SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

Pulmonary Disease Detection And Classification Using Feature Wise Attention Residual Expansion Network Using Chest X-Ray Images

Jannu Swetha Priyanka[1,2]* , Senguttuvan Thenmozhi[3]

[1]Department of Electronics and Communication Engineering, Dayananda Sagar College of Engineering, Bangalore, India
[2]Department of Electronics and Communication Engineering, Vardhaman College of Engineering, Hyderabad, India
[3]Department of Electronics and Communication Engineering, Dayananda Sagar College of Engineering, Bangalore, India
* Corresponding author's Email: swethapriyanka1823@gmail.com

Abstract
Chest X-Ray (CXR) images are one of the most widely utilized medical imaging techniques to diagnose pulmonary diseases due to its effectiveness and accessibility. However, accurately classifying pulmonary disease is challenging due to subtle differences and overlapping features that leads to suboptimal performance. In this research, Feature wise Attention Residual Expansion Network (FARENet) is proposed in Multi Layer Perceptron (MLP) to detect and classify pulmonary disease accurately. The expansion and projection layers in MLP enhance feature representation through increasing dimensionality for capturing complex patterns and then minimizing to retain only most appropriate information that enhance efficiency. Feature wise Channel Attention Mechanism (CAM) refines this process via assigning importance to significant features that ensure better discrimination between pulmonary disease classes. Moreover, residual skip connection provides effective gradient flow and conserve essential information from CAM that avoids model degradation. These components enhance classification accuracy, learning ability, and minimize redundancy. Hence, proposed method achieves high accuracy of 99.15%, 95.23%, and 99.41% on CXR, PulmoNet, and COVID-19 datasets compared to existing methods like InceptionResNetV2.

Keywords: Chest X-Ray, COVID-19, Feature wise Attention Residual Expansion Network, Multi Layer Perceptron, Pulmonary diseases.

1. Introduction

The lungs are primary organs responsible for respiration providing the intake of oxygen and removal of carbon dioxide during inhalation. This function under scores the significant role in human life and overall well being [1] [2]. Lung diseases including COVID-19, pneumonia, tuberculosis, emphysema, atelectasis and Acute Respiratory Distress Syndrome (ARDS) are significant global health concerns which contribute to less quality of life and high mortality rates. In this research, pulmonary disease [3] detection and classification is performed by considering pneumonia, COVID-19, normal, and lung opacity which are observed commonly in Chest X-Ray (CXR) imaging [4]. Pneumonia is caused by viruses, bacteria, and fungi occurring when alveoli fill with fluid which leads to breathing difficulties [5]. Likewise, lung opacity and COVID-19 conditions also exhibit radiographic patterns that is identified via imaging. According to World Health Organization (WHO), population like children under 5 and elderly individuals above 65 has respiratory diseases specifically pneumonia which is a leading cause of mortality worldwide [6]. Hence, detecting and classifying the pulmonary disease is crucial for diagnosis, effective treatment, and minimizing overall public health burden. To detect and classify pulmonary, CXR [7] images are used due to its efficiency and visualizing lung abnormalities and infections [8].

Bacterial and viral pneumonia are the two types of pneumonia used based on source of infection. Bacterial pneumonia needs antibiotic treatment to remove the infectious agents. Viral pneumonia follows a different progression where recovery primarily based on the body's immune system to eliminate invading

viruses [9]. A global pandemic is brought on by COVID-19 which produce numerous acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [10]. In recent years, Artificial Intelligence (AI) [11] including Machine Learning (ML) and Deep Learning (DL) techiques has represented remarkable performance in pulmonary detection from CXR images [12]. DL [13-15] is preferred over ML [16] methods due to automatically learn discriminative features and eliminates the needs for manual feature engineering [17] [18]. Also, DL capture intricate spatial patterns and subtle variations in medical images which leads to high accuracy, improved robustness, and better generalization in detecting and classifying different pulmonary diseases [19] [20].

1.1 Problem Statement and Objective

Accurately classifying pulmonary disease is challenging due to subtle differences and overlapping features that leads to suboptimal performance. To solve this issue, Feature wise Attention Residual Expansion Network (FARENet) is proposed in Multi Layer Perceptron (MLP) for accurate detection and classification of pulmonary disease. DenseNet201 and Vision Transformer (ViT) are used for feature extraction to capture fine-grained local information and global contextual details that enable accurate discrimination between visually similar patterns. The feature expansion layer enhance feature representation while projection layer eliminates redundant information for retaining only discriminative features. Moreover, feature wise CAM assigns higher importance to appropriate features that minimize overlapping and less informative regions. Residual connection conserve low-level details and enable stable learning that leads to improved classification accuracy.

1.2 Contribution

A main contribution of this research is described as:

- In MLP, expansion layer is used to increase the dimensionality of feature space to capture intricate and different feature interaction for enhanced learning. The projection layer minimize this expanded feature space to remove redundant information that enhance overall classification performance and computational effectiveness.
- The feature wise CAM assign importance weights to each features to concentrate on most relevant information for pulmonary disease classification which improve feature refinement and enhance class separability.
- Residual skip connection layer preserve original feature information by passing to deeper layers to ensure minimal information loss. This layer improves gradient flow, address vanishing gradient issue, enhance training stability, and overall model performance.

The remaining section is represented as: Section 2 denotes literature review; Section 3 provides proposed method; Section 4 contains experimental results; and Section 5 includes conclusion.

2. Literature Survey

Petra Radocaj et al. [21] developed a Convolutional Neural Network (CNN) technique with mish activation function to diagnose pediatric pneumonia by utilizing CXR images. Initially, resizing was applied to ensure uniform dimension whereas rotation, zooming, shifting, and horizontal flipping were employed for augmentation to balance classes. CNN model like InceptionResNetV2 was used to effectively extract the features and diagnose the pneumonia disease while mish activation provides improved information flow and gradient propagation via the network. However, CNN with mish activation struggle to distinguish subtle variations between pneumonia and other lung conditions due to overlapping visual features that leads to misclassification.

Mudasir Ali et al. [22] introduced a EfficientNetV2L method to detect the pneumonia. To enhance model's robustness and avoid overfitting, data augmentations like shifting, random rotation, zooming and flipping were used which generate additional training samples. Then, resizing and normalization were applied to ensure uniform dimension and consistent pixel intensity distribution across all input images. At last, EfficientNetV2L capture the visual characteristics related with pneumonia that enable accurate identification of subtle patterns. Nevertheless, EfficientNetV2L struggle with class imbalance and subtle visual differences in pneumonia patterns that affects sensitivity and leads to suboptimal performance.

Sukhendra Singh et al. [23] suggested a ViT with self attention mechanism to detect pneumonia effectively. To capture spatial relationship and global contextual information, ViT employs self-attention mechanism. ViT captures long-range dependencies through modeling interactions between image patches which enable

more comprehensive understanding of intricate anatomical structures. Self-attention enable model to select most informative regions that enhance feature representation and diagnostic accuracy.

Faisal Alshanketi et al. [24] presented a Visual Geometry Group-16 (VGG-16) with class weights and data augmentation to diagnose pneumonia. To improve computational effectiveness, high-resolutional images were resized to ensure uniform dimension which maintains essential diagnostic features. The incorporation of class weights were addressed the issue of class imbalance by assigning higher importance to underrepresented classes that leads to enhanced generalization. However, VGG-16 with class weights struggle to capture intricate variations in CXR images that results in misclassification.

AbdulRahman Tosho Abdulahi et al. [25] developed a Deep CNN (DCNN) with PulmoNet to detect the pulmonary disease. The size of the obtained images were minimized by rescaling to acquire rapid model training process and then, the rescaled images were converted into greyscale. Rotation, scaling, translation, and flipping were applied in data augmentation to balances the classes. At last, PulmoNet method was used for detecting pulmonary disease which was inspired by Wise Residual Network (WRN). This method solves issues rapidly through adding shortcuts between different layers that avoid distortion.

Burcu Oltu et al. [26] suggested a multi-head attention based network with DenseNet201 and Vision Transformer (ViT) for CXR classification. DenseNet201 effectively extracts the features whereas ViT capture long-distance dependencies which enhance classification accuracy. Then, Global Average Pooling (GAP) was used to retain spatial information that minimize dimensionality and enhance feature learning. However, multi-head attention introduce redundancy across attention heads where multiple heads learn similar patterns that minimize overall efficiency of feature representation.

Ankon Ghosh Argho et al. [27] presented a modified efficient CNN (EfficientConvNet) to classify lung disease accurately. Data augmentation like horizontal flipping, lighting, rotation, and zooming were utilized to balance the classes. In EfficientConvNet, Leaky Rectified Linear Unit (ReLU) activation function was used to solve dying neuron problem which enhance learning process. At last, dropout layer was used to solve overfitting and diverse dense layer were employed to improve classification accuracy.

Sanli Yi et al. [28] presented a pulmonary disease diagnosis named BSD by utilizing CXR images. At first, influence of bone structure and other tissues were eliminated based on extracted boneless pulmonary parenchyma. Then, improved U-Net was used to segment the pulmonary disease. At last, enhanced functional module was incorporated depending on boneless pulmonary parenchyma to extract the features which enhance accuracy.

Overall, existing methods struggle to distinguish subtle variations, inability to effectively capture intricate feature patterns, and presence of overlapping features that leads to suboptimal performance. To address this issue, FARENet is proposed to detect and classify the pulmonary disease effectively. DenseNet201 and ViT captures local and global features that enable better discrimination of subtle and overlapping features. Expansion layer, projection layer, feature wise CAM, and residual connection maximize feature refinement and training stability that results in enhanced classification performance.

3. Proposed Methodology

In this research, FARENet is proposed for accurate pulmonary disease and classification using CXR images. Initially, CXR, PulmoNet, and COVID-19 datasets are utilized to evaluate model performance. During pre-processing, resizing, Anisotropic Diffusion Filtering (ADF), Contrast Limited Adaptive Histogram Equalization (CLAHE), min-max, and augmentation are used to enhance image quality and ensure uniformity. Later, DenseNet201 and ViT extract the features whereas proposed FARENet performs detection and classification of pulmonary disease accurately. Figure 1 demonstrates working process of proposed FARENet method.

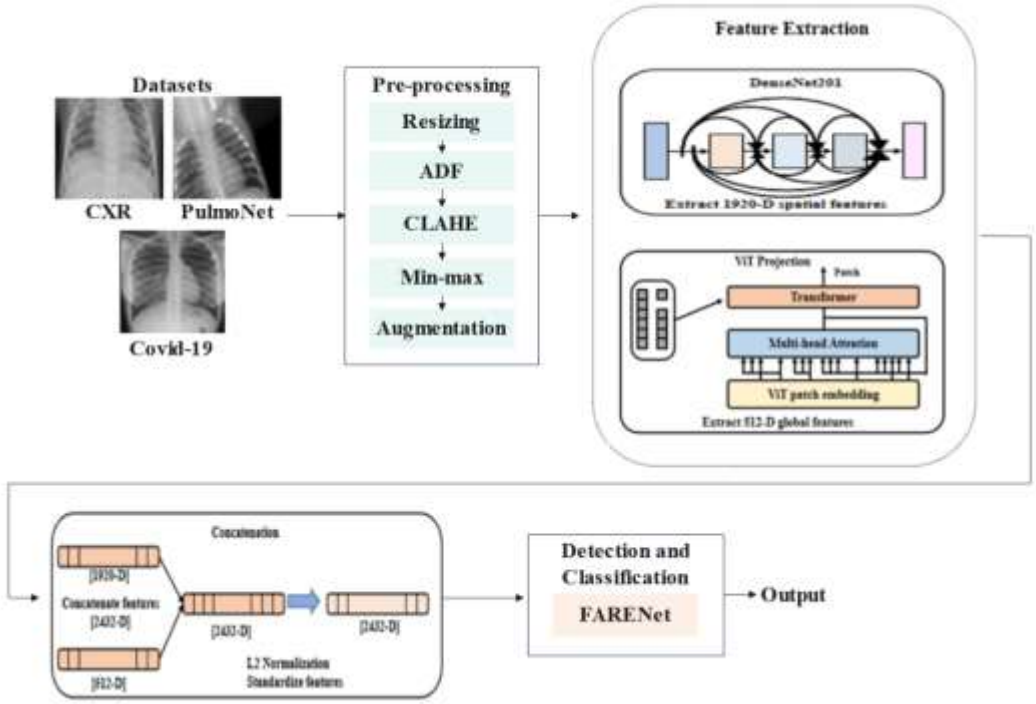


Figure 1. Working process of proposed FARENet for pulmonary disease detection and classifications

3.1 Datasets

This research employs CXR, PulmoNet, and COVID-19 datasets to estimate model performance for pulmonary disease detection and classification. Each dataset is treated as an independent classification task with own label space. The proposed method is trained and evaluated separately on CXR (binary classification) and PulmoNet as well as COVID-19 (multi-class classification) which ensures consistency. A detailed description of these datasets are presented below.

CXR: It involves a total of 5,863 CXR [29] images collected from pediatric patients aged 1 to 5 years at Guangzhou Women and Children's Medical Center. These images are classified into 4,280 pneumonia cases and 1,583 normal cases which provides a binary classification task. Figure 2 represents representative sample images from CXR.

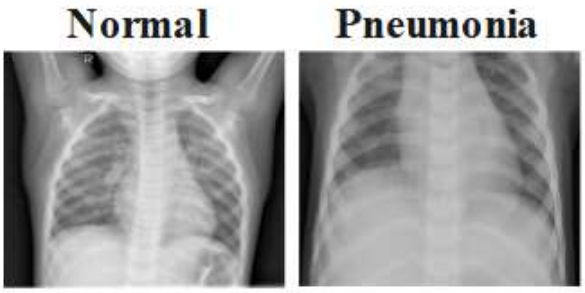


Figure 2. Sample images from CXR dataset

PulmoNet: This dataset contains a total of 16,435 images involving 883 bacterial pneumonia, 10,325 normal, 3,749 COVID-19, and 1,478 viral pneumonia cases. PulmoNet [30] rovides a different representation of pulmonary conditions which enable multi-class classification of lung diseases. Figure 3 denotes sample images from PulmoNet dataset.

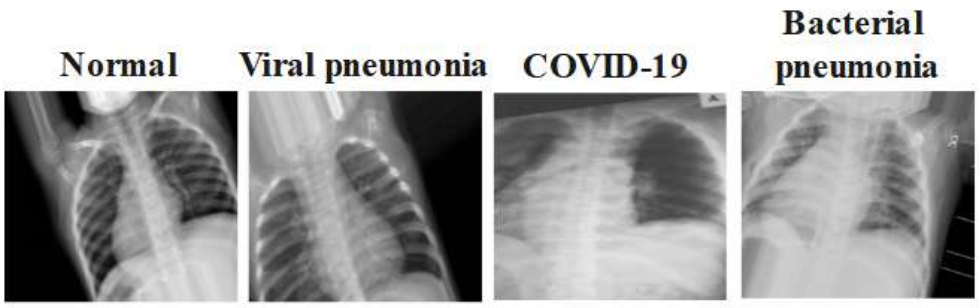


Figure 3. Representative sample images from PulmoNet dataset

COVID-19: It is established by Qatar University and Dhaka University which contains a different collection of CXR images for multi-class classification. This dataset has 4 major categories like 10,192 normal, 3,616 COVID, 6,012 lung opacity, and 1,345 viral pneumonia images. Figure 4 demonstrates sample images from COVID-19 [31] dataset and table 1 represents distribution of datasets used for pulmonary disease detection and classification. The obtained images are passed through pre-processed stage.

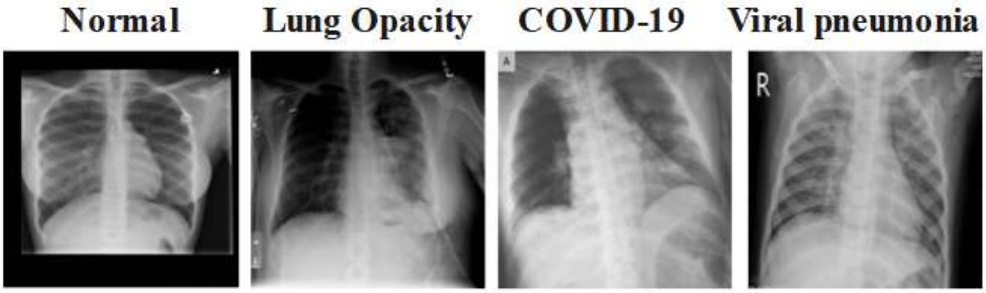


Figure 4. Sample images from COVID-19 dataset

Table 1. Dataset distribution showing number of classes, total images, training testing split for all datasets

Datasets	Classes	Total Images	Train (80%)	Test (20%)
CXR	2	5,863	4,690	1,173
PulmoNet	4	16,435	13,148	3,287
COVID-19	4	21,165	16,932	4,233

3.2 Pre-processing

After obtaining images, resizing, ADF, CLAHE, min-max, and augmentation are used to improve image quality and ensure uniformity. These methods minimize noise, improve contrast, standardize pixel intensity values, and increase data diversity for robust model training. A detailed description of each pre-processing methods are presented below.

Resizing: Due to datasets contains images with varying resolutions, resizing is applied to transform all CXR images into uniform dimensions of 224 × 224. This ensures consistency across datasets, enable effective batch processing and stable model training.

ADF: This method is applied to CXR images for minimizing noise while preserving important anatomical structures. Unlike traditional smoothing methods, ADF [32] performs edge-preserving diffusion that prevents blurring of critical regions in image. This method operates through selecting smoothing regions by maintaining edges and boundaries using low-gradient areas for diffusion and limiting it in high-gradient regions. In this research, ADF is applied with 10 iterations, conductance parameter k=15, and time step = 0.25. This configuration effectively minimize small-scale intensity variations by preserving edge sharpness which leads to cleaner and less noisy images.

CLAHE: It is employed to enhance the local contrast of CXR images which makes subtle pathological regions like opacities and infiltrates more visible. CLAHE [32] operates on small regions of image defined by tile size 8 × 8 and employs clip limit 0.02 to manage contrast amplification which prevents noise over-

enhancement. This method enhance the visibility of low-contrast lung regions while maintaining image quality and provides more effective learning. Figure 5 depicts the CLAHE images from different datasets.

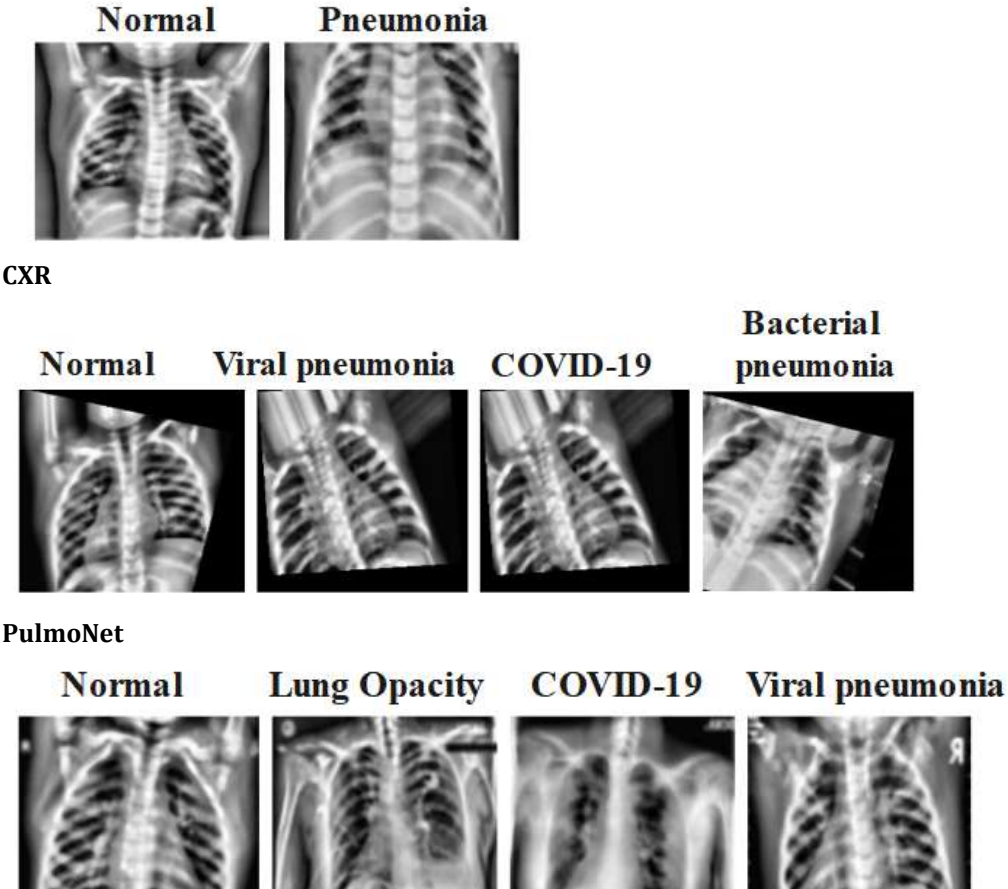


Figure 5. Representative CLAHE images from different datasets

Min-max: This method is used to scale a pixel intensity values of CXR images to fixed range [0,1] which ensures uniform distribution. Since CXR images have varying intensity ranges, min-max is performed individually for each image where minimum and maximum values are calculated per image using equation (1). It also stabilize the input for DL method, improve convergence during training and overall classification performance.

$$X_{new} = (X - X_{min})/(X_{max} - X_{min}) \quad (1)$$

Where X demonstrates input, X_{new} represented normalized value, X_{min}, X_{max} indicates minimum and maximum value.

Data Augmentation: Random rotation [$\pm$20], horizontal flipping with a probability of 0.5, and shifting [$\pm$3 pixels in x-translation and $\pm$3 pixels in y-translation] are used to increase the diversity of training samples and enhance model generalization. Since datasets are limited in size, augmentation assist the model learn from spatial transformation of same image without modifying clinical relevance. This minimize overfitting, enhance robustness, and enable to learn invariant features for accurate pulmonary disease detection and classification. To avoid data leakage, dataset is initially split into training 80% and testing 20%. Augmentation is applied only to training set during online training which ensures that no overlap between training and testing samples. Table 2 shows the impact of data augmentation across different datasets. The pre-processed images are passed through feature extraction stage for further process.

Table 2. Comparison of total images before and after augmentation across different datasets

Datasets	Before Augmentation	After Augmentation
CXR	5,863	351,750
PulmoNet	16,435	986,100
COVID-19	21,165	1,269,900

3.3 Feature Extraction

After pre-processing, DenseNet201 and ViT are used to extract the features. Feature extraction is a process of identifying and extracting the appropriate features from previous stage that is utilized for classification. DenseNet201 captures local features whereas ViT captures global features and the detailed working process of both methods are discussed below.

DenseNet201: This method is used to extract the features in pulmonary disease due to effective information flow and feature reuse ability. DenseNet201 [33] contains 4 four dense blocks and three transition layers where each dense block has 6, 12, 48, and 32 densely connected convolutional layers. Each layer within a dense block employs a combination of 1×1 convolution followed by 3×3 convolution to minimize complexity and extract the features. Each layer output is concatenated with all following layers which enable to learn rich and diverse representations that solve vanishing gradient issue. Between dense blocks, transitition layers are included by containing Batch Normalization (BN), 1×1 convolution, and 2×2 average pooling with stride 2 to minimize spatial dimension. This dense connectivity ensures maximum information propagation and enhance model performance for accurate classification. The final classification layer of DenseNet201 is removed and features are extracted from output of last convolution block. Global Average Pooling (GAP) is used to transform feature maps into 1920 dimensional feature vector that represents high-level spatial representation learned from CXR images.

ViT: It extracts features through dividing image into patch size of16×16 that are flattened into patch embeddings. Positional embeddings are incorporated to these patch representations for retaining spatial information. The sequence of embedded patches are passed through multiple transformer encoder layers consisting of multi-head self-attention and feed-forward neural network. The transformer encoder contains 12 layers with 12 attention heads per layer to capture global contextual relationships. The resulting feature representation is aggregated and fed into linear projection layer to get 512 dimensional feature vector that indicates global information from CXR images.

The extracted features from DenseNet201 are 1920 dimension while ViT produces 512 dimensional feature vector. These features are concatenated to form a combined feature vector of 2432 which is passed to the detection and classification process.

3.4 Design of Feature-wise Attention Residual Expansion Network (FARENet) for Pulmonary Disease Detection and Classification

After extraction, FARENet is used in MLP to detect and classify the pulmonary disease effectively. MLP [34] is selected because of its ability to learn intricate non-linear relationships from high dimensional feature representation. Compared to traditional classifiers, MLP offers better flexibility in modeling complex patterns and interactions between features that enhance the detection and classification of pulmonary disease. Before concatenation, the feature representations gathered from DenseNet201 and ViT are aligned via learning process. Although normalization is not applied before concatenation, expansion and projection layers in FARENet transform the combined feature vector. Initially, a feature expansion layer is used to expand the feature dimension to 4,864 to capture more complex and discriminative patterns followed by Rectified Linear Unit (ReLU) for introducing non-linearity and improve learning ability. Then, projection layer is applied to minimize feature dimension back to 2,432 which eliminate redundant information that enhance computational efficiency. To further refine the features, feature-wise CAM is used to assign adaptive weights to significant features. A residual skip connection is utilized to ensure effective gradient flow and to retain feature information which avoids degradation performance. ReLU enhance feature transformation followed by dropout layer to minimize overfitting through randomly deactivating neurons during training. At last, Fully Connected (FC) layer with softmax function is used to perform classification. This design effectively leads to robust and accurate classification method. Figure 6 shows layerwise architecture of proposed FARPNet and its detailed working process is presented below.

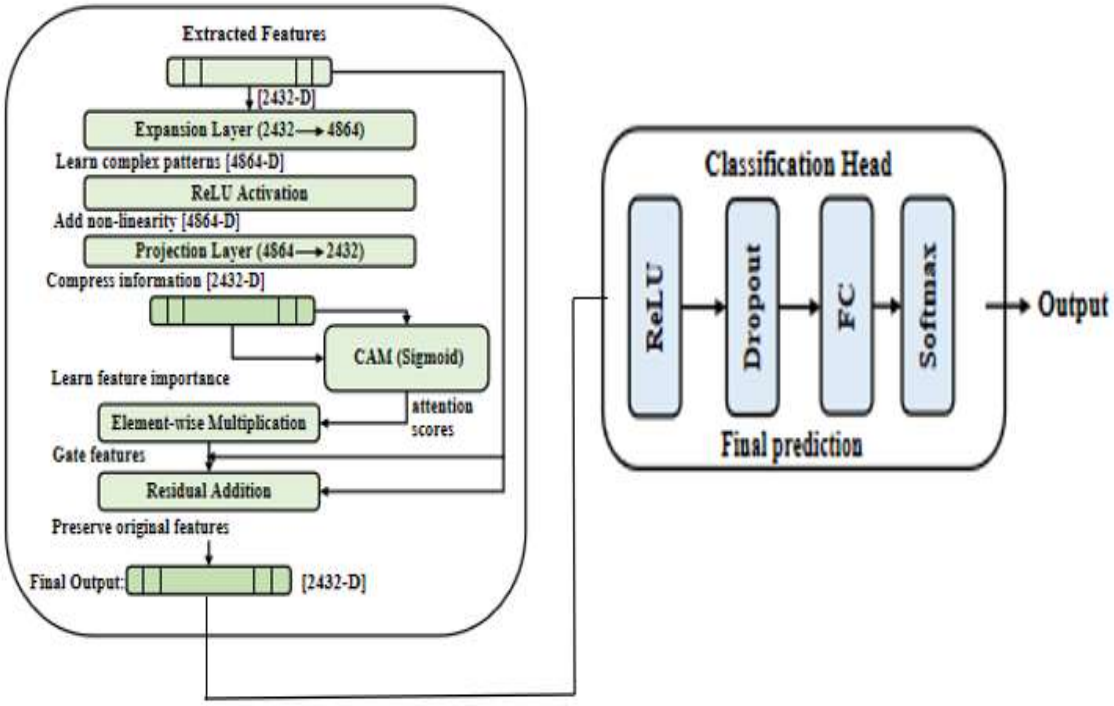


Figure 6. Layerwise architecture of proposed FARPNet for pulmonary disease detection and classification

Input Layer: The MLP considers the concatenated feature vector size of 2,432 obtained from DenseNet201 and ViT as input $I = I_1, I_2, \ldots, I_n$ which assigns each feature a weight. The feature vector with n values are converted into n input features to indicate presence of values. Equation (2) represents input feature undergoes the function $f_s(I)$ where it is computed as sum of input sequence.

$$f_s(I) = \sum_{i=1}^{n} \omega_i L_i \quad (2)$$

Where ω_i demonstrates weight assigned to appropriate data item i. Weights are represented as ω_{ih} which connects input and hidden layers, t_h indicates threshold at hidden layers.

Expansion layer: This layer is introduced in the hidden layer to increase the dimensionality of input feature vector from 2,432 to 4,864 which enhance representational ability of model. This layer learns more complex and different feature interactions by projecting features into higher dimensional space which effectively distinguish between different pulmonary disease classes through enriched feature representation. The selection of an expansion factor of 2 ensures a balance between improved feature diversity and computational efficiency which avoids redundancy and overfitting associated with larger expansion. Also, ReLU is used to introduce non-linearity for learning complex patterns from feature space which minimize vanishing gradient issue.

Projection layer: It is used to minimize the expanded feature dimension from 4,864 back to 2,432 through learnable linear transformation that ensure efficient representation. The projection layer assist to eliminate less informative features introduced during expansion step and improve computational efficiency. Instead of simple compression, this layer selectively retains only most discriminative information required for classification by assigning higher importance to relevant features. This process enhance generalization by avoiding overfitting caused by high-dimensional representations. Hence, the projection layer maintain an optimal balance between dimensional efficiency and feature richness that improve model performance.

Feature-wise Channel Attention: This method is used to assign importance weights (0 to 1) to each feature by focusing on most relevant feature channel. Compared to Multi Head Attention Mechanism (MHAM), channel attention is computationally efficient that minimize model complexity and captures feature importance. This enable more appropriate to refine high-level features in classification process that leads to enhanced performance. The feature-wise channel attention employs Squeeze and Excitation (SE)

module for designing the dependencies between diverse feature channels. In SE module, each one dimensional feature with channel length T is aggregated into real number which forms global feature map w with feature size ($C \times 1$). The mathematical formula of i^{th} channel w is represented in equation (3)

$$w_i = F_{squeeze}(u_i) = \frac{1}{T}\sum_i^T u_i(j) \quad (3)$$

Where $F_{squeeze}$ represents squeeze operation, u_i denotes feature map of i^{th} channel, $u_i(j)$ determines feature map of i^{th} channel in j^{th} element. Then, correlation between channels are modeled via feedback structure which contains 2 FC neural network to produce appropriate weight α using equation (4)

$$\alpha = F_{excitation}(w, W) = \sigma(W_2\eta(w.W_1)) \quad (4)$$

Where W_1 indicates weight of 1st FC of convolution process, W_2 represents weight of 2nd FC layer, $F_{excitation}$ determines excitation function, C denotes channels, r determines hyper-parameter in training process, σ denotes sigmoid activation function, and η refers non-linear activation function ReLU. The mathematical formula of refined feature map U' in CAM is formulated in equation (5)

$$U' = \alpha.u_c \quad (5)$$

Where u_c denotes original feature map of channel c. Based on this SE module, CAM is processed by learning feature-wise weights which are then utilized to recalibrate the importance of each channel. In CAM, two features maps are acquired through global maximum and average pooling of size $C \times T$. Later, these 2 feature maps are passed through shared neural network (MLP) which has 2 layers of 1 dimensional CNN with C/r and C neurons. The corresponding elements are added to 2 output vectors to acquire new feature map. At last, channel attention weights are gathered using sigmoid function and the refined feature map is generated by multiplying the elements associated with initial input feature map. The mathematical formula of weight value in 1 dimensional CAM is presented in equation (6)

$$\beta(U) = \sigma(W_1W_2(U_{avg} + U_{max})) \quad (6)$$

Where U_{avg} refers average pooling feature maps, U_{max} determines maximum pooling feature map, $\beta(U)$ represents channel attention weight vector obtained from CAM. The refined feature map is denoted as U' which is represented in equation (7).

$$U' = \beta \otimes U \quad (7)$$

Residual skip connection: It is utilized to preserve original feature information and improve gradient flow. The input feature vector of FARENet is added to output of feature refinement process. This residual connection ensures that essential low-level information is retained and provides efficient gradient propagation which solves vanishing gradient issue and enhance overall training stability.

Output layer: A network learns relationships between input and predicted output through modifying the weight and bias. Equation (8) shows predicted output for g^{th} node over v^{th} neuron with inputs $\{q_1, q_2, \ldots, q_i\}$ at corresponding layer is expressed in equation (8).

$$O_i(g) = \sum_{v=1}^{g} w_h^2 f(\sum w_{vh}^1 q_v(g) + t_h) \quad (8)$$

The following sub-layer within hidden layer obtain output $O_h(x)$ is used as input for further process. The bias associated with hidden and output layer values b_{ho} are utilized to calculate neuron's weight O_v using equation (9). where w_{ho} represents weight for hidden to output layer.

$$O_v = \sum_{j=1}^{g} w_{ho} \times O_j(g) + b_{ho} \quad (9)$$

Softmax layer: This layer is employed to convert final output into normalized probability values for each class. It ensures that output values present in [0,1] range that enable model to interpret the outcome as class-wise probabilities. This provide effective binary and multi-class pulmonary disease classification through selecting the class with highest probability as final prediction.

The hyperparameters are selected based on Grid search to ensure optimal model performance. An expansion factor of 2 is applied to enhance feature representation while ReLU is used in hidden layer to introduce non-linearity. The softmax activation function is used to generate normalized feature weights between 0 and 1. A dropout rate of 0.5 is utilized to minimize overfitting through randomly deactivate neurons during training. Adam optimizer with learning rate of 0.001 is employed for efficient weight updates while batch size of 64 with 100 epochs ensures stable performance and rapid convergence. Categorical cross entropy loss function is utilized for multi-class classification by calculating between predicted probabilities and true label. Algorithm 1 shows psuedo code of proposed method and table 1 demonstrates notation descriptions used in methodology to ensure transparency and reproducibility.

Table 1. Description of notation used throughout the proposed methodology

Symbols	**Description**
X	input
X_{new}	normalized value
X_{min}, X_{max}	minimum and maximum value
ω_i	weight assigned to appropriate data item i
ω_{ih}	weights that connects the input and hidden layers
t_h	threshold at hidden layers.
w_i	global feature map w of i^{th} channel
W_1	weight of 1st FC layer of convolution process
$F_{squeeze}$	squeeze operation
u_i	feature map of i^{th} channel
$u_i(j)$	feature map of i^{th} channel in j^{th} element
W_2	weight of 2nd FC layer
channels	C
r	hyper-parameter in training process
σ	sigmoid activation function
U_{avg}	average pooling feature maps
U_{max}	maximum pooling feature map
U'	feature map after channel wise weighting
b_{ho}	bias corresponding to hidden and output layer values
O_v	output neuron's weight
w_{ho}	weight for hidden to output layer
u_c	original feature map of channel c
$F_{excitation}$	excitation function
η	non-linear activation function ReLU
$\beta(U)$	channel attention weight vector obtained from CAM

Algorithm 1 of proposed method

Input: Datasets D={CXR, PulmoNet, COVID-19}
Output: Predicted class labels

Start
Step 1: Datasets and Pre-processing

Load datasets D
Split the datasets into 80% training and 20% testing

Step 2: Pre-processing
For each input image in D do:
Resize image to uniform dimension
Apply ADF to minimize noise and CLAHE to enhance contrast
Normalize image using min-max normalization
Apply data augmentation like random rotation, horizontal flipping, shifting
Store pre-processed image
End For

Step 3: Feature Extraction

For each pre-processed image do:
Employ DenseNet201 to capture local features
Utilize ViT to capture global features
Concatenate the obtained 2432 feature vectors
End For

Step 4: Detection and Classification

Initialize model parameters (weights and biases) and set hyperparameters
For epoch =1 to N do:
For each batch in training data do:
For each feature vector in batch do:
Input the concatenated feature vector
Pass through expansion layer to increase feature dimension
Apply activation function to introduce non-linearity
Pass through projection layer to minimize feature dimension
Calculate channel-wise attention weights using global pooling process, generate attention scores, and recalibrate feature importance
Refine features using attention weights
Add residual connection to conserve original features
Use dropout to minimize overfitting
Pass through FC layer
Generate class probabilities using softmax
Calculate loss using true labes
Update model parameters
End For
End For
End For

Return predicted labels
End

4. Results

The proposed method is simulated using MATLAB environment with DL toolbox, 64 GB Random Access Memory (RAM), Windows 10 operating system, Intel i5 processor, 6 GB Graphic Processing Unit (GPU), and 1 TB Solid State Drive (SSD). Accuracy, specificity, recall, f1-score, and precision are used to compute proposed method performance using equations (10) to (14).

$$\text{Accuracy}= \frac{TP+TN}{TP+TN+FP+FN}\times 100 \quad (10)$$

$$\text{Specificity}= \frac{TN}{TN+FP}\times 100 \quad (11)$$

$$\text{Recall}= \frac{TP}{TP+FN}\times 100 \quad (12)$$

$$\text{F1-Score}= \frac{2TP}{2TP+FP+FN}\times 100 \quad (13)$$

$$\text{Precision}= \frac{TP}{TP+FP}\times 100 \quad (14)$$

Where TN refers True Negative, TP demonstrates True Positive, FP illustrates False Positive, *FN* defines False Negative.

4.1 Performance Analysis

Table 3 represents analysis of numerous DL methods on diverse datasets. The proposed FARENet achieves high accuracy due to employing DenseNet201 and ViT which capture local spatial information and global contextual details unlike CNN and swin transformer focus on limited representation. The feature expansion and projection layers enhance discriminative feature learning that minimize redundancy and enhance

efficiency compared to MLP. Also, feature-wise CAM focuses on significant features whereas residual skip connections enable better gradient flow that leads to enhanced classification performance.

Table 3. Performance analysis of different DL methods for pulmonary disease detection and classification

Datasets	Methods	Accuracy (%)	Recall (%)	Precision (%)	Specificity (%)	F1-score (%)
CXR	CNN	94.75	94.50	94.30	94.80	94.40
	Swin transformer	95.73	95.56	95.49	95.71	95.52
	MLP	98.21	97.16	97.12	97.23	97.14
	Proposed FARENet	99.15	98.97	99.01	99.11	98.95
PulmoNet	CNN	83.85	83.60	84.40	83.90	84.50
	Swin transformer	91.72	91.55	92.48	93.70	91.50
	MLP	92.10	92.05	92.90	93.12	93.42
	Proposed FARENet	95.23	95.98	95.25	99.15	98.98
COVID-19	CNN	93.45	94.16	93.63	95.32	93.89
	Swin transformer	95.82	95.34	95.37	95.70	95.35
	MLP	98.90	98.32	98.55	98.97	98.43
	Proposed FARENet	99.41	99.15	99.08	99.31	99.11

Table 4 demonstrates analysis of k-fold validation for proposed method. The k=5 achieves better accuracy due to providing optimal balance between variance and bias which ensure sufficient training data and maintain reliable performance. Compared to lower folds (k=3), k=5 minimize bias through allowing more diverse training samples and compared to higher folds (k=7,9), it avoids excessive variance and overfitting from smaller validation splits. Also, it ensures stable learning through maintaining limited data in each fold for consistent feature representation. Therefore, k=5 achieves balanced trade-off between bias and variance that results in enhanced generalization and higher accuracy.

Table 4. Analysis of k-fold validation which represents classification performance of proposed method

Datasets	k-fold	Accuracy (%)	Recall (%)	Precision (%)	Specificity (%)	F1-score (%)
CXR	k=3	96.58	96.40	96.32	96.55	96.36
	k=5	99.15	98.97	99.01	99.11	98.95
	k=7	98.32	98.19	98.24	98.29	98.21
	k=9	97.42	97.34	97.28	97.35	97.31

PulmoNet	k=3	91.25	90.25	90.25	92.25	90.25
	k=5	95.23	95.98	95.25	99.15	98.98
	k=7	92.12	93.12	92.45	98.12	92.78
	k=9	94.68	93.68	94.68	95.68	94.17
COVID-19	k=3	96.28	96.28	96.28	96.28	96.28
	k=5	99.41	99.15	99.08	99.31	99.11
	k=7	97.82	97.82	97.82	97.82	97.82
	k=9	97.05	97.05	97.05	97.05	97.05

Table 5 shows analysis of computational complexity and statistical analysis for different methods. Computational complexity is evaluated using run time, memory usage, training time, and inference time while statistical analysis is measured through Confidence Interval (CI) over 5 runs with 42 random seeds to determine the model reliability. The FARENet obtains high training time, inference time, and memory consumption due to feature expansion layer increases dimensionality which includes computational overhead and memory usage. Also, feature-wise attention and residual connection enhance performance but increased complexity. CI measures the consistency and stability of model performance which shows that obtained accuracy is not due to random variation.

Table 5. Analysis of computational complexity and statistical analysis for different DL methods

Datasets	**Methods**	**Runtime (s)**	**Memory usage (MB)**	**Training time (s)**	**Inference time (s)**	**CI (%)**	**STD**
CXR	CNN	1.25	18.50	95.40	0.0011	(95.73-93.77)	(±1.12)
	Swin transformer	2.10	32.40	145.20	0.0019	(96.47-94.99)	(±0.85)
	MLP	5.45	84.60	310.80	0.0048	(98.69-97.73)	(±0.55)
	Proposed FARENet	6.85	92.40	375.50	0.0052	(99.33-98.97)	(±0.21)
PulmoNet	CNN	1.56	18.50	145.50	0.0010	(85.12, 82.58)	(±1.45)
	Swin transformer	2.86	32.40	215.30	0.0021	(92.68, 90.76)	(±1.10)

	MLP	7.23	84.60	410.60	0.0045	(92.84, 91.36)	(±0.85)
	Proposed FARENet	8.65	95.40	482.40	0.0053	(95.53, 94.93)	(±0.35)
COVID-19	CNN	1.35	18.50	110.50	0.0012	(94.54, 92.36)	(±1.25)
	Swin transformer	2.30	32.40	165.20	0.0014	(96.65, 94.99)	(±0.95)
	MLP	6.10	84.60	350.40	0.0043	(99.46, 98.34)	(±0.65)
	Proposed FARENet	7.38	93.18	415.62	0.0051	(99.56, 99.26)	(±0.18)

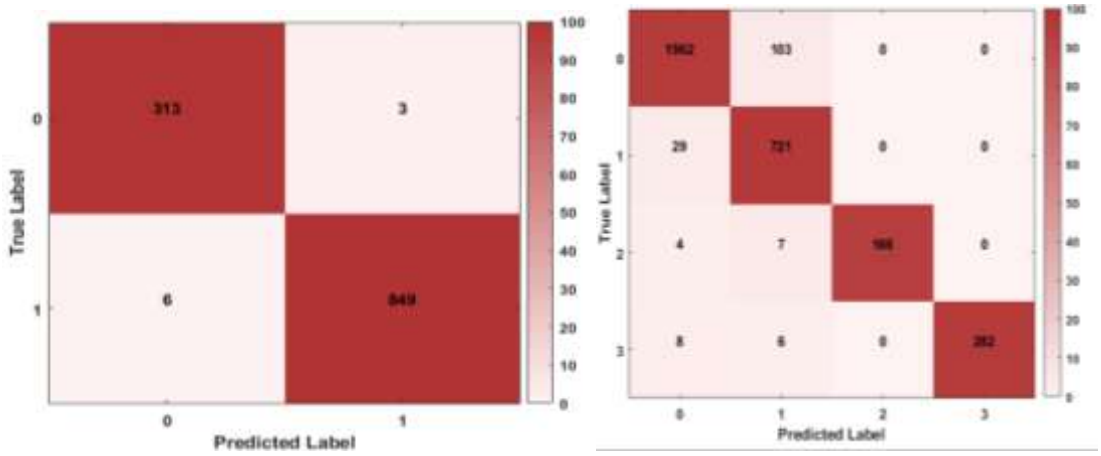
Table 6. Ablation study analysis to show the impact of individual components for proposed method. Compared to other variants, proposed FARENet achieves high accuracy of 99.15%, 95.23%, and 99.41% on CXR, PulmoNet, and COVID-19 datasets due to capturing richer and more discriminative representations. The integration of feature wise CAM and residual skip connection refines significant features which enhance gradient flow. Also, expansion and projection layer improve feature diversity and eliminates redundancy that leads to high classification performance across all datasets.

Table 6. Ablation study to show the impact of individual components for proposed method

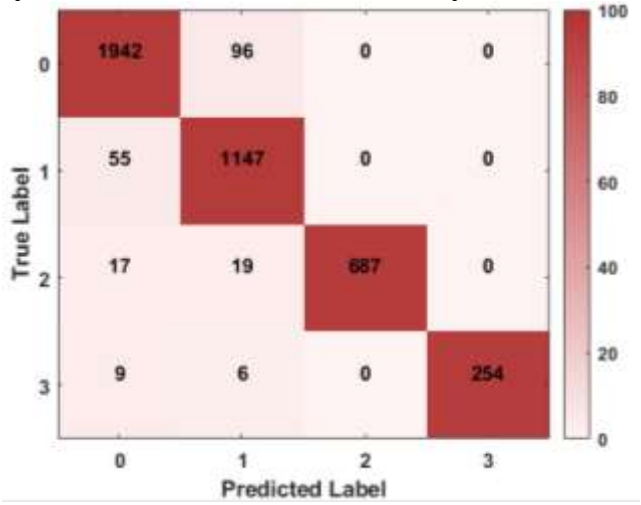
Datasets	**Methods**	**Accuracy (%)**	**Recall (%)**	**Precision (%)**	**Specificity (%)**	**F1-score (%)**
CXR	DenseNet201 + MLP	94.75	94.30	94.50	94.80	94.40
	DenseNet201 + ViT+MLP	96.10	95.85	96.00	96.15	95.90
	DenseNet201 +ViT+residual skip connection +CAM+MLP	97.65	97.50	97.45	97.70	97.48
	FARENet	99.15	99.01	98.97	99.11	98.95
PulmoNet	DenseNet201 + MLP	83.85	84.40	83.60	83.90	84.50
	DenseNet201 + ViT+MLP	87.45	88.10	86.85	88.30	87.60
	DenseNet201 +ViT+residual skip connection +CAM+MLP	91.50	92.10	90.75	92.80	91.85
	FARENet	95.23	95.98	95.25	99.15	98.98
COVID-19	DenseNet201 + MLP	93.45	93.63	94.16	95.32	93.89
	DenseNet201 + ViT+MLP	95.10	95.05	95.20	96.15	95.12

	DenseNet201 +ViT+residual skip connection +CAM+MLP	97.45	97.20	97.55	97.80	97.37
	FARENet	99.41	99.08	99.15	99.31	99.11

Figure 7 represents confusion matrix of proposed method showing pulmonary classification performance a) CXR b) PulmoNet c) COVID-19 datasets. In CXR dataset, model obtains high TP rates with minimal misclassification between pneumonia and normal classes. For PulmoNet, most of the samples are classified correctly across all four classes with minimal misclassification due to similar related categories. Likewise COVID-19 obtains accurate multi-class prediction with limited errors. As a result, outcomes shows that proposed method obtains high reliability with minimal FP and FN across different datasets.



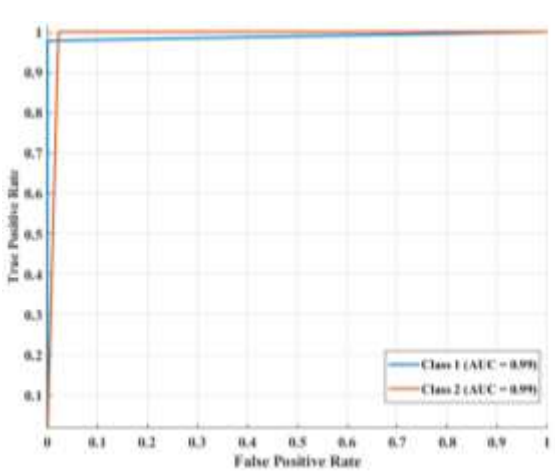
a) CXR b) PulmoNet



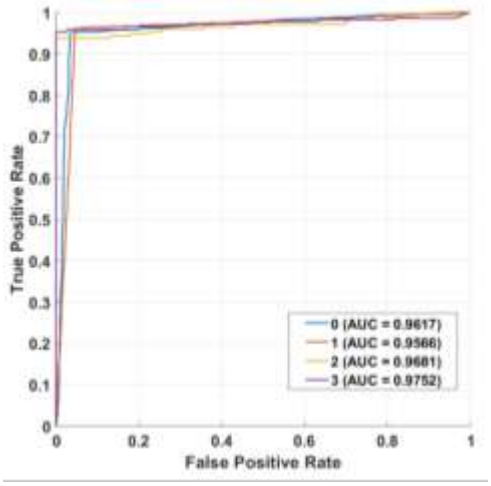
c) COVID-19

Figure 7. Confusion matrix of proposed FARENet which shows classification performance a) CXR b) PulmoNet c) COVID-19 datasets

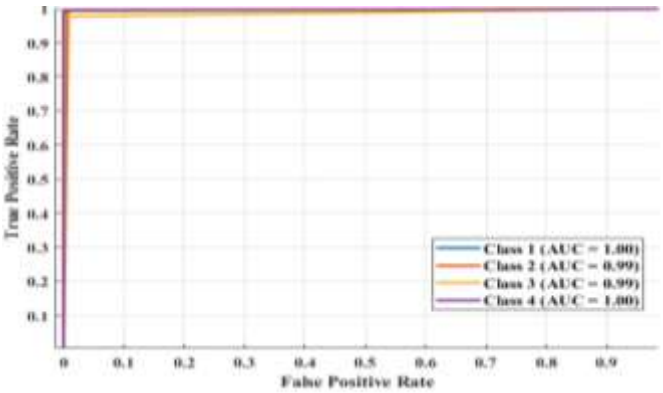
Figure 8 demonstrates Receiver Operating Characteristics (RoC) curve of proposed FARENet showing classification performance a) CXR b) PulmoNet c) COVID-19 datasets. CXR dataset obtains high Area Under Curve (AUC) values around 0.99 while PulmoNet maintains robust multi-class separability with AUC values of above 0.95. The COVID-19 dataset exhibits ideal performance with AUC values which shows reliable predictions. Across all datasets, minimal gap between curve determines stable and generalized learning ability. Overall, the outcomes validate that FARENet achieves better and consistent diagnostic accuracy.



b) CXR



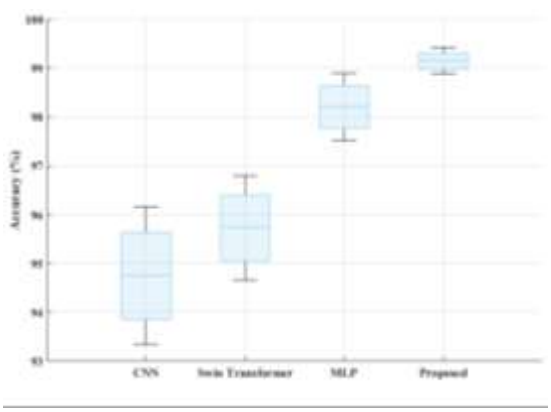
b) PulmoNet

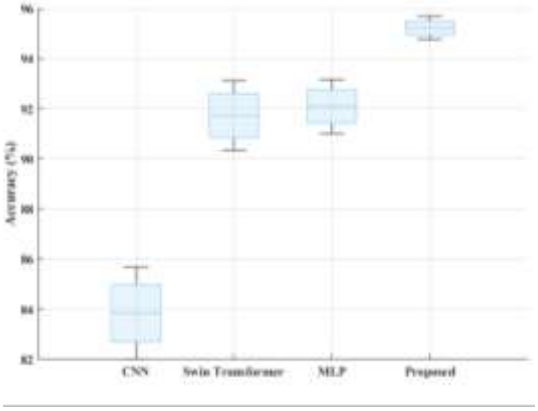


c) COVID-19

Figure 8. RoC curve analysis of proposed FARENet representing classification performance a) CXR b) PulmoNet c) COVID-19 datasets

Figure 9 depicts statistical comparison of pulmonary disease classification accuracy with baseline and proposed method a) CXR b) PulmoNet c) COVID-19 datasets. The box plots represent that FARENet achieves high accuracy with minimal variations across all datasets. Statistical analysis of proposed FARENet demonstrates significant enhancement with CXR (T = 6.182, p = 0.00174), PulmoNet (T = 13.998, p = 0.00008), and COVID-19 (T = 2.426, p = 0.03613). Hence, FARENet achieves better accuracy, stability, and generalization for pulmonary disease detection and classification.





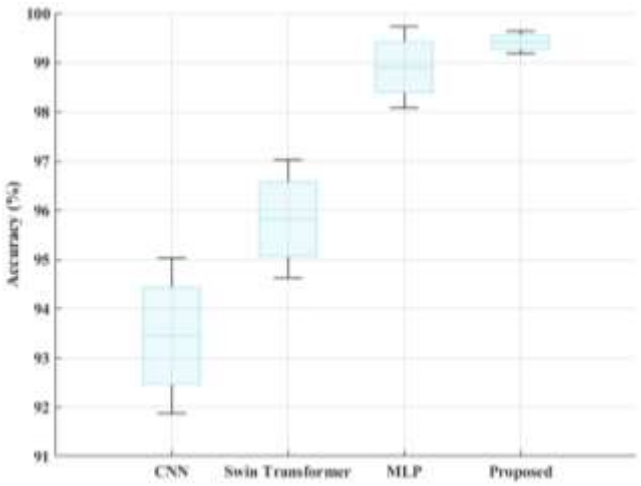


Figure 9. Statistical comparison based on pulmonary disease classification accuracy a) CXR b) PulmoNet c) COVID-19 datasets

Table 7 shows cross-dataset validation where the model is trained on COVID-19 dataset and tested on CXR dataset to analyze generalization on unseen data. The COVID-19 dataset is chosen for training which has comparatively fewer images. This validation is conducted by ensuring label space compatibility between COVID-19 and CXR datasets through label alignment. Particularly, multi-class labels in COVID-19 dataset is mapped to binary label space to match with CXR dataset. Classes like COVID-19, lung opacity, and viral pneumonia are grouped under disease category whereas normal class remain unchanged which ensures consistency across datasets. The results represents that proposed method maintains strong classification performance when evaluated on different datasets which shows robustness and has ability for learning generalized feature representation.

Table 7. Cross-dataset validation where the model is trained on COVID-19 dataset and tested on CXR dataset

Methods	Accuracy (%)	Recall (%)	Precision (%)	Specificity (%)	F1-score (%)	CI (%)	STD
CNN	82.40	81.15	80.50	83.10	80.82	(80.55, 84.25)	(±1.85)
Swin Transformer	85.15	84.80	83.90	85.45	84.35	(83.75, 86.55)	(±1.40)
MLP	88.60	87.90	88.10	89.05	88.00	(87.65, 89.55)	(±0.95)
Proposed FARENet	91.25	90.80	90.95	91.60	90.87	(90.65, 91.85)	(±0.60)

4.2 Comparative Analysis

Tables 8 to 10 demonstrates comparative evaluation of existing techniques on CXR, PulmoNet, and COVID-19 datasets. The DCNN [25] and ViT-DenseNet201 [26] values are presented in decimal form which is converted into percentage to match with proposed method value. Compared with existing methods, FARENet achieves high accuracy due to employing feature expansion, CAM, projection layer, and residual connection that enhance discriminative feature learning and classification performance. Moreover, DenseNet201 and ViT capture local spatial and global contextural features effectively that leads to obtain better performance.

Table 8. Comparative evaluation of existing techniques on CXR dataset

Methods	Accuracy (%)	Recall (%)	Precision (%)	F1-score (%)
InceptionResNetV2 [21]	97.61	97.60	97.60	97.61
EfficientNet-V2L [22]	94.02	97.24	94.40	95.80
ViT [23]	97.61	N/A	N/A	95.00
BSD [28]	98.73	98.73	98.73	98.73

Proposed FARENet	99.15	98.97	99.01	98.95

Table 9. Comparative evaluation of existing techniques on PulmoNet dataset

Methods	Accuracy (%)	Recall (%)	Precision (%)	F1-score (%)
DCNN [25]	94	95.01	90.52	88.65
Proposed FARENet	95.23	95.98	95.25	98.98

Table 10. Comparative evaluation of existing techniques on COVID-19 dataset

Methods	Accuracy (%)	Recall (%)	Precision (%)	F1-score (%)
ViT-DenseNet201 [26]	98.58	98.58	98.59	98.59
EfficientCovNet [27]	96.12	N/A	N/A	N/A
BSD [28]	95.76	95.76	95.76	95.75
Proposed FARENet	99.41	99.15	99.08	99.11

4.3 Discussion

The experimental results shows that proposed FARENet achieves better performance across CXR, PulmoNet, and COVID-19 datasets in terms of accuracy, precision, recall, specificity, and f1-score. A key finding is that integration of DenseNet201 and ViT effectively captures local and global features while feature expansion, CAM, and residual connection enhance discriminative feature learning. The statistical analysis shows improved stability with minimal variation and cross-dataset validation represents robust generalization ability on unseen data. Although the model obtain high computational cost, the substantial enhancement in classification performance and reliability enable FARENet highly suitable for accurate pulmonary disease detection and classification.

5. Conclusion

This research proposed FARENet to detect and classify the pulmonary diseases using CXR images. By integrating DenseNet201 and ViT, the model effectively captures global and local features to differentiate between multiple lung diseases. The inclusion of feature expansion, projection layers, CAM, and residual connection enhance feature representation and classification performance through learning more discriminative patterns. The projection layer eliminates redundant and less informative features which ensures computational efficiency and better generalization. Moreover, CAM focus on most relevant features which enhance model's ability. Residual connections conserve original feature information and provides smooth gradient flow that avoids degradation and improve training stability. Experimental outcomes shows that CXR, PulmoNet, and COVID-19 datasets represents better accuracy, robustness, and generalization compared to existing techniques. The proposed method obtains high accuracy of 99.15%, 95.23%, and 99.41% on CXR, PulmoNet, and COVID-19 datasets compared to existing methods like InceptionResNetV2. In future, advanced oversampling method will be considered to solve imbalance issue by generating more realistic and diverse minority class samples that enhance model robustness and minimize bias toward majority classes.

Data Availability Statement

The datasets generated during and/or analysed during the current study are available in CXR, PulmoNet, and COVID-19 datasets [https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia], [https://www.kaggle.com/datasets/abdulahiabdulrahman/pulmonet-dataset], and [https://www.kaggle.com/datasets/preetviradiya/covid19-radiography-dataset]

References

[1] Salam, M.A., Abdellatif, A., Abdallah, M. and Salam, N.A., 2025. A Hybrid Deep Learning and Machine Learning Model for Multi-Class Lung Disease Detection in Medical Imaging. International Journal of Intelligent Engineering & Systems, 18(1).

[2] Dardouri, S., 2025. A Deep Convolutional Neural Network Model for Lung Disease Detection Using Chest X-Ray Imaging. Pulmonary Medicine, 2025(1), p.6614016.
[3] Prasath, J., Prabu, S., Mayil, V.V. and Saini, S., 2025. Optimized double transformer residual super-resolution network-based X-ray images for classification of pneumonia identification. Knowledge-Based Systems, 311, p.113037.
[4] Deepak, G.D. and Bhat, S.K., 2025. A multi-stage deep learning approach for comprehensive lung disease classification from x-ray images. Expert Systems with Applications, 277, p.127220.
[5] Munir, K., Tanveer, M.U., Alyamani, H.J., Bermak, A. and Rehman, A.U., 2025. Pneux-net: An enhanced feature extraction and transformation approach for pneumonia detection in x-ray images. IEEE Access.
[6] Subramaniam, K. and Palanisamy, N., 2026. DeiT-GAN: a segmentation-driven transformer framework for multi-class chest X-ray disease classification. Automatika, 67(1), pp.19-38.
[7] Yanar, E., Hardalaç, F. and Ayturan, K., 2025. Pelm: A deep learning model for early detection of pneumonia in chest radiography. Applied Sciences, 15(12), p.6487.
[8] Usman, C., Rehman, S.U., Ali, A., Khan, A.M. and Ahmad, B., 2025. Pneumonia disease detection using chest X-rays and machine learning. Algorithms, 18(2), p.82.
[9] Vyas, R. and Khadatkar, D.R., 2025. Ensemble of deep learning architectures with machine learning for pneumonia classification using chest X-rays. Journal of Imaging Informatics in Medicine, 38(2), pp.727-746.
[10] Rao, G.E., Rajitha, B., Srinivasu, P.N., Ijaz, M.F. and Woźniak, M., 2024. Hybrid framework for respiratory lung diseases detection based on classical CNN and quantum classifiers from chest X-rays. Biomedical Signal Processing and Control, 88, p.105567.
[11] Gtifa, W., Mhaouch, A., Alsharif, N., Althobaiti, T. and Sakly, A., 2025. Nature-inspired multi-level thresholding integrated with CNN for accurate COVID-19 and lung disease classification in chest x-ray images. Diagnostics, 15(12), p.1500.
[12] Fu, X., Lin, R., Du, W., Tavares, A. and Liang, Y., 2025. Explainable hybrid transformer for multi-classification of lung disease using chest X-rays. Scientific Reports, 15(1), p.6650.
[13] Geroski, T., Ranković, V., Pavić, O., Dašić, L., Petrović, M., Milovanović, D. and Filipović, N., 2025. Enhancing COVID-19 disease severity classification through advanced transfer learning techniques and optimal weight initialization schemes. Biomedical Signal Processing and Control, 100, p.107103.
[14] Saleem, S., Sajid, M.Z., Sharif, A., Amjad, J., UlHaq, A. and Aldossary, H., 2026. An integrated deep learning framework leveraging NASNet and vision transformer with mixprocessing for accurate and precise diagnosis of lung diseases. SLAS technology, p.100394.
[15] Rahman, M.S., Chowdhury, M.E., Rahman, H.R., Ahmed, M.U., Kabir, M.A. and Sarmun, R., 2026. Self-DenseMobileNet: A robust framework for lung nodule classification using Self-ONN and stacking-based meta-classifier. Biomedical Signal Processing and Control, 112, p.108588.
[16] Kumar, S. and Bhowmik, B., 2025. EffiCOVID-net: A highly efficient convolutional neural network for COVID-19 diagnosis using chest X-ray imaging. Methods, 240, pp.81-100.
[17] Jafari Tabaghsar, B., Forghani, Y., Sheibani, R. and Tavoli, R., 2026. Lung Disease Detection Using Hybrid CNN-XGBoost Classifier With New Loss Function. IET Software, 2026(1), p.6391804.
[18] Xie, J., Xie, L., Xiao, T., Zhu, R., Hu, C., Qian, K. and Zhang, C., 2026. DiemaNet: An automated detection model for three typical respiratory diseases. Biomedical Signal Processing and Control, 112, p.108504.
[19] Narmadha, A.P. and Gobalakrishnan, N., 2025. Het-rl: Multiple pulmonary disease diagnosis via hybrid efficient transformers based representation learning model using multi-modality data. Biomedical Signal Processing and Control, 100, p.107157.
[20] Rabbah, J., Ridouani, M. and Hassouni, L., 2025. Improving pneumonia diagnosis with high-accuracy CNN-Based chest X-ray image classification and integrated gradient. Biomedical Signal Processing and Control, 101, p.107239.
[21] Radočaj, P., Radočaj, D. and Martinović, G., 2025. Optimizing convolutional neural network architectures with optimal activation functions for pediatric pneumonia diagnosis using chest X-rays. Big data and cognitive computing, 9(2), p.25.
[22] Ali, M., Shahroz, M., Akram, U., Mushtaq, M.F., Altamiranda, S.C., Obregon, S.A., Díez, I.D.L.T. and Ashraf, I., 2024. Pneumonia detection using chest radiographs with novel efficientnetv2l model. IEEE Access, 12, pp.34691-34707.
[23] Singh, S., Kumar, M., Kumar, A., Verma, B.K., Abhishek, K. and Selvarajan, S., 2024. Efficient pneumonia detection using Vision Transformers on chest X-rays. Scientific reports, 14(1), p.2487.
[24] Alshanketi, F., Alharbi, A., Kuruvilla, M., Mahzoon, V., Siddiqui, S.T., Rana, N. and Tahir, A., 2025. Pneumonia detection from chest x-ray images using deep learning and transfer learning for imbalanced datasets. Journal of imaging informatics in medicine, 38(4), pp.2021-2040.

[25] Abdulahi, A.T., Ogundokun, R.O., Adenike, A.R., Shah, M.A. and Ahmed, Y.K., 2024. PulmoNet: a novel deep learning based pulmonary diseases detection model. BMC Medical Imaging, 24(1), p.51.
[26] Oltu, B., Güney, S., Yuksel, S.E. and Dengiz, B., 2025. Automated classification of chest X-rays: a deep learning approach with attention mechanisms. BMC Medical Imaging, 25(1), p.71.
[27] Argho, A.G., Maswood, M.M.S., Mahmood, M.I. and Mondol, N., 2024. EfficientCovNet: A CNN-based approach to detect various pulmonary diseases including COVID-19 using modified EfficientNet. Intelligent systems with applications, 21, p.200315.
[28] Yi, S., Qin, S., She, F. and Shao, D., 2025. BSD: A multi-task framework for pulmonary disease classification using deep learning. Expert Systems with Applications, 259, p.125355.
[29] CXR dataset link: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia (Accessed on 31 March 2026)
[30] PulmoNetdataset link:https://www.kaggle.com/datasets/abdulahiabdulrahman/pulmonet-dataset] (Accessed on 31 March 2026)
[31] COVID-19 dataset link: https://www.kaggle.com/datasets/preetviradiya/covid19-radiography-dataset (Accessed on 31 March 2026)
[32] Ayalew, A.M., Bezabh, Y.A., Abuhayi, B.M. and Ayalew, A.Y., 2024. Atelectasis detection in chest X-ray images using convolutional neural networks and transfer learning with anisotropic diffusion filter. Informatics in Medicine Unlocked, 45, p.101448.
[33] Sharma, A. and Parvathi, R., 2025. Enhancing cervical cancer classification: Through a hybrid deep learning approach integrating DenseNet201 and InceptionV3. IEEE Access, 13, pp.9868-9878.
[34] Srinivasu, P.N., Jaya Lakshmi, G., Gudipalli, A., Narahari, S.C., Shafi, J., Woźniak, M. and Ijaz, M.F., 2024. XAI-driven CatBoost multi-layer perceptron neural network for analyzing breast cancer. Scientific Reports, 14(1), p.28674.