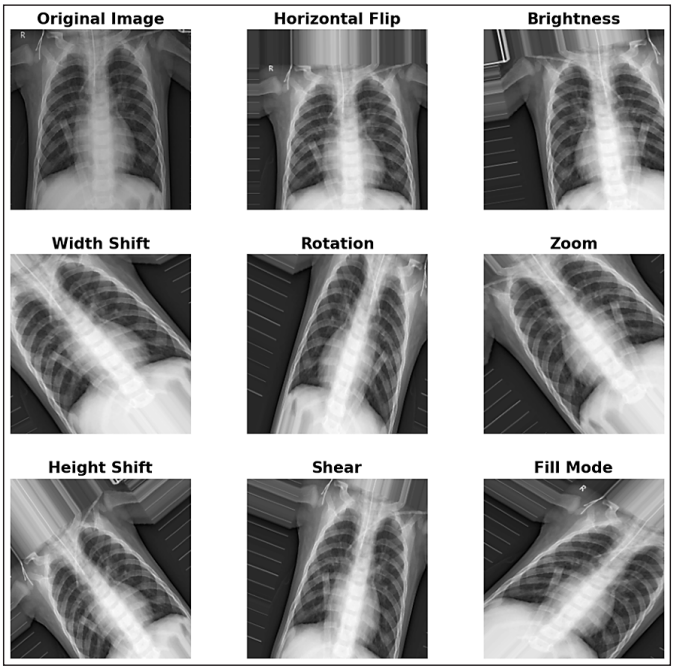

Figure 3. Example augmented images illustrating the application of data augmentation techniques to increase dataset variability

Feature Extraction

Due to its effective depth-wise separable convolutions, lightweight design, and robust performance on embedded and mobile vision tasks, MobileNetV2 is utilised for feature extraction in the proposed research. The MobileNetV2 model utilises Inverted Residuals with Linear Bottlenecks, which enables quicker computation and reduced memory consumption without sacrificing accuracy. When working with large amounts of medical imaging data, such as MRI, CT, and X-ray images, this is especially helpful. Additionally, MobileNetV2 achieves a balance between model accuracy and complexity, preventing overfitting and facilitating efficient feature learning from small datasets. It is a good fit for the proposed classification challenge due to its demonstrated ability to extract reliable and discriminative features from medical images.

In this research, the MobileNetV2 model is utilised to extract both local and global features of chest X-ray images, aiming to enhance classification performance. The MobileNetV2 model improves performance across a range of model sizes on various tasks and benchmarks. MobileNet models employ an excellent concept, utilising depth-wise separable convolutional blocks in place of expensive convolutional layers.

Feature Selection

In this research, the IPDO is employed to reduce the dimensionality of features. A metaheuristic algorithm that simulates prairie dog foraging is called the IPDO algorithm. Every day, prairie dogs participate in social activities such as foraging, cavern construction, cave maintenance, and predator defence. Thus, four-time intervals comprise the IPDO algorithm, based on the routines of prairie dogs. Based on a fixed mirror lifestyle, we then divide exploration and exploitation. In the spatial dimension, each prairie dog's foraging activity is represented by a $1 \times dim$ vector. Prairie dogs' mobility range is limited by the upper-bound UB and lower-bound LB, which prevents them from straying from their trajectory when foraging. A difficulty can be solved by placing each prairie dog in its own set at various places.

Classification

This research presents supervised deep learning for multi-class Classification of the most prevalent chest diseases. The multi-class chest X-ray images are classified using an enhanced GoogLeNet (EGoogLeNet) model.

Among the most widely used CNN models for image detection and classification is the GoogLeNet model, which has the advantages of low parameters and high classification and recognition accuracy. GoogLeNet successfully captures the features of different scales by using pooling layers of varying widths and parallel multiple convolution kernels. It presents techniques that facilitate network training and optimisation, including the

structure of Inception, design of depth and width, parameter efficiency, average pooling, and auxiliary classifiers. Figure 4 represents the standard Inception module introduced in the GoogLeNet (Inception v1) architecture. It is a powerful CNN model that successfully increases the network's capacity for classification, accuracy, and computational efficiency.

As per the attention mechanism, an enhanced GoogLeNet network model is proposed, building upon the GoogLeNet model to classify chest X-ray images precisely. By discovering the weights of each channel and space, the network can concentrate more attention on the information. It is of greater importance for the process of classifying images due to the attention mechanism, as different disease types correlate to separate locations in the X-ray image feature map. As a result, the attention mechanism selection enhances the network's capacity to adjust to variations in different images, allowing for more emphasis on the features that are most suitable for the classification process within the network. This enhances the model's capacity for generalisation and recognition accuracy.

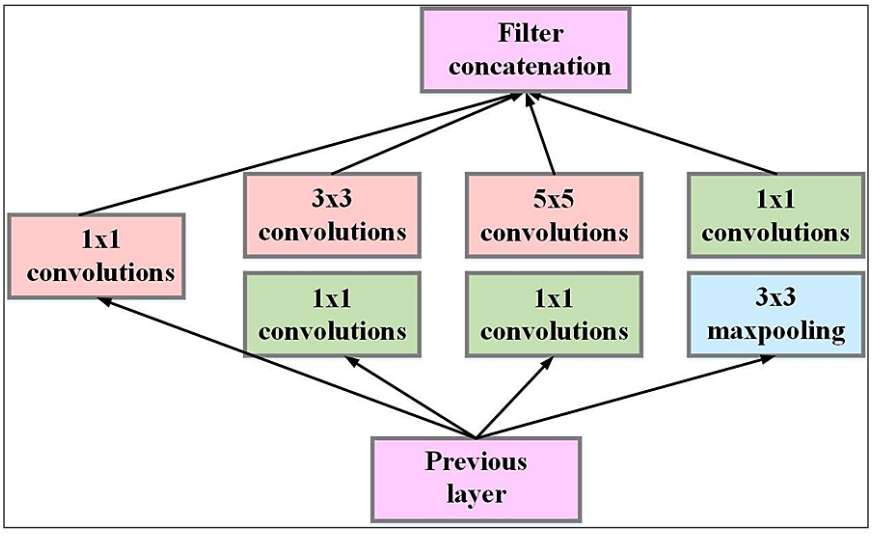

Figure 4. Inception module structure used in the GoogLeNet architecture

GoogLeNet is used by the model as the primary network. Prior to the input image reaching GoogLeNet's first convolutional layer, a weighted feature map is created by introducing a spatial attention mechanism. Following the final GoogLeNet stage, the weighted feature representations are routed to the GoogLeNet model's first convolutional layer and processed. The Channel Attention Mechanism module follows the Inception module. By combining convolution processes of various kernel sizes (1×1, 3×3, and 5×5) in parallel, the Inception module significantly improves multi-scale feature extraction in this situation. Because disease patterns vary in size and shape, a model is needed that can simultaneously capture both local and global features, which is especially useful in

medical imaging. The Inception module ensures that a rich collection of spatial features is first collected when combined with the Channel Attention Mechanism. Then, by emphasising the most relevant channels, the attention module can refine these features. The model's capacity to identify intricate and delicate lung disease patterns in chest X-rays is enhanced by this sequential structure, which improves classification accuracy. On the feature maps processed, the different channels' weights are modified adaptively, which is the main purpose of the GoogleNet model and the spatial attention mechanism in the network, prioritising the characteristics pertinent to the categorisation task. Additionally, this reduces the impact of redundant features, thereby enhancing the network's ability to adapt. Additionally, by reducing the network parameters and the model's complexity, this attention mechanism is used.

Figure 5 illustrates our proposed EGoogLeNet architecture, which incorporates spatial and channel attention mechanisms. In the proposed EGoogLeNet model, both the spatial and channel attention mechanisms are applied sequentially to enhance the feature map before it is passed into the main classification layers. Immediately following the processing of the input features by the first convolutional and pooling layers, the spatial attention mechanism is implemented. By highlighting the locations of the pertinent features in the image, this method calculates spatial weights that enable the model to focus on significant areas within the X-ray image.

The output of the spatial attention module is then subjected to the channel attention mechanism. To help the model concentrate on the most essential pathological features for classification, this technique adaptively recalibrates the relevance of each feature channel.

These two attention mechanisms are processed in a cascaded manner, with spatial attention followed by channel attention, which allows the model to learn the most essential disease feature maps. The spatial and channel weighting are combined into a single refined feature representation by fusing the feature maps produced by the two methods using element-wise multiplication as illustrated in Figure 5. For final classification, this fused attention-enhanced feature map is subsequently sent via the GoogLeNet architecture's final layers.

Attention Mechanism

The poor performance of direct feature map recognition can be attributed to several issues, including the image's low resolution and the slight variation in the tumour region across various types of chest X-ray images. Since different types of lung diseases have distinct characteristics, the classification model must determine the importance of image information and where it should be placed in the image to create a more useful feature map by assigning distinct weights to each component of the input feature map in the GoogLeNet model.

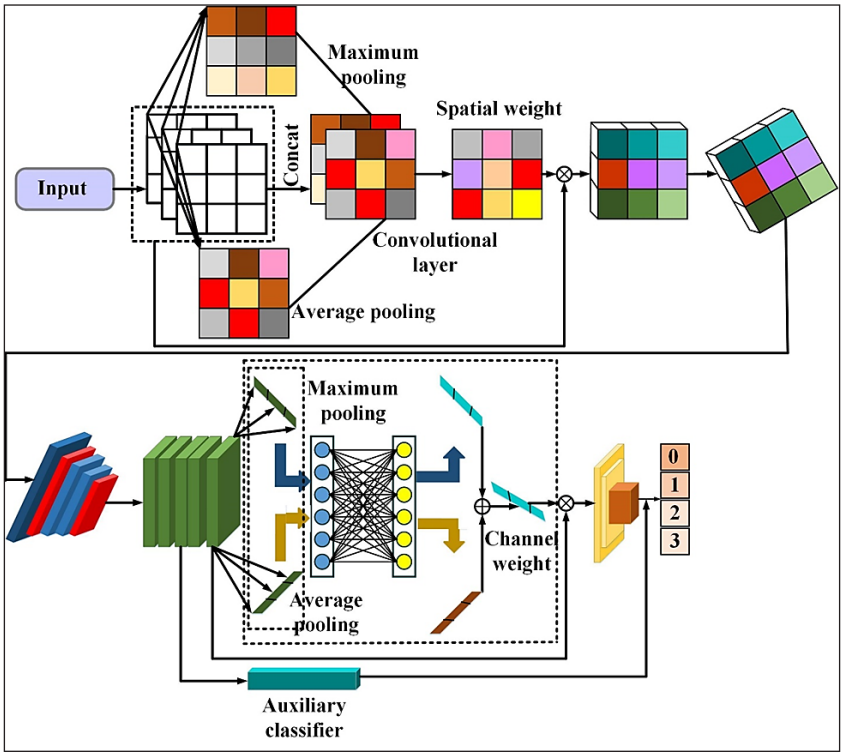

Figure 5. The attention mechanism-based general architecture of the EGoogLeNet model

The proposed EGoogLeNet design enhances the model's focus on the most important features, both spatially and across channels, by integrating the spatial and channel attention mechanisms sequentially to create a single attention module. There are two steps in the fusion process. The spatial attention mechanism is first applied to the input feature map, which highlights the locations of the image's most informative parts. To accomplish this, global max pooling and average pooling are applied along the channel dimension. The spatial attention map is then created by concatenating the data and feeding them into a convolutional layer, which is followed by a sigmoid activation. The original feature map and this spatial attention map are then multiplied element by element to highlight significant spatial regions.

Then, the output from the spatial attention is fed into the channel attention mechanism, which learns what features (channels) are most important for classification. Across spatial dimensions, it applies global max pooling and average pooling, and feeds a shared MLP with the combined features, using a ReLU activation. A sigmoid function is applied to the output to compute channel-wise weights. These weights are multiplied by the input feature map to enhance the discriminative features.

Experimental Settings

Table 3 presents the detailed computational environment used to implement and evaluate the proposed EGoogLeNet-based chest disease classification framework. We used the Python programming language.

Table 3
Summary of the computational environment and system configuration used to implement and evaluate the proposed EGoogLeNet framework

Name	Values/Types
Language	Python
GPU	NVIDIA RTX 3050 8 GB
RAM	16 GB
Operating System	Windows 10 Pro

To ensure appropriate training without excessive overfitting, our model was trained for over 100 epochs. The adaptability of the Adam optimiser in modifying network weights led to its selection. The model was able to converge gradually by using a 0.0001 learning rate and a 32 mini-batch size, which were chosen for their superior convergence qualities. Cross-entropy was the loss function employed, a common choice for classification problems.

RESULTS AND DISCUSSION

The experimental results and discussion on the multi-disease classification model based on the EGoogLeNet model are presented in this section.

Evaluation Metrics

Within the field of deep learning, model evaluation for a specific task depends on specific quality metrics. Recall, F1-score, precision, and accuracy provide data on the model's overall performance.

Equation 1 defines accuracy as the proportion of effectively classified images in the model to the total number of evaluated images. Equation 2 defines precision as the percentage of real positives in the subset of data the model has classified as positive. Equation 3 provides recall, also known as sensitivity, which measures the model's precision in identifying all pertinent samples within the true positive category. Furthermore, as demonstrated by Equation 4, the F1-score is a composite metric that balances recall and accuracy to provide a single measure of the classifier's accuracy.

$$Accuracy = \frac{(TN+TP)}{(TN+FP+FN+TP)} \quad [1]$$

$$Pr\,e\,cision = \frac{TP}{(FP+TP)} \quad [2]$$

$$Re\,c\,all = \frac{TP}{(FN+TP)} \quad [3]$$

$$F_1 - Score = 2 \times \frac{(Pr\,ecision \times Re\,call)}{(Pr\,ecision + Re\,call)} \quad [4]$$

where TN: True Negative, TP: True Positive, FN: False Negative, and FP: False Positive.

Training and Validation Evaluation

The training and validation accuracy and loss curves of the proposed EGoogLeNet model are depicted in Figure 6. The accuracy curve (Figure 6a) shows a sharp improvement over the first 15 epochs, with training accuracy rising from about 70% to over 90%. The model's great generalisation capacity and low overfitting are demonstrated by the training and validation accuracies, which progressively converge after 20 epochs and reach a peak of approximately 99%. The model appears to maintain constant performance throughout unknown data, as seen by the close relationship between the validation accuracy and the training accuracy.

During the early training phase, the training loss drops sharply from above 0.50 to below 0.15 in the first 20 epochs, as shown by the loss curve (Figure 6b). This pattern shows effective model parameter optimisation using the Adam optimiser and is consistent with the fast accuracy gains. After 60 epochs, the validation loss stabilises at a level below 0.10, closely tracking the training loss throughout the process. Even after extensive training over 100 epochs, the EGoogLeNet model does not experience overfitting, as evidenced by the lack of discernible divergence between the two loss curves.

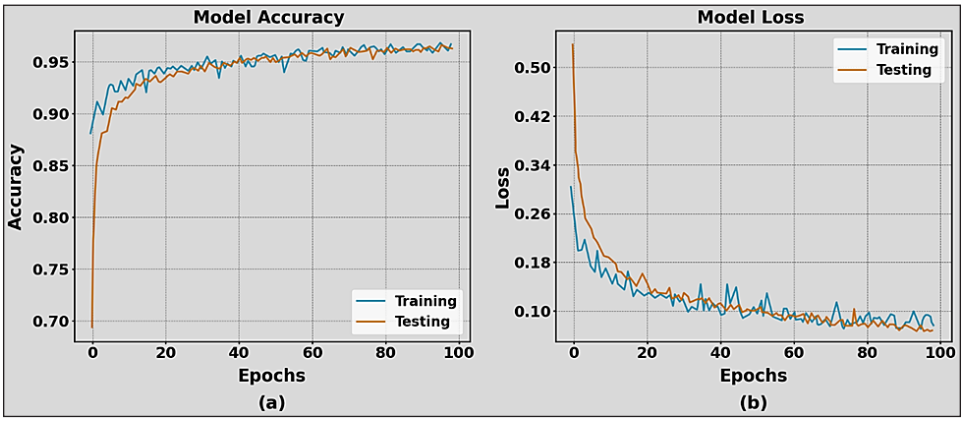

Figure 6. Training and validation accuracy and loss curves of the proposed EGoogLeNet model

These findings support the efficiency of the proposed hybrid framework, which combines the IPDO algorithm with MobileNetV2-based feature extraction to provide an optimal feature set that enhances classification accuracy and accelerates convergence. The proposed method's resilience for accurate and dependable multi-class lung disease diagnosis is further supported by the final stabilised high accuracy (99.41%) and low loss values across both training and validation stages.

Evaluation Results

The test set results are used to evaluate the proposed research's performance and compare the findings with those of recent existing deep learning models.

The classification results of the proposed EGoogLeNet model are summarised in Table 4. On the C19RD test set, the proposed EGoogLeNet model demonstrated a remarkable capacity to distinguish between multiple lung disease stages, achieving an overall classification accuracy of 99.41%. Precision and recall values are consistently high across all classes, as shown in the classification report, indicating that there are few false positives and false negatives. COVID-19 cases were specifically detected with 0.98 precision and 0.99 recall, guaranteeing dependable detection with minimal misclassification. Additionally, cases of lung opacity and viral pneumonia had a 0.99 recall, demonstrating the model's ability to identify subtle radiographic abnormalities that are frequently difficult for human interpretation.

The model can accurately identify healthy chest X-rays while minimising unnecessary false positives, as demonstrated by the Normal class's high precision of 0.99 and recall of 0.98. Even with differences in class size, the weighted-average precision of 0.99 and the macro-average precision of 1.00 show equal performance across all categories. The stability and dependability of the classification findings are confirmed by the consistently high F1-scores (0.98–0.99) for each class. The proposed improved deep learning model for multi-class chest disease classification yields an overall precision of 99.50%, recall of 99.50%, and F1-score of 99.25%. The model's promise as a reliable decision-support tool for fast and precise multi-class chest disease diagnosis in clinical settings is highlighted by its high precision and recall across all disease categories.

We conducted a series of controlled experiments by varying each hyperparameter within its tested range and monitoring the corresponding classification performance of the proposed EGoogLeNet model, thereby supporting the justification summary in Table 4. The outcomes demonstrated that, when compared to other settings, the ideal values indicated in Table 4 consistently yielded higher accuracy, F1-score, and convergence stability. For example, increasing the learning rate beyond 0.0001 resulted in unstable training with noticeable fluctuations in validation accuracy, while lower values slowed convergence without improving final performance. Similarly, batch sizes above 32 led to suboptimal

F1-scores due to the underrepresentation of minority classes within each batch, whereas sizes below 16 produced noisy gradients and slower optimisation. The optimal epoch value (100) was chosen because the validation accuracy reached saturation at this point, and further training did not yield additional improvement. These experimental observations validate the selection of the optimal hyperparameters presented in Table 4.

The class-wise prediction performance of the proposed EGoogLeNet model is illustrated through the confusion matrix in Figure 7. The matrix's diagonal dominance indicates that the great majority of samples were appropriately categorised. In 745 out of 753 cases,

Table 4
Classification results of the proposed EGoogLeNet model showing F1-score, recall, precision, and overall accuracy

Class	Precision	Recall	F1-score	Support
COVID-19	0.98	0.99	0.98	753
Lung Opacity	0.99	0.99	0.99	1221
Normal	0.99	0.98	0.99	2047
Viral Pneumonia	0.98	0.99	0.99	262
Accuracy	-	-	0.9941	4232
Macro Avg	1	0.99	0.99	4232
Weighted Avg	0.99	0.99	0.99	4232

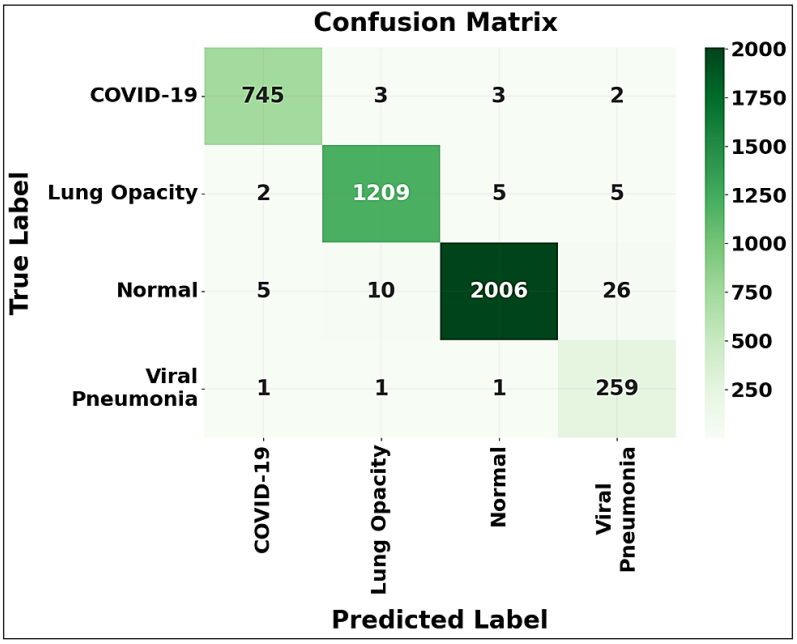

Figure 7. Confusion matrix illustrating the class-wise prediction performance of the proposed EGoogLeNet model

COVID-19 patients were accurately recognised, with only slight misclassifications into the categories of lung opacity, normal, and viral pneumonia. Lung Opacity achieved similarly high accuracy, with 1,209 correct predictions out of 1,221 cases, and only a few samples were misclassified into other classes.

For the Normal category, 2,006 out of 2,047 instances were correctly recognised, with the small number of errors primarily involving misclassification into Viral Pneumonia, likely due to overlapping radiographic features. Viral pneumonia exhibited the smallest sample size but still achieved 259 correct predictions out of 262 cases, showing the model's robustness even with limited training data.

These results, combined with the high recall, precision, and F1-scores, confirm the proposed hybrid framework's efficiency in learning highly discriminative features for each category. The minimal off-diagonal values indicate that integrating MobileNetV2 feature extraction, IPDO feature selection, and the EGoogLeNet classifier enables precise differentiation between visually similar lung disease patterns, making it well-suited for practical diagnostic applications.

The Receiver Operating Characteristic (ROC) curves and AUC values for each disease category obtained using the proposed EGoogLeNet model are shown in Figure 8. Each class has exceptionally high Area Under the Curve (AUC) values; COVID-19 achieved an AUC of 0.9973, indicating a nearly perfect trade-off between sensitivity and specificity. Pneumonia and Normal classes also exhibit strong separability, with AUC scores of 0.9959 and 0.9906, respectively. Despite being a challenging category due to its visual resemblance to other diseases, lung opacity achieved an AUC of 0.9863, confirming the robustness of the model's learned features.

For all classes, the ROC curves' sharp rise toward the upper-left corner indicates high true favourable rates and few false positives. The precision and recall scores of the classification report are consistent with these high AUC values, demonstrating the effectiveness of the MobileNetV2 feature extraction and IPDO feature selection process in producing highly separable feature representations. The proposed framework is a dependable tool for multi-class chest disease identification in actual clinical settings, as the ROC analysis demonstrates that it consistently maintains robust decision boundaries across various disease types.

Sample prediction outputs generated by the proposed EGoogLeNet model on test images are displayed in Figure 9. The green labels indicate correctly classified images, while red labels show misclassified samples. Several chest X-ray samples from each class are displayed to help visualise the proposed model's classification performance. The model's capacity to identify pertinent patterns for each disease category is demonstrated by the fact that most images are correctly labelled.

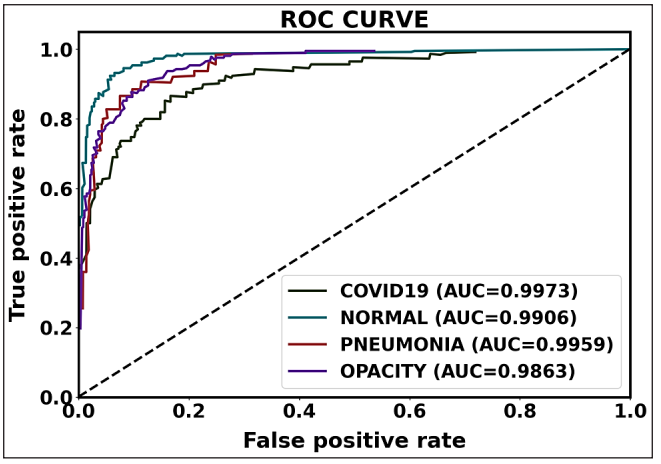

Figure 8. ROC curves with AUC values for each class obtained using the proposed EGoogLeNet model

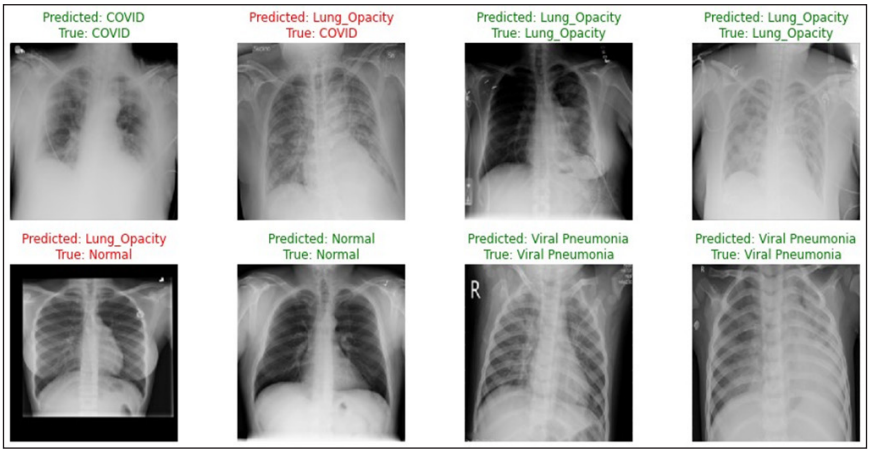

Figure 9. Sample predictions generated by the proposed EGoogLeNet model on the test images

Performance Comparison

Table 5 summarises a comprehensive performance comparison between the proposed EGoogLeNet framework and several recent state-of-the-art deep learning models for multi-class lung disease classification. The results clearly demonstrate that the proposed model outperforms the baseline on each evaluation parameter. In particular, the proposed model achieves an overall accuracy of 99.41%, significantly surpassing popular architectures such as CX-RaysNet (97.20%) and EfficientNet-B4 (96%). With F1-score, recall, and precision of 99.25%, 99.50%, and 99.50%, respectively, our proposed framework outperforms the nearest competitor (DenseNet-201) by significant margins.

Table 5
Comparative performance analysis of the proposed EGoogLeNet model against recent state-of-the-art approaches for multi-class lung disease classification

Reference	Classes	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
(Patel et al., 2024)	5	EfficientNet-B4	96	93	95.6	94.6
(Ali et al., 2024)	5	CX-RaysNet	97.20	97.02	96.43	97.25
(Thavasimuthu et al., 2024)	5	Multi-layer Perceptron Neural Network (MPNN)	92	95	97	98
(Karaddi et al., 2024)	6	Modified ResNet50	-	-	-	98.03
(Alshmrani et al., 2023)	5	CNN	96.48	97.56	93.75	-
Proposed Approach	4	EGoogLeNet	99.41	99.50	99.50	99.25

The IPDO algorithm, which minimises redundancy and improves class separability, and MobileNetV2, which enables effective and discriminative feature extraction, are responsible for the improvement. Furthermore, the EGoogLeNet classifier produces high-confidence predictions in all four categories by efficiently utilising the revised feature space.

Our framework achieves superior performance while maintaining computational efficiency when compared to models that use deeper or more sophisticated backbones (e.g., EfficientNet-B4, XceptionNet). This indicates that the proposed hybrid model not only performs better in terms of prediction accuracy than existing models, but it also holds great promise for real-world clinical use, where accuracy and inference speed are crucial.

Discussion

With an overall accuracy of 99.41%, the proposed EGoogLeNet-based multi-class classification framework shows very promising performance in recognising COVID-19, Lung Opacity, Normal, and Viral Pneumonia cases from chest X-ray images. The improved resilience and generalisation capability of the model are mainly due to the better pre-processing pipeline, multi-level feature extraction, and effective classification of abnormal patterns.

Even if the performance was excellent, some restrictions must be noted. The evaluation is restricted to the C19RD dataset, which may limit generalisability to other populations and imaging situations, as it may not accurately reflect the variances among multi-institutional clinical environments. Furthermore, despite the model's excellent discriminative capacity, its real-world interpretability remains limited. To boost clinical trust, explainability approaches, such as Grad-CAM or SHAP, must be further integrated. The diagnostic depth of the model may be limited by the lack of multi-modal imaging and the exclusive use of chest X-ray data, particularly in complex instances where CT or clinical metadata could provide complementary insights.

However, the results of this research have significant implications for both clinical practice and future research. The proposed model demonstrates potential for deployment as a computer-aided diagnosis tool to support radiologists, particularly in resource-limited settings, and can enhance triage efficiency during large-scale screening scenarios such as respiratory outbreaks. Expanding validation across various cohorts, integrating multi-modal data fusion for improved diagnosis accuracy, and incorporating explainable AI frameworks to enhance model transparency and clinical acceptance should be the primary goals of future research.

CONCLUSION

Using chest X-ray images, this research classified multiple lung diseases. An enhanced deep learning architecture was the foundation for the system's development and evaluation. The proposed research consists of four primary phases. To improve the dataset's quality, the initial pre-processing step handles image scaling, normalisation, and data augmentation techniques. Using the MobileNetV2 model, the second phase extracts the effective local and global features relevant to the disease. To reduce network complexity, important disease features are selected in the third phase using the IPDO algorithm. Finally, the EGoogLeNet model is used to classify the multi-class chest X-ray images. Our evaluation process on a large-scale C19RD dataset demonstrates the framework's outstanding potential, outperforming other recent state-of-the-art techniques for lung disease classification. The 99.41% accuracy achieved is evidence of the framework's accuracy and dependability. The framework performed exceptionally well, showing superior accuracy on key metrics such as AUC, F1-score, recall, and precision, with average values of 0.9925, 0.9925, 0.9950, and 0.9950, respectively.

We plan to expand this research in multiple ways in the future by utilising GANs to generate a more representative and extensive range of artificial medical images, improving the approaches' ability to be applied to several patient populations, combining it with electronic health records and additional diagnostic instruments to enable more comprehensive patient evaluations, and extending its use to cover further respiratory and thoracic diseases.

ACKNOWLEDGEMENT

We are grateful to everyone who has supported and encouraged us during this research. Your invaluable advice and assistance have been instrumental in completing this work.

REFERENCES

Abhishek, G. S. V. R., Singla, A., & Eshwar, D. (2025). DICA-Net: Optimising chest X-ray classification with attention U-Net and pigeon local search. *Network Modeling Analysis in Health Informatics and Bioinformatics, 14*(1), Article 20. https://doi.org/10.1007/s13721-025-00513-5

Ali, A., Wang, Y., & Shi, X. (2024). Detection of multi-class lung diseases based on customised neural network. *Computational Intelligence, 40*(2), e12649. https://doi.org/10.1111/coin.12649

Alshmrani, G. M. M., Ni, Q., Jiang, R., Pervaiz, H., & Elshennawy, N. M. (2023). A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images. *Alexandria Engineering Journal, 64*, 923-935.https://doi.org/10.1016/j.aej.2022.10.053

Balaha, H. M., Ahmed, R. A., & Balaha, M. H. (2025). LDD-VTS: An AI-based framework for lung disease diagnosis using vision transformers and SHAP. *Health Information Science and Systems, 13*(1), Article 47. https://doi.org/10.1007/s13755-025-00363-5

Bhardwaj, P., & Kaur, A. (2025). DLSAC-Net: An automated enhanced segmentation and classification network for lung diseases detection using chest X-ray images. *Multimedia Tools and Applications*, 1-26. https://doi.org/10.1007/s11042-025-20763-9

Bhosale, R. D., & Yadav, D. M. (2024). Customised convolutional neural network for pulmonary multi-disease classification using chest X-ray images. *Multimedia Tools and Applications, 83*(6), 18537-18571. https://doi.org/10.1007/s11042-023-16297-7

Chen, C., Isa, N. A. M., Liu, X., Ding, J., & Lu, L. (2026). MSA-Net: Multi-scale attention-based DenseNet for multi-label chest X-ray image classification. *Biomedical Signal Processing and Control, 113*, Article 109069. https://doi.org/10.1016/j.bspc.2025.109069

Egala, R., & Sairam, M. V. S. (2025a). Multi-instance Riemannian residual neural network with mountaineering team-based chest disease detection using chest X-ray images. *Research on Biomedical Engineering, 41*(3), Article 44. https://doi.org/10.1007/s42600-025-00423-5

Egala, R., & Sairam, M. V. S. (2025b). Multi-layer stacked residual coordinate termite alate network for multi-class lung diseases detection from chest X-ray images. *Applied Soft Computing*, Article 113393. https://doi.org/10.1016/j.asoc.2025.113393

Gonçalves, L. A., Junior, G. B., Bessa, M. L., Matos, C. E., & Fernandes, A. G. (2025). DualAttentionNet: A convolutional neural network for thoracic disease classification in chest X-rays. *Procedia Computer Science, 256*, 797-804. https://doi.org/10.1016/j.procs.2025.02.181

Juneja, A., Kumar, V., Kaur, M., Singh, D., & Lee, H. N. (2024). XcepCovidNet: Deep neural networks-based COVID-19 diagnosis. *Multimedia Tools and Applications, 83*(37), 85195-85225. https://doi.org/10.1007/s11042-024-19046-6

Kaggle. (2024). *COVID-19 radiography dataset*.

Karaddi, S. H., & Sharma, L. D. (2024). Classification of lung disorders in chest multi-modal images using hyper-parameter tuning and modified ResNet50. *Multimedia Tools and Applications*, 1-25. https://doi.org/10.1007/s11042-024-20097-y

Kaushik, S., & Verma, R. (2025). LD-DBGRTM: Vision transformer enabled lightweight attention-based deep learning model for pneumonia detection. *Biomedical Materials & Devices*, 1-25. https://doi.org/10.1007/s44174-025-00548-2

Kumar, S., & Bhowmik, B. (2025). ADConv-Net: Advanced deep convolution neural network for COVID-19 diagnostics using chest X-ray and CT images. *SN Computer Science, 6*(5), 1-22. https://doi.org/10.1007/s42979-025-03923-4

Maniruzzaman, M., Sami, A., Hoque, R., & Mandal, P. (2024). Pneumonia prediction using deep learning in chest X-ray images. *International Journal of Science and Research Archive, 12*(1), 767-773. https://doi.org/10.30574/ijsra.2024.12.1.0880

Odeh, A. A. R., & Mustafa, A. (2024). Explaining transfer learning models for the detection of COVID-19 on X-ray lung images. *International Journal of Electrical and Computer Engineering (IJECE), 14*(4), 4542-4550. https://doi.org/10.11591/ijece.v14i4.pp4542-4550

Okolo, G. I., Katsigiannis, S., & Ramzan, N. (2025). CLN: A multi-task deep neural network for chest X-ray image localisation and classification. *Expert Systems with Applications*, Article 128162. https://doi.org/10.1016/j.eswa.2025.128162

Patel, A. N., Murugan, R., Srivastava, G., Maddikunta, P. K. R., Yenduri, G., Gadekallu, T. R., & Chengoden, R. (2024). An explainable transfer learning framework for multi-classification of lung diseases in chest X-rays. *Alexandria Engineering Journal, 98*, 328-343. https://doi.org/10.1016/j.aej.2024.04.072

Prasath, J., Prabu, S., Mayil, V. V., & Saini, S. (2025). Optimised double transformer residual super-resolution network-based X-ray images for classification of pneumonia identification. *Knowledge-Based Systems, 311*, Article 113037. https://doi.org/10.1016/j.knosys.2025.113037

Satapathy, S. K., Cho, S. B., Mishra, S., Sah, S., & Mohanty, S. N. (2025). A federated learning approach for classifying chest diseases from chest X-ray images. *Biomedical Signal Processing and Control, 100*, Article 107107. https://doi.org/10.1016/j.bspc.2024.107107

Sayeed, A., Khansur, N. O., Srizon, A. Y., Faruk, M. F., Alyami, S. A., Azad, A. K. M., & Moni, M. A. (2024). An effective screening of COVID-19 pneumonia by employing chest X-ray segmentation and attention-based ensembled classification. *IET Image Processing, 18*(9), 2400-2416. https://doi.org/10.1049/ipr2.13106

Shanthi, S., & Mahalingam, M. (2025). Advanced chest X-ray image classification for early detection and treatment monitoring of respiratory conditions. *Biomedical Signal Processing and Control, 110*, Article 107990. https://doi.org/10.1016/j.bspc.2025.107990

Singh, K., Gaur, A., Kumar, S., Shastri, S., & Mansotra, V. (2025). Deep CP-CXR: A deep learning model for classification of COVID-19 and pneumonia disease using chest X-ray images. *Annals of Data Science*, 1-24. https://doi.org/10.1007/s40745-025-00601-3

Slika, B., Dornaika, F., Merdji, H., & Hammoudi, K. (2024). Lung pneumonia severity scoring in chest X-ray images using transformers. *Medical & Biological Engineering & Computing, 62*(8), 2389-2407. https://doi.org/10.1007/s11517-024-03066-3

Slimi, H., Cherif, I., Abid, S., & Sayadi, M. (2025). Enhanced pneumonia classification from chest X-ray images using hybrid capsule and LSTM networks with attention mechanisms and GAN integration. *Progress in Artificial Intelligence*, 1-14. https://doi.org/10.1007/s13748-025-00383-y

Tekin, V., Tekinhatun, M., Özçelik, S. T. A., Fırat, H., & Üzen, H. (2025). A deep learning-based EffConvNeXt model for automatic classification of cystic bronchiectasis: An explainable AI approach. *Journal of Imaging Informatics in Medicine*, 1-26. https://doi.org/10.1007/s10278-025-01688-z

Thavasimuthu, R., Hanumanthakari, S., Sekar, S., & Kirubakaran, S. (2024). Enhancing multi-class lung disease classification in chest X-ray images: A hybrid manta-ray foraging volcano eruption algorithm boosted multi-layer perceptron neural network approach. *Network: Computation in Neural Systems*, 1-32. https://doi.org/10.1080/0954898X.2024.2350579

Varadharajan, I., & Rathinavelayutham, S. (2025). An efficient lung disease classification using dove swarm optimisation based multi-scale faster-RCNN model. *Evolving Systems, 16*(3), Article 89. https://doi.org/10.1007/s12530-025-09708-7

Yadav, S., Rizvi, S. A. M., & Agarwal, P. (2024). Detection of lung diseases for pneumonia, tuberculosis, and COVID-19 with artificial intelligence tools. *SN Computer Science, 5*(3), Article 303. https://doi.org/10.1007/s42979-024-02617-7

Zhao, X., & Wang, X. (2025). Multi-label chest X-ray image classification based on long-range dependencies capture and label relationships learning. *Biomedical Signal Processing and Control, 100*, Article 107018. https://doi.org/10.1016/j.bspc.2024.107018

Zhu, J., Al-Qaness, M. A., Dalal, A. A., Tao, H. Z., & Alsamhi, S. H. (2025). Multiple lung diseases detection using advanced deep learning model with attention mechanisms and upsampling features. *Engineering Applications of Artificial Intelligence, 156*, Article 111038. https://doi.org/10.1016/j.engappai.2025.111038