Journal of Theoretical and Applied Information Technology
15th February 2025. Vol.103. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

CHESTX-PNEUMONET: OPTIMIZED CNN WITH NOVEL LOSS FUNCTION FOR PNEUMONIA DETECTION

S SUGUNA MALLIKA [1], CH.SARADA[2], GAYATRI MANTRI [3], ARUNA KUMARI BOMMU [4], CH RAJYALAKSHMI [5]

[1]Department of CSE, CVR College of Engineering, Hyderabad, Telangana, India
[2]Department of CSE, CVR College of Engineering, Hyderabad, Telangana, India
[3]Department of CSE, Malla Reddy College of Engineering and Technology, Hyderabad, Telangana, India
[4]Department of CSE, Malla Reddy College of Engineering and Technology , Hyderabad, India,
[5]Department of CSE, B V Raju Institute of Technology, Hyderabad, Telangana, India
[1]suguna.kishore@gmail.com , [2]sharada.ch@gmail.com , [3]gayatricse312@gmail.com ,
[4]mailme2aruna@gmail.com , [5]lakshmi.chatragadda@bvrit.ac.in

ABSTRACT

The automated and accurate detection of pneumonia poses a significant problem in medical analysis due to the subtlety of its indicators in X-ray or CT images. This task is of paramount importance, as pneumonia claims millions of lives annually. Leveraging sophisticated techniques such as deep learning is essential to enhancing both diagnostic precision and operational efficiency. However, adapting existing neural network frameworks for clinical imaging tasks often results in overfitting and limited transferability. To mitigate these limitations, we introduce PneumoClassifyNet, an innovative lightweight convolutional network architecture tailored specifically for Pneumonia Detection in chest X-ray analysis. This architecture is more compact yet more potent than traditional fine-tuning approaches. Additionally, we introduce a new loss function, RadCE-loss, designed to effectively extract distinguishing characteristics from incorrectly classified and fuzzy images. Furthermore, the convolutional kernels are optimized within the convolutional neural network (CNN) model to improve accuracy of classification. The paper presents PneumoClassifyNet, an optimized CNN for chest X-ray classification, and RadCE-loss, a function that enhances accuracy by handling misclassified images. Results of Experiments indicate that the lightweight PneumoClassifyNet, coupled with the RadCE-loss, achieves superior performance across key metrics, including F1-score, recall, accuracy and AUC. These findings affirm that a carefully optimized convolutional neural network (CNN) architecture can outperform fine-tuned deep learning models.

Keywords: *Pneumonia Detection, Convolutional Neural Network, Optimized Loss Function, Radce-Loss , F1 Score, Accuracy*

1 INTRODUCTION

Over the past decade, computer vision has become an indispensable asset in tackling numerous complex challenges of the 21st century. One of its most promising applications lies in medical image diagnosis, where AI has demonstrated significant potential.Pneumonia continues to be a leading cause of mortality, claiming over 4 million lives annually, and constituting more than 15% of fatalities in children under five years old[1,2]. The risk factors include chronic illnesses such as asthma, cystic fibrosis, diabetes, cardiac failure and reduced immunity. Accurate and swift diagnosis is essential for better patient care.In medical imaging, chest X-rays are widely utilized for diagnosis, favoured for their non-invasive nature and cost-effectiveness. However, interpreting chest X-rays for pneumonia diagnosis is susceptible to subjective interpretability/variability [3]. While clinical evaluation and imaging form the cornerstone of pneumonia detection, even skilled radiologists may struggle with interpretation due to overlapping visual indicators with other thoracic conditions or the subtlety of the signs.

This underscores the need for automated detection systems. Numerous studies have developed automated systems for pneumonia detection via chest radiography using deep learning, which enables end-to-end classification by autonomously extracting features from raw data, unlike traditional machine learning requiring manual feature selection [4]. CNNs(Convolutional neural networks) are particularly adept at recognizing patterns in images for classification, as it inherently captures the translationally invariant information through

convolutional processes. Sharma et al.[5] and Stephen et al.[6][25] designed primitive CNN models for classifying pneumonia in chest X-rays, incorporating augmentation of data to address the scarcity of available data. Utilizing a dataset created by Kermany et al.[7][26], Stephen et al. secured 93.73% accuracy, while Sharma et al. attained an accuracy of 90.68%. Rajpukar et al.[8] applied the DenseNet121 convolutional neural network (CNN) model for classification of pneumonia, obtaining an F1-score of 76.8%. They attributed the suboptimal performance to the absence of patient history, which affected both their deep network model and the medical experts against whom the model's performance was compared.Panwar et al. [9][27] integrated the VGG-19 architecture alongside GradCAM, resulting in a 95.6% accuracy within a 3-class pneumonia classification problem. Likewise, Brunese et al. [10] utilized a VGG-16-based architecture in conjunction with the GradCAM for visual rectification, securing a 96.2% accuracy on a data collection comprising of 6,523 chest X-rays. Adi and Kemal et al.[11][29] implemented a DenseNet-201 model and accomplished recall and accuracy metrics of up to 95%, along with a precision of nearly 90% on over 6,000 chest X-rays sourced from Kaggle. Mahmud et al. [12] introduced CovXNet, a CNN architecture that employs depth-wise convolutions with variable dilation rates to obtain a diverse array of features from chest X-rays. This model, trained on 6,161 chest X-rays, attained a accuracy of 97.4% in a binary class pneumonia detection scenario.

Ouchicha et al[24][30]. established a residual CNN model that was trained on 2,905 X-ray images, resulting in an accuracy of 96.7% in pneumonia classification task.Wang and Bao et al.[13][31] introduced a prior attention residual learning block, which they integrated into two 3DResNets. Their model resulted in an accuracy of 93.3% on a pneumonia classification task, utilizing a dataset of 4,697 X-rays. Das et al.[14][32] explored transfer learning with established CNN architectures, specifically DenseNet201, ResNet50V2, and InceptionV3. By leveraging a weighted average ensembling method, they achieved 91.62% accuracy in differentiating between viral pneumonia and healthy lung tissues in a binary classification scenario, employing a data set of 1,004 X-ray scans. Nikolaou and Massaro et al.[15][33] enhanced the pre-trained EfficientNetB0 model by adding a dense layer, resulting in a more robust CNN architecture. This model, trained on 15,153 X-ray images, attained a accuracy of 95%, effectively differentiating between healthy and viral pneumonia cases. Singh and Tripathi et al.[16][34] employed a Quaternion Convolutional Neural Network (QCNN), specifically using a Quaternion Residual Network, to categorize chest X-rays as either heathy or abnormal cases. By treating the RGB color channels as a single unit for capturing patterns, their model, trained on 5,856 images, accomplished a accuracy of 93.75%. Joshi and Yadav et al[17] proposed a DL(deep learning) system derived from DarkNet-53, a 53-layer CNN pre-trained on ImageNet. After fine-tuning on dataset comprising 6,884 chest X-ray images, the model demonstrated an average testing accuracy at 97.11%, accompanied by an inference time measuring 0.137 seconds per image, making it a rapid diagnostic tool. Dash and Mohapatra et al.[18][28] presented a modified VGG-16 architecture, eliminating the fully connected layers and substituting them with a simplified set to optimize the model for pneumonia detection. Trained on a dataset of 1,272 X-rays, the model attained an accuracy of 97.12%.

Despite these breakthroughs, CNN models often encounter challenges such as overfitting on limited medical datasets and poor generalization to unseen data, limiting their utility in clinical settings. To overcome these limitations, we propose PneumoLiteNet, a compact and efficient CNN architecture particularly tailored for detection of Pneumonia in chest radiography(X-ray) images. Unlike traditional methods that rely on fine-tuning pre-trained deep learning models, PneumoLiteNet is built from scratch to optimize performance for this particular task. Additionally, we propose a novel loss function, RadCE-loss, which enhances model learning by focusing on distinguishing features in misclassified and ambiguous images. This approach not only boosts classification accuracy but also mitigates overfitting, making the model more adaptable to diverse patient data.

Our contributions in this paper are twofold: First, we present PneumoClassifyNet, a lightweight CNN model that categorizes chest X-ray images into two classes: healthy or abnormal. The model is optimized for efficiency, reducing unnecessary parameters, rearranging layer orders and incorporating elements from conventional CNN architectures to improve feature extraction. Its lightweight nature ensures faster training and lower computational costs. Second, we propose the RadCE-loss function, which enhances the

performance of PneumoClassifyNet by effectively handling misclassified and indistinct images through a self-adjusting coefficient.This paper is structured as follows: 2nd Section outlines the proposed technique, 3rd Section compares our results with the modern advanced approaches, and the last 4th Section concludes the study.

2 DATA RESOURCES AND PROCEDURES

2.1. Dataset

The Images utilized in this research originates from a deep learning competition hosted on Kaggle. It comprises X-rays of the lungs of young children between ages of one and five. These X-rays are sourced from the Guangzhou Children's Medical Centre and are verified by healthcare professionals. The dataset comprises a total of 5,856 labeled images, with 4,273 displaying indications of pneumonia, while the remaining 1,583 are identified as healthy. Each Image is a grayscale scan with dimensions varying between 1,346×1,044 and 2,090×1,858 pixels. Typically, the severely affected pneumonia Images are often characterized by prominent alveolar consolidations, primarily resulting from bacterial infections and fluid accumulation in the lungs. While in some Images, the differences between healthy and pneumonia-affected X-rays are generally discernible, in certain instances, the variations are less pronounced.

2.2. CNN Method Foundations

This section outlines two primary contributions: enhancing the loss/cost function and designing a more efficient structure or design of the model. Drawing inspiration from the characteristics of specific radical power functions, we developed a self-adapting component to refine the process of training/learning of model, with particular focus on enhancing the recognition of indistinct and wrongly classified images. Given the limitations of fine-tuning and the significant differences of images between chest X-ray and those used to train pretrained models, we opted to design a novel architecture from scratch. This new model is thinner, lighter, and optimized for binary classification. Below, detailed descriptions of the novel loss function, followed by the model architecture and the learning process is provided.

2.2.1. Problem Formulation of the Novel Loss Function

Anomaly detection in chest X-rays is formulated as a problem of binary classification, where the chest X-ray image is provided, and the outcome is a dual label signifying either the existence (1) or absence (0) of illness. For binary image classification tasks, the SoftMax cross entropy cost function is commonly utilized as the optimization criterion.

$$L = -\frac{1}{N}\sum\left[t_i \ln q_j + (1 - t_i)\ln\left(1 - q_j\right)\right] \quad (1)$$

Where ti is ground truth value taking value 1 or 0, N is the total no. of images employed for training the model, and qj ∈ [0, 1] is the SoftMax probability defined as :

$$q_j = \frac{e^{s_j}}{\sum_{c=1}^{C} e^{s_c}} \quad (2)$$ Where s denote the value that is fed to the SoftMax and C represents the total no. of categories and j∈[0,C−1], In this specific case, C=2. When examining the image X, the outcome of the final layer Qi is transformed by the model network as follows:

$$q_i = g(X|w_f), \quad q_{ij} \in Q_j \quad (3)$$ Where g represent the entire nonlinear model network and w_f denote the parameter vector across all network layers. Finding the best parameter configuration by iteratively modifying them to minimize the following function, as shown in Equation (4) is the aim of the training procedure.

$$argmin_{w_f \in W_F} \frac{1}{N}\sum L_i(Q_i, t_i|\, w_f) \quad (4)$$

Differentiating between healthy and pathological chest X-rays poses a significant challenge, particularly when normal images contain noise and the diseased areas in abnormal images are not readily apparent

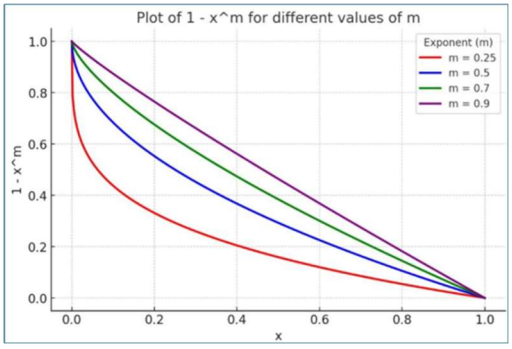


Fig 1. Graph plot of various radical power functions for distinct values of m

To address this issue, we introduce a self-adapting coefficient which is multiplied with the loss

function during the processing of the training image X, as depicted in Equation (5):

$$L_i(Q_i, t_i|\, w_f) = K_i \quad [t_i \ln q_j + (1 - t_i) \ln(1 - q_j)]. \quad (5)$$

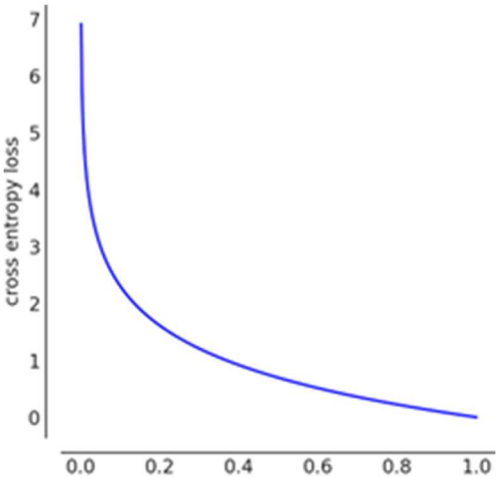
We seek an appropriate K_i with a basic shape that is monotonically decreasing and convex within the interval [0,1]. This inspiration comes from the behaviour of a simple radical power function $1-x^m$, particularly when $0<m<1$. It is observed that the convex function curve declines from 1 to 0 when x ranges from 0 to 1 as illustrated in Figure 1.

Utilizing this radical function enables us to get learning more effectively from misclassified and indistinguishable chest X-rays. The coefficient is outlined as follows:

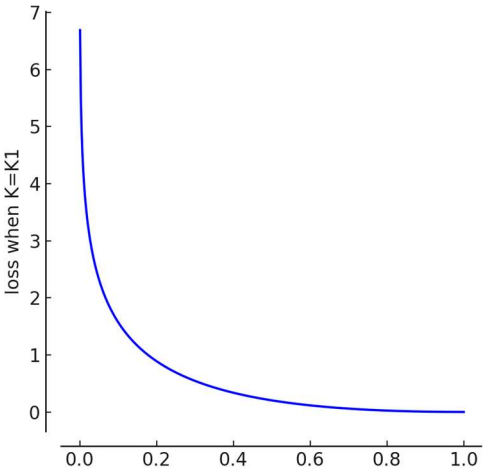
$$K_i = 1 - q_{ij}^m \quad (6)$$

In our study, $q_{ij} \in [0,1]$ and $m=0.5$. Figure 2(a) illustrates the standard CE loss when $K_i = 1$. In contrast, Figure 2(b) represents a modified form of Figure 2(a), where K_i is defined as per Equation (6).

It is evident that the curve in Figure 2(b) exhibits a steeper decline compared to the curve in 2(a) when $q_{ij}<0.5$, indicating a higher loss value when q_{ij} is small. This characteristic allows the curve in Figure 2(b) to prioritize images that are wrongly classified during the learning process.



(a)



(b)

Fig 2. Loss function curves based on varying Ki value, (a) is the standard CE loss when Ki =1, (b) is a variant of (a) with Ki defined by Equation (6)

Consequently, the RadCE-loss is expressed as in (7):

$$L_{new}(Q_i, t_i|\, w_f) = -\frac{1}{N} \Sigma\, L_i(Q_i, t_i|\, w_f) \quad (7)$$

The final loss is computed as the average loss across all data samples. Averaging the loss normalizes the output, contributing to the stabilization of the training process.

The parameters w_f are iteratively modified by using backpropagation to optimise L_{new} as shown in (8):

$$w_f = w_f - \eta \frac{\delta L_{new}}{\delta w_f} \quad (8)$$

After initializing w_f randomly, $\frac{\delta L_{new}}{\delta w_f}$ is obtained through backpropagation, and w_f is revised with the gradient descent approach, and η denotes the learning rate ; $\frac{\delta L_{new}}{\delta w_f}$ represents the gradient of $L_{new}(Q_i, t_i|\, w_f)$.

To determine the optimal parameters w_f to make $L_{new}(Q_i, t_i|\, w_f)$ minimal, Equation (8) is iteratively executed, with both w_f and $L_{new}(Q_i, t_i|\, w_f)$ continuously updated until $L_{new}(Q_i, t_i|\, w_f)$ converges to its minimum value. Executing the updates in Equation (8) makes the network to focus more on misclassified images, resulting in the development of more discriminative features and improving classification accuracy.

3.2.2 Model Architecture and Training Process

Numerous studies have endeavoured to improve on current convolutional neural network (CNN) architectures, including ResNet, VGGNet,

DenseNet, for the training of medical images [19]-[21]. Nevertheless, transfer-learning may prove suboptimal when substantial differences exist between the original and target datasets. The chest X-rays in the utilized dataset are high-dimensional grayscale images, which differ significantly from the natural images in ImageNet. With advancements in computational power, training a domain-specific classifier from scratch has become feasible. In this study, we introduce an innovative network named PneumoLiteNet (Pneumonia Detection Light/Compact Network).

As respresented in Fig. 3, the architecture of the model comprises a repetitive sequence of convolutional layers succeeded by pooling layers. The convolutional layers act as the primary mechanism for this extraction process, with each convolutional block supplemented by a batch-normalisation layer and a subsequent ReLU activation function.The model utilizes a streamlined design with only three consecutive blocks dedicated to feature extraction from images. The convolutional kernels are configured to 64, 128, and 256 in each respective block. To emphasize small lesion regions and capture intricate characteristics, a kernel size of 3x3 is employed. After these convolutional blocks, the architecture includes an additional convolutional layer (convlayer-4), succeeded by a dropout layer, batch-normalization, a global pooling layer, and the final output convolutional layer (convlayer-5).

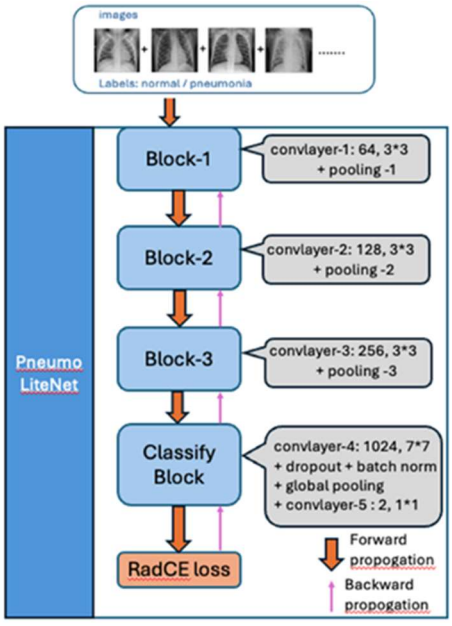


Fig 3. PneumoLiteNet

The inclusion of these layers, rather than fully connected layers, decreases the parameter count, while the dropout and batch normalization[23] layers are designed to mitigate overfitting. Pooling layers generally succeed convolutional layers, serving to diminish the dimensionality and the no. of parameters within the feature maps. Pooling achieves this by summarizing the most salient features of the feature map. In this model, max-pooling with a stride and filter size of 3x3 was used to downsample the features. ReLU was the activation function employed in every layer with the exception of the output final layer, where standard SoftMax function was utilized to classify whether the lung in an image is affected by pneumonia, by predicting probabilities in the final step.

Unlike many widely-used networks, PneumoLiteNet consists of only five convolutional layers, with three layers dedicated to feature extraction, significantly fewer than established models such as ResNet(50 layers) and DenseNet(121 Layers). Additionally, a dropout, batch-normalization, and a global pooling layer are sequentially integrated between the final two convolutional layers to prevent overfitting, a configuration not commonly found in other model architectures. Throughout the training process, the Chest X-rays are input into PneumoLiteNet, where the parameters are optimized through backpropagation to achieve a minimal value of the RadCE cost function.

To enhance the efficiency of the training process, ensure convergence, and minimize training time, two optimization mechanisms were strategically utilized. The first, known as "reduce learning rate on plateau," decreases the learning rate by a predefined factor when learning stagnates. This prevents gradient descent from overshooting the global minimum. In the proposed network, the patience parameter was established at a value of 5, meaning learning rate was reduced after five epochs of no improvement in accuracy. The reduction factor was established at 0.9, and the cooldown period, during which adjustments to the learning rate are paused, was designated as 5 epochs. This method aims to prevent overshooting and improve convergence efficiency.The second mechanism, "early stopping," monitors the validation loss and halts training when no improvement is observed over a defined interval. This method significantly accelerates the training process, as it avoids unnecessary epochs once convergence is achieved. For instance, although our model was

scheduled to train for 100 epochs, it converged after 35, saving time and preventing overfitting.

We employed the Adam optimizer[22] to improve optimization efficiency. In addition to these mechanisms, careful tuning of hyperparameters—like the batch size, learning rate and optimization algorithm played a critical role. We selected a learning rate of 4x10^-4, determined through hyperparameter tuning. It is essential to balance this parameter, as a large learning rate risks converging to a local minimum, while a smaller one slows down the training process. Experimentation showed that a mini-batch size of 32 yielded optimal performance. The batch size refers to the count of samples that are passed through the model training before the parameters of the model are updated, and introducing some update through mini-batch gradient descent helps overcome local minima and saddle points in non-convex optimization landscapes.

3 RESULTS AND DISCUSSION

3.1 Results

For result analysis, Training accounted for 80% of the available chest X-ray images, with 70% going towards real training and 10% going towards validation. The rest, or 20%, was set aside for testing. A ten-fold cross validation was conducted for each CNN architecture utilized in this evaluation study. There are four commonly employed evaluation measures for classification scenarios: recall, precision, F1 score, and accuracy. Given their prevalence in classification research, detailed definitions of these metrics have been omitted. To offer a more complete assessment of the model's performance, the Area Under the Curve (AUC) was also calculated. To offer a more complete assessment of the model's performance, the Area Under the Curve (AUC) was also calculated. The AUC value, constrained between 0.5 and 1, provides a single numeric value to represent the model's classification ability, where a higher AUC signifies a superior classifier. This makes the AUC value a more intuitive and direct evaluation metric than the ROC curve itself.

Table 1: Assessment of network models used in this evaluation study, along with their benchmark results of testing The metrics—precision, recall, accuracy, F1 score, AUC—are shown as mean values ± standard error/deviation.

	Proposed Network		Transfer learning		
Network	Using standard CE-loss	Using RadCE-loss	VGG-19	ResNet50	Inception V3
Parameters	8.7 M	8.7 M	145.2 M	24.8 M	26.2 M
Accuracy (%)	95.46 ± 1.42	97.46 ± 1.23	60.92 ± 1.24	88.98 ± 1.58	90.82 ± 1.68
Recall (%)	95.48 ± 1.68	97.59 ± 1.48	61.64 ± 1.72	86.16 ± 2.18	89.05 ± 1.52
Precision (%)	95.54 ± 1.35	97.63 ± 1.19	62.08 ± 0.96	91.24 ± 1.72	91.78 ± 1.06
F_1 score (%)	95.51 ± 1.56	97.60 ± 1.28	61.86 ± 1.25	88.63 ± 1.84	90.39 ± 1.18
AUC	0.967 ± 0.007	0.985 ± 0.005	0.671 ± 0.018	0.915 ± 0.010	0.928 ± 0.012
Training time (100 epochs)	4686 s	4862 s	13,156 s	11,458 s	12,454 s
Single inference time	115 ms	115 ms	328 ms	302 ms	315 ms

Table 1 displays the key statistical metrics for the suggested PneumoLiteNet model utilizing RADCELoss, along with a variant that employs the standard cross-entropy loss in place of RadCE-loss in the convolutional layers. These results are compared to three popular pre-existing architectures implemented via transfer learning. Figure 5 illustrates the average performance indicators for each network model individually. Notably, the Additionally, due to its reduced number of trainable parameters, the proposed model demonstrates greater computational

efficiency, requiring an average of only 115 milliseconds to generate a single inference for an image. This prediction time remains consistent for both the RadCE-loss and standard loss function models, as the loss function is applied only during the training phase. In comparison, transfer-learning model networks require approximately three times the computational resources for a typical inference process. VGG-19 model was omitted from Figure 4 to highlight the performance distinctions among the advanced architectures. Furthermore, the dropout rate in the proposed network model was critical in optimizing performance. The best results were achieved with a dropout rate of 0.4, across all evaluation metrics. Dropout values below 0.2 didn't appreciably enhance metrics outcome compared to model without dropout, while rates above 50% hindered classification accuracy. These findings suggest that a carefully adjusted dropout rate in the network layers can substantially improve accuracy and convergence speed. The integration of dropout in the convolutional layers not only enhances classification accuracy but also accelerates convergence.Proposed architecture utilizing RadCE-loss outperformed all other models tested across the four key statistical metric recall, precision, F1 score and accuracy consistently exceeding 97%. The same network model, when using the standard cross-entropy loss, achieved the second-highest performance at 95.5%, while the Inception architecture followed with approximately 90%. This demonstrates a substantial performance gap, not only between the proposed model with RadCE-loss and the transfer learning architectures but also when compared to the model using the standard loss function.A notable distinction of approximately 2% was observed across all statistical indicators when comparing the two proposed architectures, favouring the model utilizing RadCE-loss in its convolutional layers. The low standard deviation of accuracy measurements further suggests that the proposed model maintains consistent performance across all tests, leading to reliable classifications. Table 1 also includes the AUC values derived from ROC curve analysis, presented in a mean and standard error/deviation format. Figure 5 depicts the ROC curves corresponding to mean AUC values for each network. These results further highlight the proposed network model with more perfect CXR image classifications.

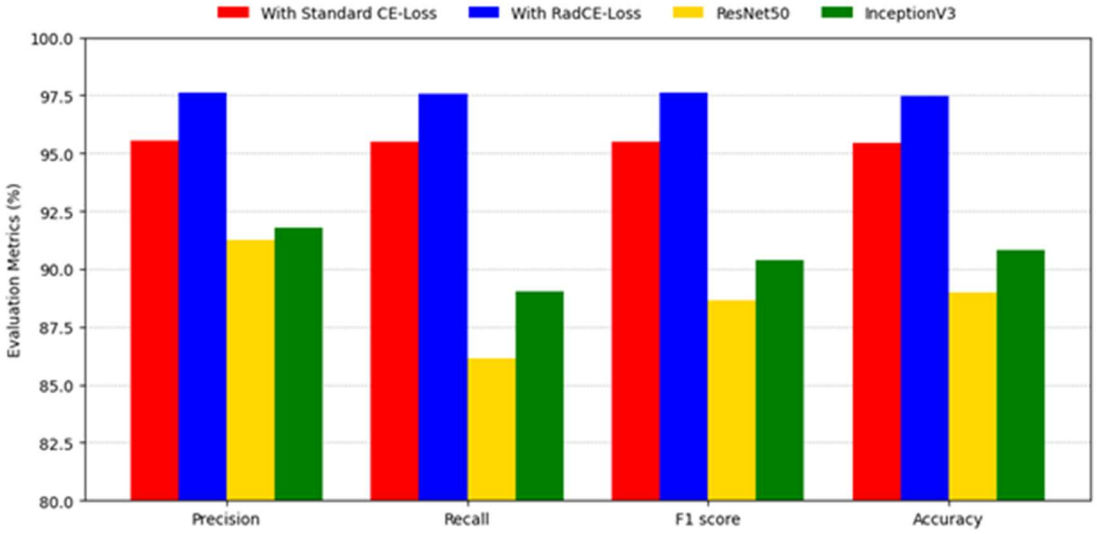


Fig 4.– Performance comparison of average precision, recall, accuracy, and F1 scores across different evaluated network models.

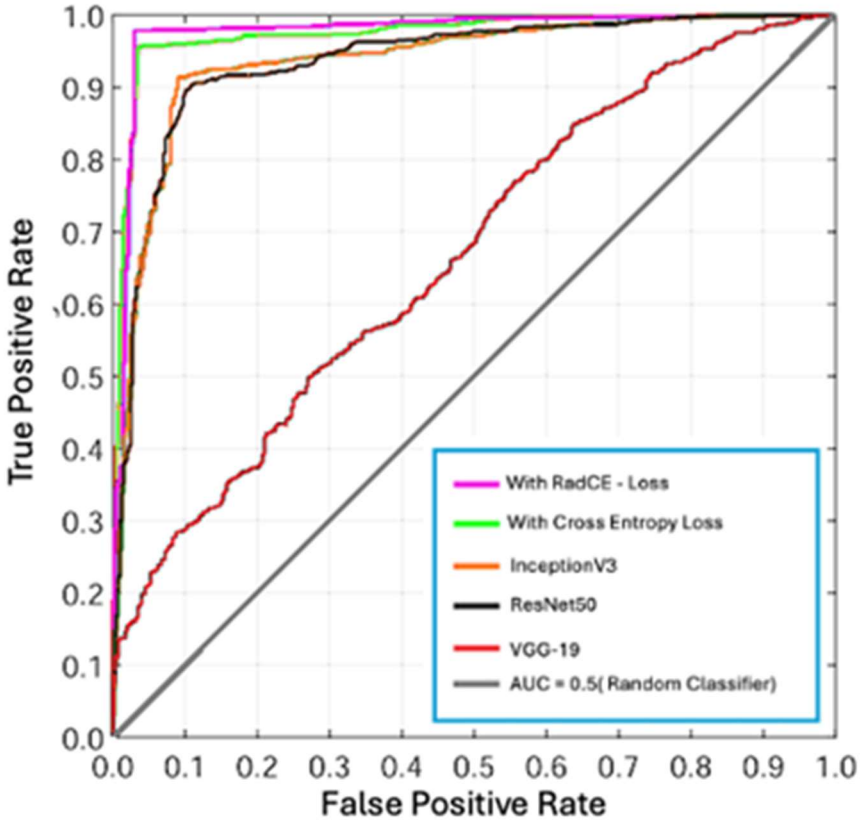


Fig 5. ROC curves for different network architectures.

Additionally, due to its reduced number of trainable parameters, the proposed model demonstrates greater computational efficiency, requiring an average of only 115 milliseconds to generate a single inference for an image. This prediction time remains consistent for both the RadCE-loss and standard loss function models, as the loss function is applied only during the training phase. In comparison, transfer-learning model networks require approximately three times the computational resources for a typical inference process. VGG-19 model was omitted from Figure 4 to highlight the performance distinctions among the more advanced architectures.

Furthermore, the dropout rate in the proposed network model was critical in optimizing performance. The best results were achieved with a dropout rate of 0.4, across all evaluation metrics. Dropout values below 0.2 didn't appreciably enhance metrics outcome compared to model without dropout, while rates above 50% hindered classification accuracy. These findings suggest that a carefully adjusted dropout rate in the network layers can substantially improve accuracy and convergence speed. The integration of dropout in the convolutional layers not only enhances classification accuracy but also accelerates convergence.

3.2 Discussion

The primary conclusion of this work is the advantageous impact of using a lightweight model with significantly fewer parameters and incorporating a novel loss function for the convolutional neural network architecture. Besides facilitating converging of the model faster during training the model, the proposed model demonstrates superior accuracy in both validation and final inference phases. The utilization of this network for pneumonia identification / classification, utilizing chest radiographs, further supports the idea that smaller, customized CNN architectures can surpass bigger models implemented by transfer learning.The efficacy of the suggested PneumoLiteNet network was compared against established models within identical experimental settings, as presented in the tables and figures above. The results indicate that the newly implemented network outperforms its counterparts in both efficiency and accurac

Table 2: Assessment of outcomes from various architectural models in the existing literature and proposed

Study	Technique	Dataset	F_1 Measure	Execution time
Panwar et al.[9]	VGG-19	6382 X-Rays + 2482 CT	95.6%	2 s
Ouchicha et al.[24]	CVD-Net	2905 X-Rays	96.7%	NA
Adi and Kemal et al[11]	DenseNet	1218 X-Rays	94.95%	'few seconds'
Wang et al.[13]	3D-ResNet	4697 X-Rays	93.3%	NA
Nikolaou and Massaro et al[15]	EfficientNet	15153 X-Rays	95%	NA
Singh and Tripathi et al[16]	Quaternion network	5856 X-Rays	93.75%	NA
Joshi and Yadav et al.[17]	DarkNet	6884 X-Rays	97.1%	0.137 s
Dash and Mohapatra et al.[18]	transfer learning + CNN	1272 X-Rays	97.1%	NA
Mahmud et al.[12]	CovXNet	6161 X-Rays	97.4%	NA
Proposed method	customized CNN (PneumoLiteNet)	5856 X-Rays	97.6%	0.122 s

Table 2 compares our architecture's performance with various contemporary advanced models. However, a direct assessment of the stated metric results may not always be unbiased, as various network models have been tested under different circumstances by their respective authors. Nonetheless, our lightweight network model with the novel loss function is highly competitive among its peers based on the metrics like F1-score and accuracy. Further analysis of the classification problem shows that the methods tested are designed to distinguish between two and four classes, depending on whether the focus is on distinguishing pneumonia from healthy lungs or categorizing various types of pneumonia. Reported F1 scores in previous studies range between 97% and 98%, which aligns with the accuracy range of our proposed network model. When evaluating performance against time, it is worth noting that only a few of the methods listed provide efficiency benchmarks. Those that do often report inference times lasting one or more seconds per test. In contrast, our model, with a 122-millisecond inference time, is highly efficient and competitive with the fastest candidate models.

To obtain a more objective comparison, our final evaluation phase included testing our network against some of the cutting-edge industry networks, such as ResNet50, InceptionV3, and VGG-19. All models were given the same pre-processed input, and the procedure of training was identical, with no changes to the parameters. The proposed network consistently outperformed the others across all metrics:, recall, precision, F1 score, and AUC. The outcomes confirm that PneumoLiteNet is effective in classifying CXR images, and the use of RadCE-loss significantly enhances performance compared to the standard cross-entropy loss.Table 1 clearly demonstrates our network's efficiency in comparison to other networks, with all models trained for 100 epochs. The smaller size of the proposed model greatly reduced training time, contributing to its overall performance superiority. This efficiency is further evident in the reduced average time required for a single prediction. In terms of model size, our architecture demonstrates a remarkable improvement over the others, especially when considering parameter count. This explains the markedly shorter training time and reduced inference duration for our proposed network.

The primary drawback of this study is its solely on a singular collection of X-ray scans. Another limitation is the exclusive use of CXR data, without incorporating CT images. Nonetheless, this can also be seen as an advantage, given that X-ray imaging is significantly more accessible than CT.

Most of the models for detection of pneumonia, including ours, were trained retrospectively using past cases from publicly available databases or private CXR collections. These datasets are unlikely to fully represent the population at large.

Therefore, clinical implementation would require periodic retraining of the main model that makes the final decisions using contemporary examples from the patient community at hand. A critical concern is the dependability of AI-driven decisions, that should be reproducible when the test is repeated under similar conditions. It is essential to ensure that AI models be trained well to base their decisions on pertinent medical image traits rather than circumstantial factors. Explainable AI models are gaining attention in medical decision-making as they offer transparency, helping clinicians understand how decisions are made, and are likely to see increasing adoption in this field[24].

4 CONCLUSION

Chest X-ray Imaging is extensively employed for detecting Lung lesions and deep learning has proven to be an effective method to assist in this diagnosis. While fine-tuning existing deep networks is promising for CXR classification tasks, especially given constraints in data size, labeling and computational resources, it can sometimes result in low transfer efficiency and overfitting, particularly when the original and the final domain datasets significantly differ. To tackle these challenges, we analyzed a database consisting of image scans of Chest X rays and introduced a novel Convolutional network, **PneumoLiteNet** tailored for anomaly detection in CXR images. Additionally, we proposed a new innovative loss function, **RadCE-loss**, which further enhances the performance of PneumoLiteNet. Experimental results demonstrate that PneumoLiteNet, despite having fewer layers and parameters, achieves superior accuracy compared to models relying on fine-tuning, whether RadCE-loss is applied or not. The inclusion of RadCE-loss further enables improving its overall performance.

Future research will focus on designing models and loss functions tailored to different classification tasks. For instance, integrating LSTM could allow the model to capture relationships between multiple labels, while further refinement of both CNN architectures and loss functions will remain a key area of our ongoing research.

REFERENCES

[1] Li YY, Zhang ZY, Dai C, Dong Q, Badrigilan S. Accuracy of deep learning for automated detection of pneumonia using chest X-Ray images: A systematic review and meta-analysis. Comput Biol Med 2020;123 . https://doi.org/10.1016/j.compbiomed.2020.103898 103898.

[2] WHO Pneumonia. World Health Organization. (2019), https://www.who.int/news-room/fact-sheets/detail/pneumonia

[3] Neuman M., Lee E., Bixby S., Diperna S., Hellinger J., Markowitz R., et al. Variability in the interpretation of chest radiographs for the diagnosis of pneumonia in children. Journal Of Hospital Medicine. 7, 294–298 (2012) https://doi.org/10.1002/jhm.955 PMID: 22009855

[4] Kadry S., Nam Y., Rauf H., Rajinikanth V. & Lawal I. Automated Detection of Brain Abnormality using Deep-Learning-Scheme: A Study. 2021 Seventh International Conference On Bio Signals, Images, And Instrumentation (ICBSII). pp. 1-5 (2021)

[5] Sharma H., Jain J., Bansal P. & Gupta S. Feature extraction and classification of chest x-ray images using cnn to detect pneumonia. 2020 10th International Conference On Cloud Computing, Data Science & Engineering (Confluence). pp. 227-231 (2020)

[6] Stephen O., Sain M., Maduh U. & Jeong D. An efficient deep learning approach to pneumonia classification in healthcare. Journal Of Healthcare Engineering. 2019 (2019) https://doi.org/10.1155/2019/4180949 PMID: 31049186

[7] Kermany D., Zhang K. & Goldbaum M. Labeled Optical Coherence Tomography (OCT) and Chest XRay Images for Classification. (Mendeley,2018)

[8] Rajpurkar P., Irvin J., Zhu K., Yang B., Mehta H., Duan T., et al. & Others Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. ArXiv Preprint ArXiv:1711.05225. (2017)

[9] Panwar H, Gupta PK, Siddiqui MK, Morales-Menendez R, Bhardwaj P, Singh V. A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. Chaos, Solitons & Fractals 2020;140 . https://doi.org/10.1016/j.chaos.2020.110190

[10] Brunese L, Mercaldo F, Reginelli A, Santone A. Explainable Deep Learning for Pulmonary Disease and Coronavirus COVID-19 Detection from X-rays. Comput Meth Prog Biomed 2020;196 .

https://doi.org/10.1016/j.cmpb.2020.105608
[11] Alhudhaif Adi, Polat Kemal, Karaman O. Determination of COVID-19 pneumonia based on generalized convolutional neural network model from chest X-ray images. Expert Syst Appl 2021;180 . https://doi.org/10.1016/j.eswa.2021.115141.
[12] Mahmud T, Rahman MA, Fattah SA. CovXNet: A multidilation convolutional neural network for automatic COVID19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. Comput Biol Med 2020;122 . https://doi.org/10.1016/ j.compbiomed.2020.103869
[13] Wang J, Bao YM, Wen YF, Lu HB, Luo H, Xiang YF, Li XM, Liu C, Qian DH. Prior-Attention Residual Learning for More Discriminative COVID-19 Screening in CT Images. IEEE Trans Med Imag 2020;39(8):2572–83. https://doi.org/10.1109/ TMI.2020.2994908.
[14] Das AK, Ghosh S, Thunder S, Dutta R, Sachin A, Chakrabati A. Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network. Pattern Anal Appl 2021;24:1111–24. https://doi.org/10.1007/ s10044-021-00970-4.
[15] Nikolaou V, Massaro S, Fakhimi M, Stergioulas L, Garn W. COVID-19 diagnosis from chest x-rays: developing a simple, fast, and accurate neural network. Health Inf Sci Syst 2021;9:36. https://doi.org/10.1007/s13755-021-00166-4.
[16] Singh S, Tripathi BK. Pneumonia classification using quaternion deep learning. Multimed Tools Appl 2022;81 (2):1743–64. https://doi.org/10.1007/s11042-021-11409-7.
[17] Joshi RC, Yadav S, Pathak VK, Malhotra HS, Khokar HVS, Parihar A, Kohli N, Himanshu D, Garg RK, Bhatt MLB, Kumar R, Singh NP, Sardana V, Burget R, Alippi C, Travis-Gonzalez CM, Dutta MK. A deep learning-based COVID-19 automatic diagnostic framework using chest X-ray images. Biocybern Biomed Eng 2021;41:239–54. https://doi.org/10.1016/j. bbe.2021.01.002.
[18] Dash AK, Mohapatra P. A fine-tuned deep convolutional neural network for chest radiography image classification on COVID-19 cases. Multimed. Tools Appl., available online 21 Sep 2021. doi: 10.1007/s11042-021-11388-9.
[19] K. Simonyan and A. Zisserman. (Sep. 2014). ‘‘Very deep convolutional networks for large-scale image recognition.’’ [Online]. Available: https://arxiv.org/abs/1409.1556
[20] C. Szegedy et al., ‘‘Going deeper with convolutions,’’ in Proc. CVPR, 2015, pp. 1–9
[21] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, ‘‘Densely connected convolutional networks,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., vol. 1, no. 2, 2017, p. 3.
[22] Kingma DP, Ba J. Adam: a method for stochastic optimization. ArXiv: 1412.6980, 2014. https://arxiv.org/abs/1412.6980
[23] Thakkar V, Tewary S, Chakraborty C. Batch Normalization in Convolutional Neural Networks – A comparative study with CIFAR-10 data. In: 5th International Conference on Emerging Applications of Information Technology. p. 1–5. https://doi.org/10.1109/EAIT.2018.8470438.
[24] Ouchicha C, Ammor O, Meknassi M. CVDNet: A novel deep learning architecture for detection of coronavirus (Covid-19) from chest x-ray images. Chaos, Solitons & Fractals 2020;140 . https://doi.org/10.1016/j.chaos.2020.110245 110245.
[25] S. Suguna Mallika, A. Sanjana, A. Vani Gayatri, S. Veena Naga Sai, Sign Language Interpretation Using Deep Learning, Morusupalli, R., Dandibhotla, T.S., Atluri, V.V., Windridge, D., Lingras, P., Komati, V.R. (eds) Multi-disciplinary Trends in Artificial Intelligence. MIWAI 2023. Lecture Notes in Computer Science, vol 14078. Springer, Cham, doi:https://doi.org/10.1007/978-3-031-36402-0_64
[26] M. Naidu Vadlamudi, N. Jayanthi, G. Swetha, P. Nishitha, G. A. Al-Salman and K. Saikumar, "IoT Empowered GNSS Tracking in Real-time via Cloud Infrastructure," 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), Pune, India, 2024, pp. 1-6, doi: 10.1109/I2CT61223.2024.10544102.
[27] Aruna, K., Gatla, S. (2024). Graph Based Semantically Extractive Tool for Text Summarization Using Similarity Score. In: Satheeskumaran, S., Zhang, Y., Balas, V.E., Hong, Tp., Pelusi, D. (eds) Intelligent Computing for Sustainable Development.

ICICSD 2023. Communications in Computer and Information Science, vol 2121. Springer, Cham. https://doi.org/10.1007/978-3-031-61287-9_10

[28] S. Bethu and N. S. Chandra, "Image Captioning with CNN and LSTM using Python," 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2023, pp. 1383-1390, doi: 10.1109/ICAIS56108.2023.10073841. keywords:

[29] Gouri Patil, Srikanth Bethu, Y V K Durga Bhavani, B Harikrishna, Adluri Vijayalakshmi, and B Sankara Babu. 2022. IoT and Cloud based Face detection application design for Surveillance systems using Deep Learning. In Proceedings of the 2022 6th International Conference on Cloud and Big Data Computing (ICCBDC '22). Association for Computing Machinery, New York, NY, USA, 49–55. https://doi.org/10.1145/3555962.3555971

[30] K. Sree Divya et al., "Implementing Blockchain Based DApp for Secure Sharing of Students' Credentials," 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), Pune, India, 2024, pp. 1-6, doi: 10.1109/I2CT61223.2024.10543559.

[31] C. Sarada, V. Dattatreya and K. V. Lakshmi, "Deep Learning based Breast Image Classification Study for Cancer Detection," *2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS)*, Raichur, India, 2023, pp. 01-08, doi: 10.1109/ICICACS57338.2023.10100206.

[32] Ch.Sarada, C., Lakshmi, K. V. ., & Padmavathamma, M. . (2023). MLO Mammogram Pectoral Masking with Ensemble of MSER and Slope Edge Detection and Extensive Pre-Processing. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 135–144. https://doi.org/10.17762/ijritcc.v11i3.6330

[33] D. Rajya Lakshmi, S. Suguna Mallika, A Review on Web Application Testing and its Current Research Directions, International Journal of Electrical and Computer Engineering, Volume 7, No.4, August 2017. DOI : 10.11591/ijece.v7i4.pp2132-2141

[34] V. S. Reddy T, S. Parvathi, C. Sarada, K. H. Mohmmed, A. Sampath Dakshina Murthy and K. Saikumar, "Enterprise Support Hierarchical Model with Secure Data Protocol," *2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, Bangalore, India, 2023, pp. 1-6, doi: 10.1109/SMARTGENCON60755.2023.10442282.