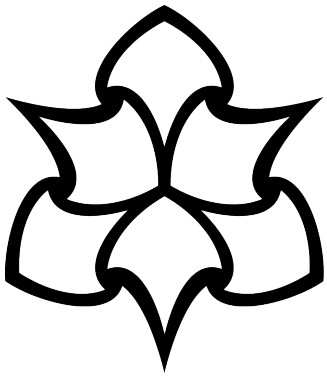


BIOCYBERNETICS AND BIOMEDICAL ENGINEERING 43 (2023) 528—550

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bbe

Original Research Article

Detection of various lung diseases including COVID-19 using extreme learning machine algorithm based on the features extracted from a lightweight CNN architecture

Md. Nahiduzzaman [a], *Md Omaer Faruq Goni* [a], *Md. Robiul Islam* [a], *Abu Sayeed* [b], *Md. Shamim Anower* [c], *Mominul Ahsan* [d], *Julfikar Haider* [e], *Marcin Kowalski* [f,*]

[a] *Department of Electrical & Computer Engineering, Rajshahi University of Engineering & Technology, Rajshahi 6204, Bangladesh*
[b] *Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Rajshahi 6204, Bangladesh*
[c] *Department of Electrical & Electronic Engineering, Rajshahi University of Engineering & Technology, Rajshahi 6204, Bangladesh*
[d] *Department of Computer Science, University of York, Deramore Lane, Heslington, York YO10 5GH, UK*
[e] *Department of Engineering, Manchester Metropolitan University, Chester St, Manchester M1 5GD, UK*
[f] *Institute of Optoelectronics, Military University of Technology, Gen. S. Kaliskiego 2, Warsaw, Poland*

ARTICLE INFO

Article history:
Received 18 December 2022
Received in revised form
4 April 2023
Accepted 16 June 2023
Available online 26 June 2023

Keywords:
COVID-19
Convolutional Neural Network
Extreme Learning Machine
Pearson Correlation Coefficient
Pneumonia

ABSTRACT

Around the world, several lung diseases such as pneumonia, cardiomegaly, and tuberculosis (TB) contribute to severe illness, hospitalization or even death, particularly for elderly and medically vulnerable patients. In the last few decades, several new types of lung-related diseases have taken the lives of millions of people, and COVID-19 has taken almost 6.27 million lives. To fight against lung diseases, timely and correct diagnosis with appropriate treatment is crucial in the current COVID-19 pandemic. In this study, an intelligent recognition system for seven lung diseases has been proposed based on machine learning (ML) techniques to aid the medical experts. Chest X-ray (CXR) images of lung diseases were collected from several publicly available databases. A lightweight convolutional neural network (CNN) has been used to extract characteristic features from the raw pixel values of the CXR images. The best feature subset has been identified using the Pearson Correlation Coefficient (PCC). Finally, the extreme learning machine (ELM) has been used to perform the classification task to assist faster learning and reduced computational complexity. The proposed CNN-PCC-ELM model achieved an accuracy of 96.22% with an Area Under Curve (AUC) of 99.48% for eight class classification. The outcomes from the proposed model demonstrated better performance than the existing state-of-the-art (SOTA) models in the case of COVID-19, pneumonia, and tuberculosis detection in both binary and multiclass classifications. For eight class classification, the proposed model achieved precision, recall and fi-score and ROC are 100%, 99%, 100% and 99.99% respectively for COVID-19 detection demonstrating its robustness. Therefore, the proposed model has overshadowed the exist-

* Corresponding author at: Military University of Technology, Warsaw, Poland.
E-mail address: marcin.kowalski@wat.edu.pl (M. Kowalski).
https://doi.org/10.1016/j.bbe.2023.06.003

ing pioneering models to accurately differentiate COVID-19 from the other lung diseases that can assist the medical physicians in treating the patient effectively.

1. Introduction

In the last few decades, several lung-related diseases have become an epidemic to humans, such as pneumonia, pleural, cardiomegaly, TB, and more recently the most life-threatening disease, COVID-19. Approximately 7% of the global population (450 million) is affected by pneumonia alone, and every year approximately 2 million people lose their lives due to the pneumonia [1]. In the past three years, almost 6.27 million people died due to COVID-19, and nearly 522 million people were affected [2]. In 2020, around 1.3 million people died due to TB; hence, in total 4 million people died due to lung-related disease every year [3]. These types of diseases are generally detected using CXR images of the lung by the radiologists as it is cheap and requires only fewer steps to detect these diseases. More than 35 million CXR images are taken every year in the US alone for medical treatment [4]. Many radiologists now have to assess more than 100 CXR images every day, resulting in an increased workload and fatigue and eventually wrong diagnosis [4]. Hence, the conventional process is very time-consuming and expensive. Furthermore, the diagnosis made by the radiologists may vary due to an error in human judgement. If COVID-19, pneumonia, TB, etc. are predicted at the early stages, it may save millions of human lives. Gradually, newer variants of COVID-19 for example, Delta, Omicron, etc. are spreading at a faster rate, hence it is too difficult for the medical physicians to cope up with the high demand for treatment on time. As a result, automated technologies that are trained to forecast the symptoms of a distinct lung-related anomalies based on a specific CXR image have the potential to aid the radiologists and medical physicians in accurate and faster diagnosis with a higher level of confidence. Hence, a computer aided intelligent system based on machine learning and deep learning has been designed to detect the lung related disease automatically based on the CXR images.

In the past, several efforts have been made to automatically identify lung-related anomalies using the CXR images due to the recent availability of the larger data sets. Loey et al. developed a CNN model for extracting the prominent features and finally a Bayesian model was used to classify the COVID-19 patient [5]. Moreover, they used 3,616 COVID-19 data and achieved an accuracy of 96%. Bhattacharyya et al. utilized a pre-trained CheXNet model and trained this model using 1,326 images of COVID-19 to develop a COVID-CXNet model [6]. For preprocessing they used Contrast Limited Adaptive Histogram Equalization (CLAHE) and achieved an accuracy of 87.88%. Ieracitano et al. distinguished between CXR images of patients with idiopathic pneumonias unrelated to COVID-19 and those of patients with COVID-19 pneumonia by suggesting a fuzzy logic-based deep learning (DL) technique [7]. In addition, they used 121 images of COVID-19 pneumonia patient and their model's accuracy was approximately 81% and provided an explainable artificial intelligence method to aiding doctors. Agrawal and Choudhary developed a deep CNN to detect the COVID-19 and pneumonia from the CXR images [8]. They handled the class imbalance problem using the oversampling technique, namely Synthetic Minority Over-sampling Technique (SMOTE). Moreover, they considered two (Covid-19 Vs Normal) and three classes (COVID-19 Vs Pneumonia Vs Normal), while they used 1,525 samples of COVID-19 patients and achieved an accuracy of 96% and 94.45% for the two and three classes respectively. Gayathri et al. extracted features using pre-trained CNN models, for instance InceptionResnetV2, Resnet101, etc., reduced the dimensioanlity using sparse auto-encoder and finally used a feed forward neural network to detect the COVID-19 [9]. In this study, they used 504 COVID-19 images to train their models and achieved an accuracy of 95.78% and AUC of 98.21%. Kassani et al. compared different transfer learning (TL) models to detect the COVID-19 from CXR and computed tomography (CT) images [10]. Firstly, they used eight TL models: MobileNet, DenseNet, Xception, ResNet, InceptionV3, Inception- ResNetV2, VGGNet, and NASNet for extracting the features from the images and several ML models were developed. Finally, 117 CXR and 20 CT images of COVID-19 patients were used to train their models and achieved the highest accuracy of 99% using the Bagging tree classifier with DenseNet121 as a feature extractor. Yousefi et al. segmented the lung lobes using a 2D U-Net model to detect COVID-19 from CXR images [11]. The authors reduced dimensions using several techniques, for instance, Laplacian scoring and principal component analysis (PCA). They used 704 CXR images for training their model and achieved 89.6%, 72.6% accuracy for two-class, multiclass classifications, respectively. Akter et al. used various TL models such as VGG19, GoogLeNet, etc., to detect the COVID-19 from the CXR images [12]. The author used the augmentation technique to balance the datasets and used 52,000 CXR images to train their models and achieved a high accuracy of 98% using MobileNetV2 for binary classification. At the same time, the compilation time was 2 h, 50 min and 21 s. Chowdhury et al. used several pre-trained TL models to detect the viral and COVID-19 pneumonia from CXR images [13]. This study combined several datasets for training the models and achieved an accuracy of 99.7%, 97.9% for two and three-class classification, respectively. Horry et al. proposed a suitable CNN to detect the COVID-19 from multimodal imaging data [14]. They also removed the noise from images and achieved a precision of 86%, 100%, and 84% for CXR, ultrasound, and CT scans, respectively. Rasheed et al. proposed two classifiers: LR and CNN, to diagnose COVID-19 [15]. The

generative adversarial network was used for data augmentation and PCA to select the most prominent features from 308 CXR images for training their model and achieved an accuracy of 97.6% using PCA. Panwar et al. utilized the VGG19 model to fast detect the COVID-19 from CXR and CT scan images [16]. Minaee et al. used 5000 CXR images to detect the COVID-19 using four TL models and achieved a specificity of 92.9% [17]. Afshar et al. developed capsule networks named COVID-CAPS to detect COVID-19 from 13,975 CXR images and achieved an accuracy of 95.7% [18]. Abbas et al. developed a Decompose, Transfer, and Compose (DeTraC) CNN model to classify the COVID-19 and achieved an accuracy of 93.1% [19]. Arias-Londono et al. used more than 79,500 CXR images for training their CNN model to detect COVID-19 and achieved an accuracy of 91.5% [20]. Alam et al. fused histogram-oriented gradient (HOG) and CNN for extracting the features from CXR images to detect COVID-19 [21].

Pandit et al. proposed the VGG19 TL model to detect COVID-19 from CXR images [22]. They used 1428 images for training their model and achieved an accuracy of 92.53%, and 96% for two- (COVID-19 vs Normal) and three-class (Normal vs Bacterial Pneumonia vs COVID-19) classification. Sekeroglu and Ozsahin et al. used several DL and ML models to detect COVID-19 and pneumonia from CXR images and performed 38 experiments using CNN, 10 experiments using five ML models, and 14 experiments using pre-trained TL models [23]. For two classifications, they used 1808 images to train their models and achieved a mean receiver operating characteristic (ROC) of 96.51%, and for the three-class classification, they used 6100 images and achieved a macro averaged F1 score of 94.10%. Khan et al. developed a CoroNet based on the pre-trained Xception model to identify COVID-19 and pneumonia [24]. For 3 class classification, their model achieved an accuracy of 95% and for 4 class (Normal vs COVID-19 vs bacterial vs Viral Pneumonia) achieved a precision and recall of 93% and 98.2% respectively. Nahiduzzaman et al. extracted the most discriminant features using a hybrid CNN-PCA from the CXR images to detect multivariant pneumonia [25]. They used ELM to discriminate viral pneumonia from bacterial pneumonia. For enhancing the contrast of the images, they used contrast limited adaptive histogram equalization (CLAHE) and trained their model using 5857 images and achieved an accuracy of 99.83% and 98.32% for two (Normal vs Pneumonia) and three-class (Normal vs bacterial vs Viral Pneumonia) classifications. Yamac et al. developed a convolutional sparse support estimator network based on a neural network to detect COVID-19, viral, and bacterial pneumonia from CXR images [26] using 6200 images and achieved an accuracy of 87.07% for a four-class classification, while 95.90% for COVID-19. Chandra et al. developed an automatic screening model to detect COVID-19 and pneumonia from CXR images [27]. 2,088 images were used for training the ML models and achieved the highest accuracy of 91.329% using support vector machine (SVM) with linear kernel. Nahiduzzaman et al. proposed a model ChestX-ray6 based on CNN to detect pneumonia from other lung-related diseases and achieved an accuracy of 97.60% [28]. Que et al. used U-NET and DenseNet to detect cardiomegaly disease from CXR images [29] through performing segmentation using U-NET and marking two separate parts namely cardiac and thoracic. By using augmentation, 2,630 images were produced for training their models and achieved the highest area under the receiver operating characteristics (AUROC) of 93.48% and accuracy of 93.75% for U-NET. On the contrary, Robiul et al. proposed an ensemble model based on ML models to detect COVID-19 with an optimistic accuracy of 99.73% [30]. Serte and Serener used a pre-trained ResNet model to classify pleural effusions (PE) from TB, pneumonia, and COVID-19 diseases [31]. They correctly detected PE with an accuracy of 99%, 75%, and 100% from pneumonia, TB, and COVID-19, respectively. For multiclass classification, they achieved an average accuracy of 83%. Sahlol et al. extracted the 50,000 features using a pre-trained MobileNet model from CXR images to detect TB disease [32]. They used an artificial ecosystem-based optimization algorithm to select 25 and 19 relevant features from the datasets named Shenzhen (SZ) and Dataset 2, respectively. Their proposed models outperformed the SOTA methods and achieved an accuracy of 90.20% and 94.10% for SZ and Dataset 2, respectively. Chandra et al. developed a computer aided diagnosis (CAD) system to detect TB disease from CXR images [33]. First, a guided image filter was used for image de-noising followed by lung segmentation. After feature extraction, SVM was employed for classification and accuracies of 95.60% and 99.40% were achieved for Montgomery (MT) and SZ datasets, respectively. Tawsifur et al. used nine TL and two U-net models to detect TB from CXR images [34]. Several databases were merged to create a single database with 7,000 CXR images for training the proposed model. They also performed augmentation and achieved accuracies of 96.7% and 98.6% using ChexNet [35] and DenseNet201 [36] respectively. Furthermore, t-distributed stochastic neighbor embedding [37] was used for data visualization. Muhammad et al. used 7,000 CXR images for extracting the features using three TL models [38] using the eXtreme Gradient Boosting [39] package to detect TB. Ayaz et al. combined CNN based features and hand-crafted features through ensemble learning to classify TB from CXR images [40]. They achieved an AUROC [41] of 99% and 97% for the MT and SZ datasets, respectively. Lopes and Valiati et al. used three TL models to predict TB from CXR images [42] and achieved an accuracy of 83.40% and 82.60% for MT and SZ datasets, respectively.

From the above literature review, it was observed that most of the studies were limited to binary classes with a small amount of data in the datasets and failed to achieve promising performance particularly for multiclass classifications with large numbers of data. Several DL and ML approaches have been conducted, but most of them require high computational times and a large number of parameters. Furthermore, most of the studies detected COVID-19 with fewer data because of the scarcity of COVID-19 CXR images. For multiclass classification, the overall accuracy is again too poor with a large number of parameters. The goal of the proposed framework was to detect lung-related diseases with novel contribution from extracting features using lightweight CNN and classifying using extreme learning machine (ELM) to obtain fewer parameters, low computational times, and high classification performance.

The main contribution of this work is outlined as follows.

1. Several databases have been combined to create a more extensive database of eight diseases (CXR8), including COVID-19, with different schemes, as shown in Table 1. In scheme 1, eight classes have been considered; in scheme 2, four classes have been considered; and so on. Binary classes such as Normal and Tuberculosis or Normal and COVID-19 were chosen mainly to verify that the binary models could still produce higher accuracy than the multiclass models [1].

2. A lightweight CNN model with three layers has been proposed to extract features from 23,690 CXR8 images, including 4,192 COVID-19 CXR images.

3. To reduce the complexity and increase the speed, the PCC has been used to remove the redundant and irrelevant features.

4. An ELM has been proposed to detect multiple lung-related diseases. Consequently, the CNN-PCC-ELM model outweighs 28 SOTA models while reducing the parameters, complexities, and processing times compared to the TL models.

5. The proposed framework was developed through different schemes and achieved optimistic results in the case of every circumstance, namely balanced and imbalanced, low-resolution CXR images.

2. Proposed framework

Fig. 1 presents the proposed framework for detecting lung diseases from the CXR images. First, for making multi-class classification, several image databases were combined to create suitable customized databases. Then the resolutions of the CXR images were converted to a uniform size and subsequently normalization was performed. After pre-processing, a lightweight and straightforward CNN model was employed to extract 512 features. A PCC algorithm was used to identify 195 prominent features by eliminating 317 irrelevant features. These features were standardized using the z-score normalization technique. Finally, an ELM model classified the features to identify the lung diseases.

2.1. Database construction

Since the number of COVID-19 images in a particular publicly available database was not so large, therefore, several databases were combined to create a customized database with a total of 4,192 COVID-19 CXR images [13,43–47]. CXR images of 1,000 cardiomegaly, 1,500 pleural, and 1,500 lung opacity diseases were included in this study [48]. On the other hand, 1,037 CXR images of tuberculosis disease were collected from three different databases [49–51]. Bacterial and viral pneumonia CXR images were differentiated from the Kaggle database. Finally, the CXR images of normal patients have been collected from two databases [52,53]. Fig. 2 demonstrates the representative CXR images from each class. Hence, a total of 23,690 CXR images were used to construct eight classes where number of CXR images for normal, bacterial pneumonia, viral pneumonia, COVID-19, cardiomegaly, pleural, lung opacity, and tuberculosis were 10,192; 2,777; 1,493; 4,192; 1,000; 1,500; 1,500 and 1,036 respectively. Though the datasets were imbalanced, the proposed framework performed satisfactorily in the case of multi-class classification, which is revealed in the result section.

In this study, six types of ML schemes were considered to evaluate the performance of the proposed framework. Scheme 1 is a 8-class classification (normal, bacterial pneu-

Table 1 – Various schemes considered for lung disease classification.

Scheme 1							
Normal	Bacterial Pneumonia	Viral Pneumonia	COVID-19	Cardiomegaly	Pleural	Lung Opacity	Tuberculosis
Scheme 2				***Scheme 3***			
Normal	COVID-19	Bacterial Pneumonia	Viral Pneumonia	Normal	Bacterial Pneumonia	Viral Pneumonia	
Scheme 4			***Scheme 5***		***Scheme 6***		
Normal	Pneumonia	COVID-19	Normal	COVID-19	Normal	Tuberculosis	

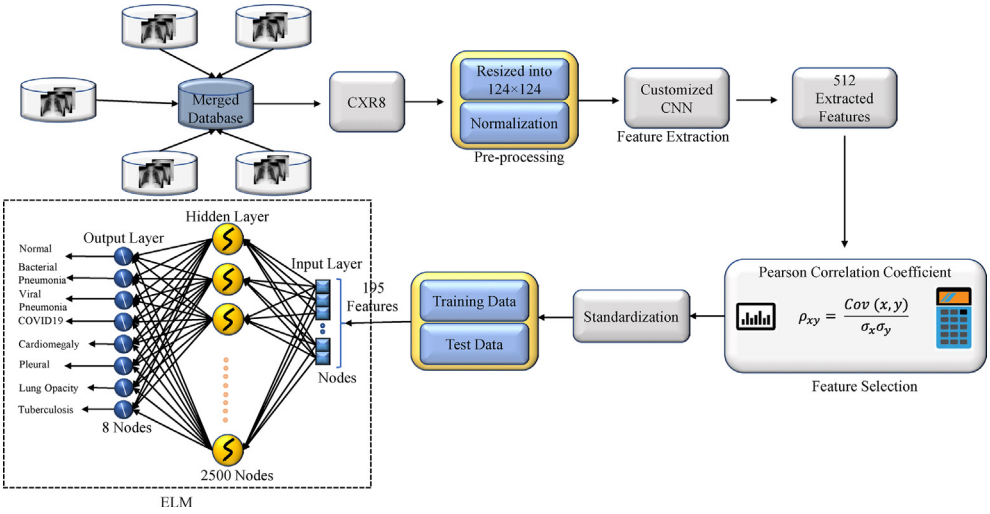


Fig. 1 – Proposed framework for lung-related disease classification.

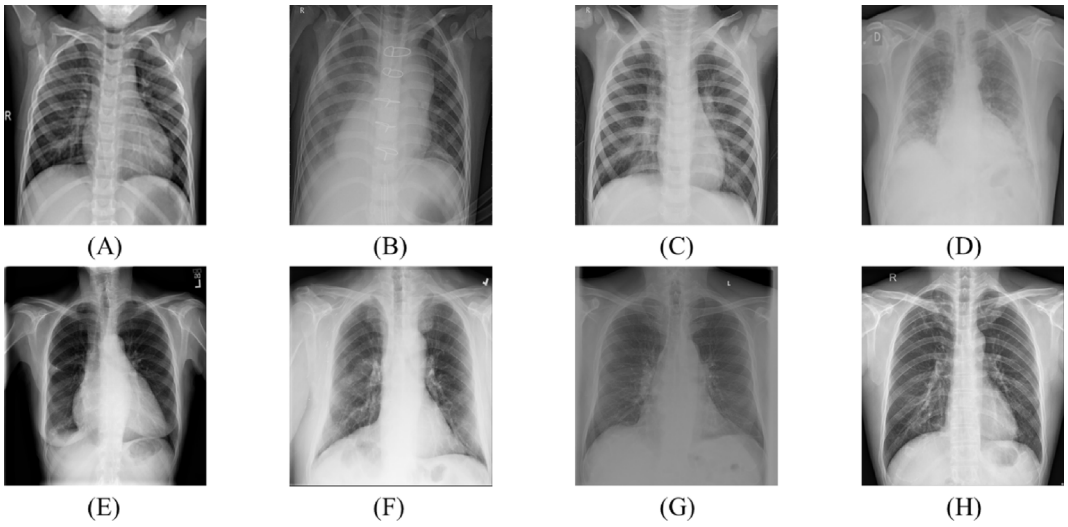


Fig. 2 – Sample CXR images of (A) Normal, (B) Bacterial Pneumonia, (C) Viral Pneumonia, (D) COVID-19, (E) Cardiomegaly, (F) Pleural, (G) Lung Opacity, and (H) Tuberculosis.

Table 2 – Datasets for different ML schemes with training and testing sets.

Scheme No	*Class type*	*Disease Types*	*Training Data*	*Test Data*
Scheme 1	8-class	Normal	8,663	1,529
		Bacterial Pneumonia	2,360	417
		Viral Pneumonia	1,269	224
		COVID-19	3,563	629
		Cardiomegaly	850	150
		Pleural	1,275	225
		Lung Opacity	1,275	225
		Tuberculosis	881	155
		Total	*20,136*	*3,554*
Scheme 2	4-class	Normal	8,153	2,039
		COVID-19	3,354	838
		Bacterial Pneumonia	2,222	555
		Viral Pneumonia	1,194	299
		Total	*14,923*	*3,731*
Scheme 3	3-class	Normal	8,153	2039
		Bacterial Pneumonia	2,222	555
		Viral Pneumonia	1,194	299
		Total	*11,569*	*2,893*
Scheme 4	3-class	Normal	8,153	2,039
		Pneumonia	1,194	299
		COVID-19	3,354	838
		Total	*12,701*	*3,176*
Scheme 5	2-class	Normal	8,153	2,039
		COVID-19	3,354	838
		Total	*11,507*	*2,877*
Scheme 6	2-class	Normal	8,153	2,039
		Tuberculosis	829	207
		Total	*8,982*	*2,246*

monia, viral pneumonia, COVID-19, cardiomegaly, pleural, lung opacity, tuberculosis), Scheme 2 is a 4-class classification (normal, COVID-19, bacterial pneumonia, viral pneumonia), Scheme 3 is a 3-class classification (normal, bacterial pneumonia, viral pneumonia), Scheme 4 is a 3-class classification (normal, COVID-19, pneumonia), Scheme 5 is a 2-class classification (normal, COVID-19), and Scheme 6 is a 2-class classification (normal, tuberculosis). Since the proposed framework was designed to achieve promising accuracy in case of multi-class classification (3, 4 and 8), Scheme 5 was also implemented to check whether the same framework could be able to achieve higher classification performances in case of

Table 3 – List of layers, output shapes and parameters for the model during feature extraction.

Layer (Type)	*Shape of Output*	*Parameters*
conv1 (Conv2D)	(None, 124, 124, 16)	1,216
bn1 (BatchNormalization)	(None, 124, 124, 16)	64
av1 (Activation)	(None, 124, 124, 16)	0
mp1 (MaxPooling2D)	(None, 62, 62, 16)	0
conv2 (Conv2D)	(None, 60, 60, 32)	4,640
bn2 (BatchNormalization)	(None, 60, 60, 32)	128
av2 (Activation)	(None, 60, 60, 32)	0
mp2 (MaxPooling2D)	(None, 30, 30, 32)	0
conv3 (Conv2D)	(None, 28, 28, 64)	18,496
bn3 (BatchNormalization)	(None, 28, 28, 64)	256
av3 (Activation)	(None, 28, 28, 64)	0
mp3 (MaxPooling2D)	(None, 14, 14, 64)	0
dp1 (Dropout)	(None, 14, 14, 64)	0
ft (Flatten)	(None, 12544)	0
dense (Dense)	(None, 1024)	12,846,080
bn4 (BatchNormalization)	(None, 1024)	4,096
av4 (Activation)	(None, 1024)	0
dp2 (Dropout)	(None, 1024)	0
Feature Extraction (Dense)	(None, 512)	524,800
Total Parameters	13,399,776	
Trainable Parameters	13,397,504	
Non-trainable Parameters	2,272	

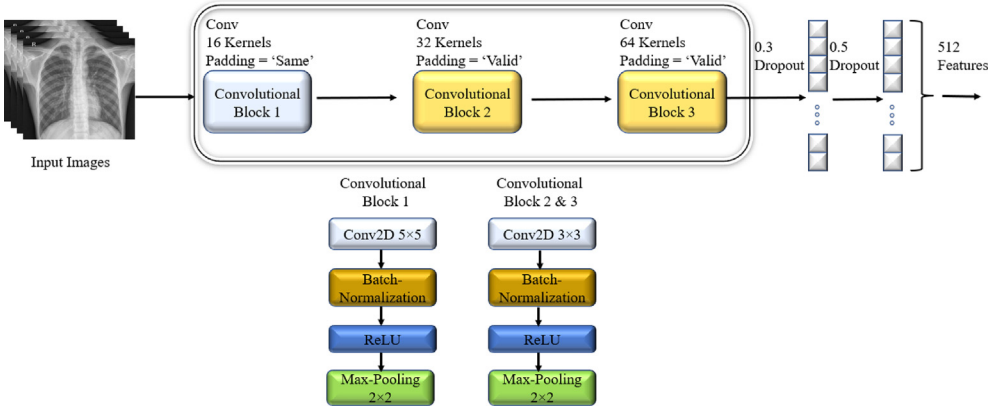


Fig. 3 – A lightweight CNN model employed for extracting features.

binary classification. The first experiment with eight-class classification was analyzed using a segregated 5-fold cross-validation (CV) procedure to ensure the model's performance was not biased and the result was demonstrated in Section 4.1. The training and testing ratios for the remaining schemes were 80:20, as the majority of SOTA models in the literature split data rather than using CV [22,23,25–27]. Table 2 shows the number of images per class split into training and test set.

Since the databases were constructed from multiple sources of CXR images, the resolutions of the images were different from each other. Hence, the CXR images were resized to a dimension of 124×124. An image is represented by a number of pixel values ranging from 0 to 255, which adds complexity. To remove this complexity, normalization was carried out by dividing each image by 255 and converting the range within 0 to 1. Finally, the CXR images were ready to feed into the CNN model for feature extraction.

2.2. *Feature extraction*

A lightweight CNN model was proposed for extracting new features from the raw pixel values of the CXR images. It has only three layers, fewer parameters than the TL model and requires less processing time. On the contrary, the TL models such as VGG19, ResNet50, ResNet152V2, InceptionResNetV2, and DenseNet201 have 19, 50, 152, 164, and 201 layers respectively. The proposed model has only 13.7 million parameters, which is 11 times lower than the VGG19 model. Since the proposed CNN model has fewer parameters and layers, the feature extractor CNN model is lightweight compared to the other TL models. It is easy to use an end-to-end system as

it does not require a separate modules for individual tasks such as feature extraction, feature selection etc. However, it has some limitations for instance an end-to-end cannot provide good accuracy with small amount of training data. Fig. 3 shows the architecture of the CNN model used in this study.

The size of the input CXR images was 124 × 124. The CNN model was comprised of three convolutional layers (CL), where each CL block consisted of a batch-normalization, an activation, and a 2 × 2 max-pooling layer. Batch normalization was utilized since it speeded up and stabilized the model by re-centering and rescaling the layers' inputs, as well as reducing overfitting. In the first CL layer, the kernel size was 5 × 5 and the padding were the'same' to check the border elements, which might contain important information. Additionally, the kernel size was kept relatively large in case any significant lesions were detected in the CXR images that needed to be collected readily. Following the initial pooling layer with an input dimension of 62 × 62, the kernel size is kept relatively small. The kernel size of the remaining CLs was 3 × 3 and used the'valid' padding. Basically, a custom CNN-based TL model was developed in this work. Most of the TL models used large kernel size at the beginning and then the kernel size was reduced in the proceeding layers [23] due to an initial large size of the CXR images and subsequent smaller sizes after each convolution layer.

To avoid gradient fading, a ReLU activation function was utilized after each batch normalization layer [54]. Following the final CL, two fully connected (FC) layers were employed, which were determined by a trial and error procedure. The first FC layer was 1,024 nodes dense, and the feature extraction was carried out from the final FC layer, which was 512 nodes deep. To reduce the complexity and overfitting, two dropout layers were included and as a result, the training speed increased significantly [55]. One dropout with a 0.3 probability was positioned after the last CL layer and another one with a 0.5 probability before the feature extraction layer. Since the amount of training data was large, the Adam optimizer was employed as it could provide improved performance for a large training dataset [56]. In order to extract the distinctive features, the model was run with the following parameters: 50 epochs; 32 batch size and 0.001 learning rate. Finally, a total of 512 important features were extracted from the last FC layer. Table 3 shows a summary of the CNN model employed. To begin, a large number of features were retrieved in order to minimize underfitting. Additionally, PCC was employed to eliminate redundant features from the derived ones in order to account for the overfitting and to improve classification accuracy.

2.3. *Feature selection*

In the current era of ML, data is very important. Every piece of information which is measurable and involved in recognizing a phenomenon or a circumstance is called a feature. To recognize each phenomenon individually, there might be a large number of features available in the real world. However, all the features are not highly related to the outcome [57]. Huge feature space makes the learning process of a ML algorithm slow and increases computational complexity. Hence, it is required to find out the optimal feature subset to resolve these issues through feature selection. Among many available FS strategies, PCC based FS technique could help in identifying distinctive features. Best feature subspace was selected from the feature extracted using the CNN for reducing complexity and increasing the performance and processing speed [57]. Correlation coefficients (CC) of all the features were determined and the first one of each pair containing a CC value above the threshold were eliminated. Finally, the standard scaler has been used to standardize the extracted features performed by subtracting the mean and scaling to mean–variance. Standardization features could help in achieving better classification performance [59,60]. The equation for the standard score for sample x [60]

$$y = \frac{x - \bar{x}}{\sigma} \tag{1}$$

Where $\bar{x}$ is the mean of the samples and σ is standard deviation of the samples.

Algorithm 1: Feature Selection using Pearson Correlation Coefficient

```
CorrMatrix = Features.cor ()
for i in range(len(CorrMatrix.columns)):
        for j in range(i):
        If abs (CorrMatrix.iloc[i, j]) > threshold:
        ColName = CorrMatrix.columns[i]
        ColCorr.add(ColName)
Features.drop(ColCorr)
```

2.4. *Extreme learning machine*

Hunag proposed ELM [61], which is a neural network (NN) based on a forward feed network. A single hidden layer has been used to classify the extracted features. Normally, ELM has shown optimistic results in case of multiclass classification and since there is no backpropagation needed in ELM, hence it also requires less time compared to the traditional NN or DL models. The training time was a thousand times faster than the traditional NN and achieved better generalization power due to the absence of backpropagation and higher classification performance [61–64]. The parameters between the input layer to the hidden layer were derived arbitrarily whereas the hidden layer to output layer parameters were derived using pseudoinverse. Fig. 4 demonstrates the basic architecture of ELM model used in this study.

The number of hidden nodes and the parameters for this study are shown in Table 4, where the hidden layer nodes were selected through trial-and-error method. The model would get overfitted if more hidden nodes were added, while if fewer hidden nodes were utilized, then the model would be under fitted. Consequently, in this study, adequate hidden nodes had employed, which produced promising classification results that resolved both the overfitting and underfitting issues. For the eight-class classification, the total trainable parameters were 13,397,504 for extracting the features using CNN and for classification, whereas the total number of parameters of ELM were 304,500, hence the total parameters were summed up to 13,702,004.

Algorithm 2: Extreme learning machine

$$X_{(n,m)} = \begin{bmatrix} x_{(1,1)} & x_{(1,2)} & \cdots & x_{(1,m)} \\ x_{(2,1)} & x_{(2,2)} & \cdots & x_{(1,m)} \\ x_{(3,1)} & x_{(3,2)} & \cdots & x_{(1,m)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{(n,1)} & x_{(n,2)} & \cdots & x_{(n,m)} \end{bmatrix} Y_{(n,t)} = \begin{bmatrix} y_{(1,1)} & y_{(1,2)} & \cdots & y_{(1,t)} \\ y_{(2,1)} & y_{(2,2)} & \cdots & y_{(1,t)} \\ y_{(3,1)} & y_{(3,2)} & \cdots & y_{(1,t)} \\ \vdots & \vdots & \ddots & \vdots \\ y_{(n,1)} & y_{(n,2)} & \cdots & y_{(n,t)} \end{bmatrix}$$

Here, X and Y be the feature and target matrix.

1. Randomly generates the input weight $W_{(m,N)}$ and bias $B_{(1,N)}$ matrix.$W(m,N) = \begin{bmatrix} w_{(1,1)} & w_{(1,2)} & \cdots & w_{(1,N)} \\ w_{(2,1)} & w_{(2,2)} & \cdots & w_{(1,N)} \\ w_{(3,1)} & w_{(3,2)} & \cdots & w_{(1,N)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{(m,1)} & w_{(m,2)} & \cdots & w_{(m,N)} \end{bmatrix} B(1,N) = \begin{bmatrix} b_{(1,1)} & b_{(1,2)} & \cdots & b_{(1,N)} \end{bmatrix}$
2. Determine the output $H_{(n,N)}$ of the hidden layer.$H_{(n,N)} = G(X_{(n,m)} \cdot W_{(m,N)} + B_{(1,N)})H(n,N) = \begin{bmatrix} h_{(1,1)} & h_{(1,2)} & \cdots & h_{(1,N)} \\ h_{(2,1)} & h_{(2,2)} & \cdots & h_{(1,N)} \\ h_{(3,1)} & h_{(3,2)} & \cdots & h_{(1,N)} \\ \vdots & \vdots & \ddots & \vdots \\ h_{(n,1)} & h_{(n,2)} & \cdots & h_{(n,N)} \end{bmatrix}$ Here, G is the activation function.
3. Determine the output weight matrix $\beta_{(N,t)}\beta_{(N,t)} = H^{\dagger}_{(N,n)} \cdot T_{(n,t)}$
4. Make prediction using $\beta_{(N,t)}$

3. Results and analysis

3.1. Evaluation matrices

To assess the performance of the proposed ML models, a number of metrics such as accuracy, precision, recall, f1-score, and AUC were considered. The metrics can be defined by Equation (2) to Equation (6) [65,66]:

$$Accuracy = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (2)$$

$$Precision = \frac{T_P}{T_P + F_P} \quad (3)$$

$$Recall = \frac{T_P}{T_N + F_P} \quad (4)$$

$$F1 - Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \quad (5)$$

$$AUC = \frac{1}{2}\left(\frac{T_P}{T_P + F_N} + \frac{T_N}{T_N + F_P}\right) \quad (6)$$

Where T_P, T_N, F_P and F_N, denote true positives, true negatives, false positives, and false negatives, respectively.

3.2. Environmental setup

Python programming language was used to write the codes, which were run on PyCharm Community Edition (2021.2.3) software. The CNN model was implemented using Keras with TensorFlow as the backend for the feature extraction. The ELM models were trained and tested on a PC with a 11th generation Intel(R) Core (TM) i9-11900 CPU @2.50 GHz, 32 GB RAM, and an NVIDIA GeForce, RTX 3090 24 GB GPU, running on a 64-bit Windows 10 Pro operating system. The code is available at https://github.com/NahiduzzamanRuet/ChestX-Ray8/blob/main/ChestX-Ray8.py.

In this study, all the schemes were performed using two models. In the first model, after the pre-processing, a light-weight CNN was applied to extract the 512 features, which were standardized. After feature standardization, the ELM was used to classify different lung diseases, and the model was named as CNN-ELM.

In the second model, after the extraction of 512 features and performing standardization on the features, PCC was employed on the extracted features to eliminate the redundant features and to select the most discriminant ones. Finally, classification was carried out by the ELM, and the model was termed as CNN-PCC-ELM. Table 5 shows the PCC values and the

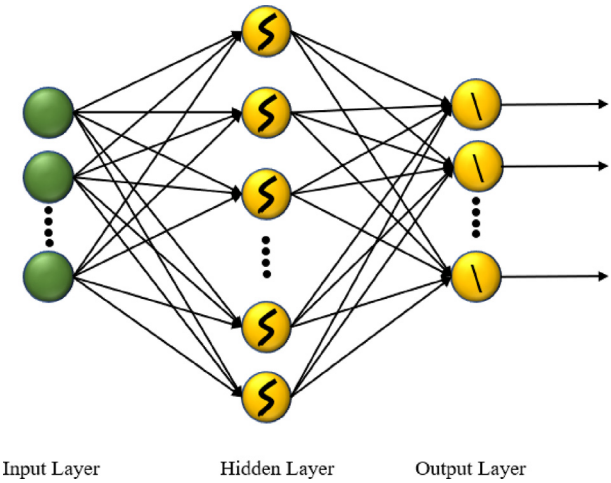


Fig. 4 – The architecture of Extreme Machine Learning.

Table 4 – Parameters of ELM for Different Schemes.

Scheme	*Total Nodes in Input Layer*	*Total Nodes in Hidden Layer*	*Total Nodes in Output Layer*	*Total No of Parameters*
Scheme 1	195	1,500	8	304,500
Scheme 2	60	700	4	44,800
Scheme 3	72	400	3	30,000
Scheme 4	59	300	3	18,600
Scheme 5	45	300	1	13,800
Scheme 6	29	400	1	12,000
Activation Function	ReLU			

Table 5 – PCC value and extracted features.

Scheme No	*PCC Value*	*Features of CNN-ELM*	*Features of CNN-PCC-ELM*
Scheme 1	0.82	512	195
Scheme 2	0.80	512	60
Scheme 3	0.80	512	72
Scheme 4	0.80	512	59
Scheme 5	0.80	512	45
Scheme 6	0.80	512	29

extracted features for different schemes. The PCC threshold value was selected using a trial-and-error method. For Scheme 1 (8 classes), which was a large class, a threshold value of 0.82 produced the most promising results than other PCC values either higher or lower. For instance, for large classes (8-classes) like the Scheme 1, with a higher value than the threshold value would increase the number of unnecessary redundant features causing less accurate results. Similarly, a lower PCC value than the threshold value would reduce the number of discriminant features leading to poor quality results again. In the cases of smaller classes (4, 3, and 2 classes), a slightly lower PCC value provided the best prominent features for effective and accurate detection of the diseases.

3.3. *Results for Scheme 1*

For Scheme 1, a total of 20,136 data were employed to train the CNN, CNN-ELM and CNN-PCC-ELM models, and 3,554 data were employed to assess the proposed framework's performance. A confusion matrix (CM) was used to calculate the precision, recall, and accuracy of the CNN, CNN-ELM models as shown in Fig. 5A.

The average precision, recall, and accuracy of the CNN model were 0.75, 0.73 and 85%, respectively. However, the average precision, recall, and accuracy of the CNN-ELM model were recorded as 0.95, 0.88, and 93.19%, respectively (Table 6). The CNN-ELM model used 512 features that were too high. Therefore, the redundant and irrelevant features were removed for improving the classification performance.

After extracting the 512 features using CNN, a PCC was used to eliminate a total of 317 redundant features leaving behind a total of 195 most prominent features. Finally, the ELM model was used for the classification, and the CNN-PCC-ELM model performance was calculated using a CM as shown in Fig. 5B. It was noticed that the proposed framework demonstrated a significantly increased accuracy of 96.22%, which was approximately 11% higher than the CNN and 3% higher than the CNN-ELM model. It shows that an end-to-

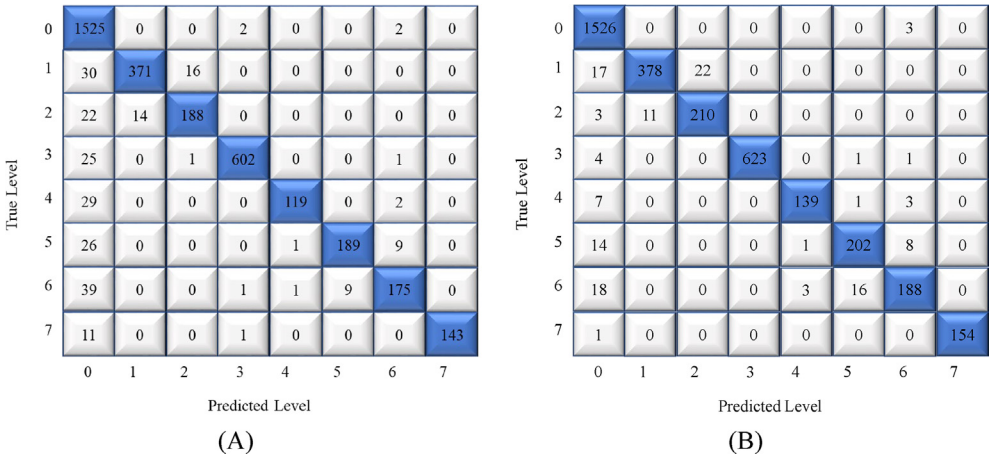


Fig. 5 – CMs for scheme 1: (A) CNN-ELM and (B) CNN-PCC-ELM.

Table 6 – Classification performance results of Scheme 1.

Type of lung diseases	Precision			Recall			F1-score			Accuracy (%)		
	CNN	CNN-ELM	CNN-PCC-ELM	CNN	CNN-ELM	CNN-PCC-ELM	CNN	CNN-ELM	CNN-PCC-ELM	CNN	CNN-ELM	CNN-PCC-ELM
Normal	0.96	0.89	0.96	0.99	1.00	1.00	0.97	0.94	0.98	–	–	–
Bacterial Pneumonia	0.87	0.96	0.97	0.79	0.89	0.91	0.82	0.93	0.94	–	–	–
Viral Pneumonia	0.66	0.92	0.91	0.73	0.84	0.94	0.70	0.88	0.92	–	–	–
COVID-19	0.91	0.99	1.00	0.93	0.96	0.99	0.92	0.97	1.00	–	–	–
Cardiomegaly	0.64	0.98	0.97	0.45	0.79	0.93	0.53	0.88	0.95	–	–	–
Pleural	0.55	0.95	0.92	0.41	0.84	0.90	0.47	0.89	0.91	–	–	–
Lung Opacity	0.46	0.93	0.93	0.61	0.78	0.84	0.52	0.85	0.88	–	–	–
Tuberculosis	0.99	1.00	1.00	0.92	0.92	0.99	0.95	0.96	1.00	–	–	–
Average	0.75	0.95	0.96	0.73	0.88	0.94	0.74	0.91	0.95	85	93.19	96.22

end model performs poorly with relatively smaller dataset. The recall of CNN-PCC-ELM also increased by 6% compared to the CNN-ELM, model demonstrating the robustness in the model's classification performance. In biomedical engineering, the recall must be high enough so that the patient with lung diseases must be correctly detected. Since the redundant and unnecessary features confused the model and reduced the classification performance, after removing the duplicate and irrelevant features, the model's performance increased with reduced complexity. Additionally, despite imbalance in the dataset, each class contributed equally (precision and recall almost higher than 91% and 84% respectively for each class) to the final outcome. The class wise ROC values for both the models presented in Fig. 6 demonstrated that the models' discriminant capabilities for each class indicating consistency of the models even with an unbalanced dataset. The calculated average AUCs of the CNN-ELM and CNN-PCC-ELM were recorded as 99.10% and 99.48% respectively clearly indicating the superiority of the latter.

To demonstrate the superior performance of the proposed framework in the case of the multiclass environment, the same dataset was used for training and testing the five TL models: VGG19, ResNet50, ResNet152V2, InceptionResNetV2, and DenseNet201. The ROCs of the TL models are presented in Fig. 7.

The AUCs of VGG19, ResNet50, ResNet152V2, InceptionResNetV2, and DenseNet201 were 98.61%, 93.08%, 98.42%, 98.30%, and 99.07%, respectively. Table 7 shows average precision, recall, and f1-score of the TL models. The accuracies of VGG19, ResNet50, ResNet152V2, InceptionResNetV2, and DenseNet201 were 74.20%, 64.02%, 73.05%, 70.45%, and 74.30%, respectively, whereas the accuracy of CNN-PCC-ELM was 96.22%, which was almost 20% higher than the TL models. The CNN-PCC-ELM model achieved a promising result because the CNN was used to extract the discriminant features. PCC was applied to find the irrelevant and redundant features that contained insignificant information, which could reduce the model's performance. Hence, the ELM model showed a favorable result with CNN-PCC for feature extraction and selection.

The CNN has three convolutional layers, and ELM has three layers; hence there are only six layers. On the contrary, the TL models VGG19, ResNet50, ResNet152V2, InceptionResNetV2, and DenseNet201 have 19, 50, 152, 164, and 201 layers respectively. The proposed model has only 13.7 million parameters, which is 11 times lower than the VGG19 model. Since the proposed CNN model has fewer parameters and layers, the features extractor CNN model is lightweight compared to the other TL models. From Table 8, it should be noted that it took 406 s to extract the 512 features using the CNN. Only 39 s were taken to remove the features using PCC and train the ELM. Therefore, the training time of the CNN-PCC-ELM model was only 445 s, whereas some of the TL models' training times were almost four times greater than the proposed model. The proposed framework took only 0.0156 s to test, whereas the InceptionResNetV2 TL model took 7.3505 s, which was relatively higher. The accuracy of the proposed framework (96.22%) was significantly higher than the accuracy of the TL models.

Fig. 8 shows a comparison of classification performance of the CNN-PCC-ELM model with the TL models. The recall and accuracy of the proposed framework were 0.94 and 96.22%,

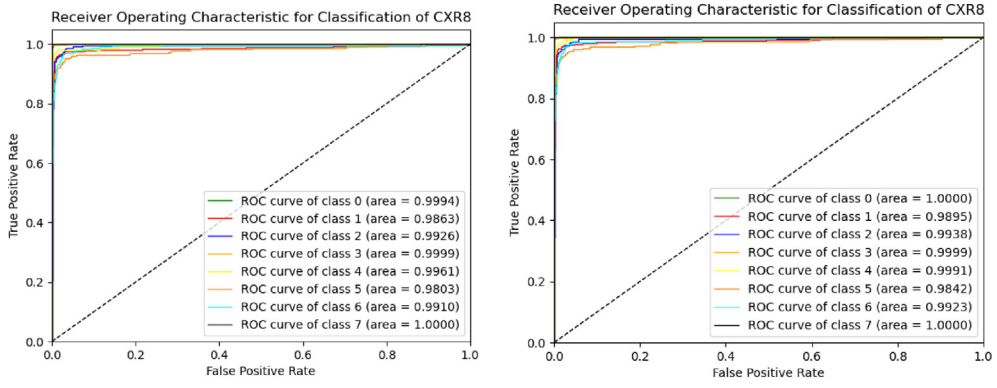


Fig. 6 – ROC curve for Scheme 1: (A) CNN-ELM and (B) CNN-PCC-ELM.

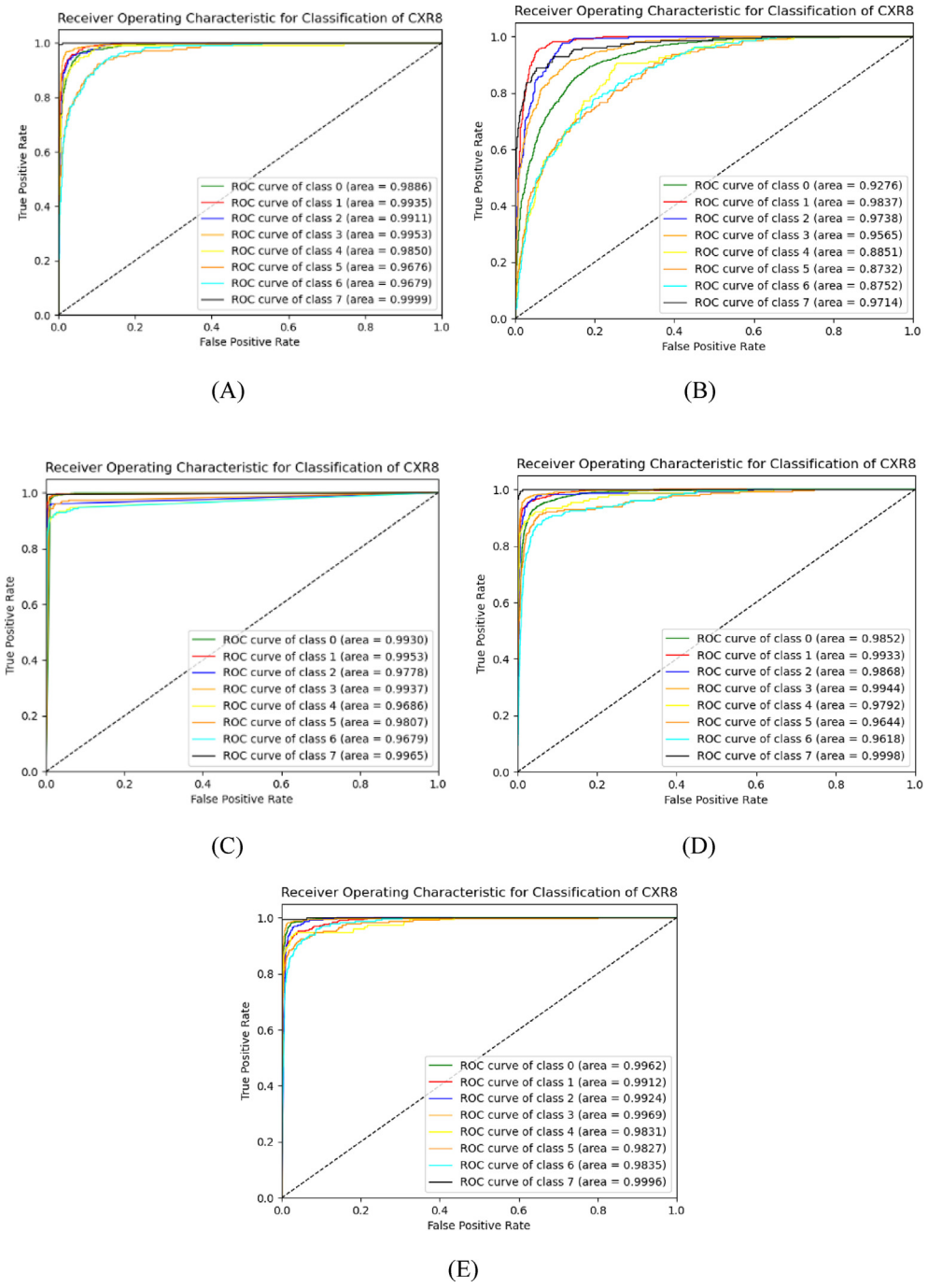


Fig. 7 – ROC curves for TL models in Scheme 1: (A) VGG19, (B) ResNet50, (C) ResNet152V2, (D) InceptionResNetV2 and (E) DenseNet201.

which were greater than that of the other TL models. From Fig. 7, it was also noticed that the ROC of CNN-PCC-ELM (99.48%) was again better than ROCs of the VGG19, ResNet50, ResNet152V2, InceptionResNetV2, and DenseNet201 models (98.61%, 93.08%, 98.42%, 98.30%, and 99.07% respectively). Since the ROC curve shows how well a classification model performs across all categorization levels, it can be concluded that the proposed framework shows its robustness in every performance criterion.

From the above results, it was clear that the proposed CNN-PCC-ELM model achieved a higher classification performance with reduced processing time, parameters, and layers compared to the TL models. Therefore, it is safe to conclude that in the case of multiclass classification, the proposed CNN-PCC-ELM is superior to the other TL models based on all types of performance criteria.

3.4. *Results for Scheme 2*

In this scheme, COVID-19 was detected from normal, bacterial, and viral pneumonia with a four-class classification. The CNN-ELM was trained using 14,923 data with 512 fea-

Table 7 – Results of Scheme 1 for VGG19, ResNet50, ResNet152V2, InceptionResNetV2, and DenseNet201.

TL models	*Precision*	*Recall*	*F1-score*	*Accuracy (%)*
VGG19	0.72	0.60	0.63	76.53
ResNet50	0.61	0.50	0.51	68.94
ResNet152V2	0.67	0.63	0.64	77.52
InceptionResNetV2	0.58	0.60	0.58	69.36
DenseNet201	0.68	0.67	0.65	77.12

Table 8 – Comparing performance of CNN-PCA-ELM with VGG19, ResNet50, ResNet152V2, InceptionResNetV2, and DenseNet201.

Models	*Accuracy (%)*	*No. of Layers*	*Parameters (million)*	*Training Time (sec)*	*Testing Time (sec)*
VGG19	76.53	19	143.6	563	1.9710
ResNet50	68.94	50	25.6	626	2.4319
ResNet152V2	77.52	152	60.3	1512	5.8072
InceptionResNetV2	69.36	164	55.8	1888	7.3505
DenseNet201	77.12	201	20.2	1491	4.4138
CNN-PCC-ELM	**96.22**	**3 + 3**	**13.7**	**445**	**0.0156**

tures. To calculate the model's classification performance 3,731 data was used where the number of normal, COVID-19, bacterial, and viral pneumonia data were 2,039, 838, 555, and 299, respectively. The classification result was calculated from the CMs shown in Fig. 9.

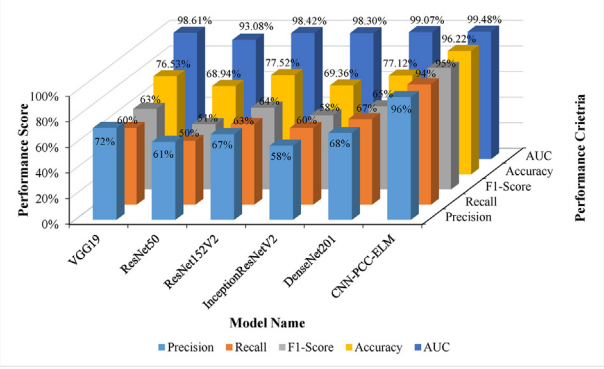


Fig. 8 – Classification performance comparison between CNN-PCC-ELM and the TL models for Scheme 1.

The recall and accuracy of the CNN-ELM were 0.90 and 94.67% as shown in Table 9. After removing the redundant features, the CNN-PCC-ELM model was trained using the same data but with only 60 features. The CNN-PCC-ELM achieved an accuracy of 96.57%, which was approximately 2% higher than the accuracy of the CNN-ELM model.

The AUCs of the CNN-PCC-ELM (99.57%) was slightly better than that of the CNN-ELM (99.41%) and Fig. 10 demonstrates the class-wise ROCs of both the models.

3.5. *Results for Scheme 3*

In this scheme, multivariate pneumonia was detected by training the CNN-ELM models with 11,569 data having 512 features. A total of 2,893 data was used for testing the model and calculating the classification performance. In the second stage, the CNN-PCC-ELM model was trained using the same

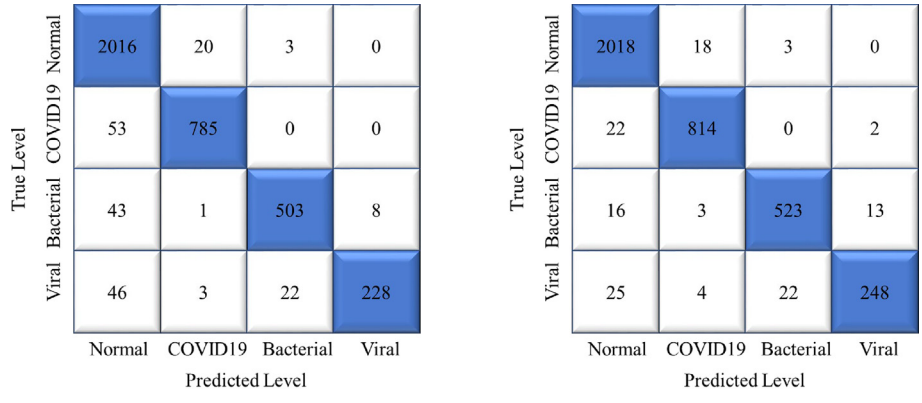


Fig. 9 – CMs for Scheme 2: (A) using CNN-ELM and (B) using CNN-PCC-ELM.

Table 9 – Classification performance results of Scheme 2.

Type of lung diseases	*Precision*		*Recall*		*F1-score*		*Accuracy (%)*	
	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM
Normal	0.93	0.97	0.99	0.99	0.96	0.98	–	–
COVID-19	0.97	0.97	0.94	0.97	0.95	0.97	–	–
Bacterial Pneumonia	0.95	0.95	0.91	0.94	0.92	0.95	–	–
Viral Pneumonia	0.97	0.94	0.76	0.83	0.85	0.88	–	–
Average	0.96	0.96	0.90	0.97	0.92	0.97	94.67	96.57

data having 72 features, which was also used for testing the model. The accuracy of the CNN-PCC-ELM model (97.33%) was almost 3% higher than that of the CNN-ELM model (94.61%) as demonstrated in Table 10. The average precision, recall, and F1-score of the CNN-PCC-ELM models were also found to be better than that of the CNN-ELM model.

Fig. 11 shows the CMs for both the CNN-ELM and CNN-PCC-ELM models.

The AUC value of the CNN-PCC-ELM model (99.33%) was also superior to that of the CNN-ELM model (98.41%) as demonstrated by the class wise ROC curves presented in Fig. 12.

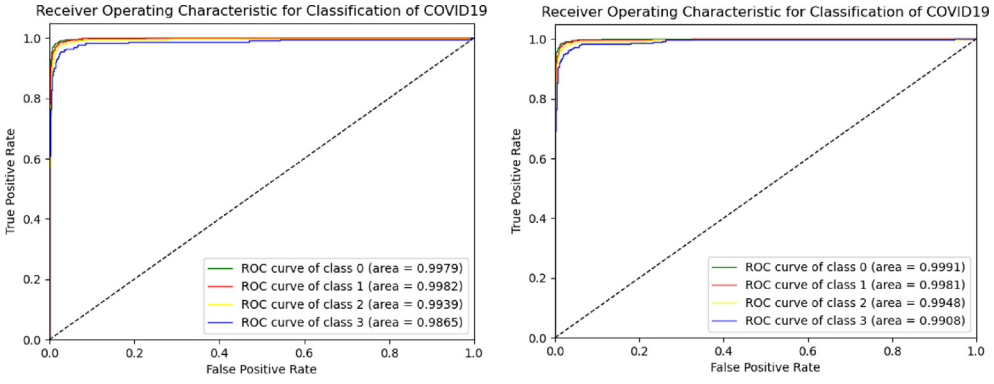
3.6. Results for Scheme 4

In this study, the main focus was to detect COVID-19 from different other lung diseases. Scheme 4 was designed to detect COVID-19 from pneumonia with a three-class classification. A total of 12,701 data with 512 features was used for training the CNN-ELM model and for assessing the performance of the model, 3,176 data was used. In the previous studies, a small number of COVID-19 CXR images were used, whereas 4,192 CXR images were employed in this study. The CNN-ELM model achieved an accuracy of 97.42% and a precision of 98% (Table 11). The CNN-PCC-ELM model achieved an optimistic accuracy of 99.55%, with a precision, recall and f1-score of 99% while testing the model using the same testing data with only 59 features. The performance criterion for Scheme 4 was calculated from the CMs shown in Fig. 13.

The AUCs of the CNN-ELM and CNN-PCC-ELM were 99.96% and 99.97%, respectively (Fig. 14).

3.7. Results for Scheme 5

In this scheme, COVID-19 was detected from the normal patient as a binary classification. The CNN-ELM model was trained using 11,507 data with 512 features and performance was measured by testing the model using 2,877 data. The accuracy and recall of the CNN-ELM model were 96.66% and 97%, respectively as shown in Table 12. In the second stage, PCC has been used for eliminating 467 features and the proposed CNN-PCC-ELM model achieved an accuracy of 98.82% while testing with 45 features.



(A) (B)

Fig. 10 – ROC curves for Scheme 2: (A) CNN-ELM and (B) CNN-PCC-ELM.

Table 10 – Classification performance results of Scheme 3.

Type of lung diseases	*Precision*		*Recall*		*F1-score*		*Accuracy (%)*	
	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM
Normal	0.95	0.99	1.00	0.99	0.97	0.99	–	–
Bacterial Pneumonia	0.92	0.93	0.91	0.95	0.91	0.94	–	–
Viral Pneumonia	0.96	0.92	0.68	0.88	0.80	0.90	–	–
Average	0.94	0.95	0.86	0.94	0.90	0.94	94.61	97.34

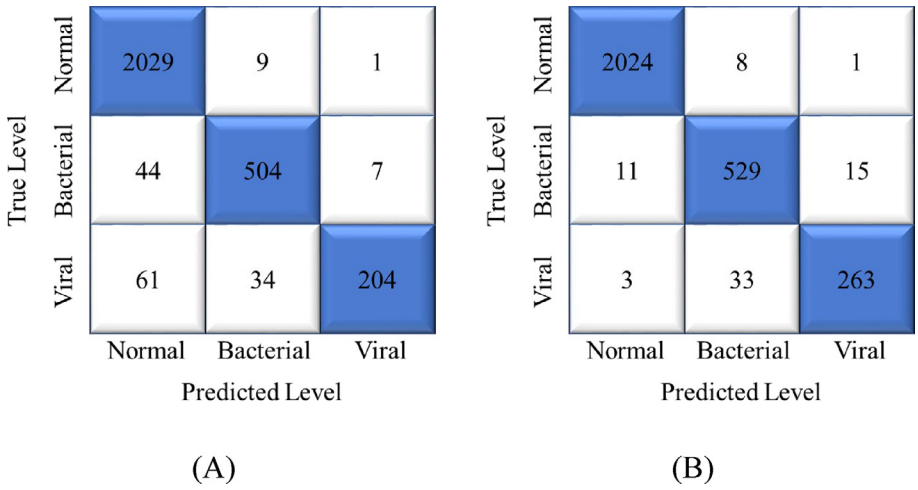


Fig. 11 – CMs for Scheme 3: (A) CNN-ELM and (B) CNN-PCC-ELM.

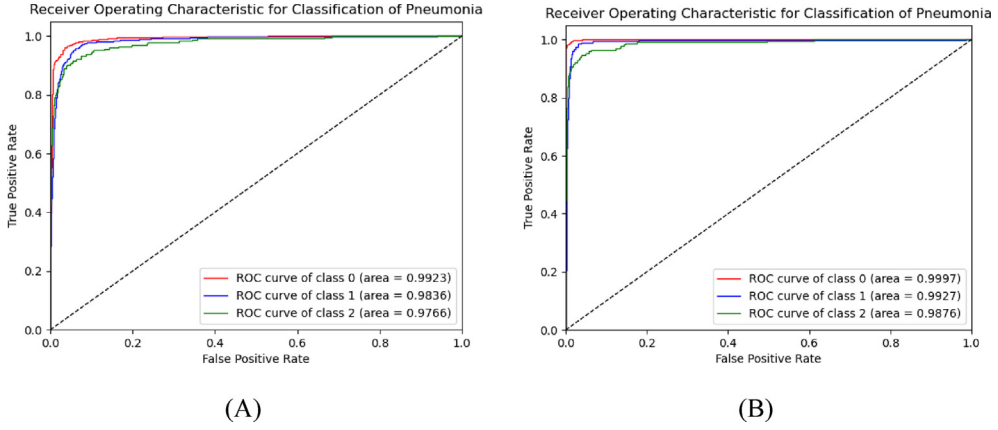


Fig. 12 – ROC curves for Scheme 3: (A) CNN-ELM and (B) CNN-PCC-ELM.

The performance criterion for Scheme 5 was calculated from the CMs shown in Fig. 15.

The AUC values of the CNN-ELM and CNN-PCC-ELM models were 99.81% and 99.88%, respectively and again this showed superior performance of the CNN-PCC-ELM model (Fig. 16).

3.8. *Results for Scheme 6*

In this final scheme, the tuberculosis disease was detected from the normal patient. In this case, 8,982 data with 512 features was used to train the CNN-ELM model and achieved an accuracy of 98.13% and a precision of 99% (Table 13 and Fig. 17) and an AUC of 99.95% (Fig. 18A).

After removing 483 irrelevant features, the CNN-PCC-ELM model was tested using 2,246 data with only 29 features and achieved an accuracy of 99.51% and an AUC of 100% which are shown in Table 13 and Fig. 18(B).

3.9. *Comparison between the six schemes*

From Fig. 19, it was identified that in every scheme, the CNN-PCC-ELM performed better than the CNN-ELM model. The

Table 11 – Classification performance results of Scheme 4.

Type of lung diseases	*Precision*		*Recall*		*F1-score*		*Accuracy (%)*	
	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM
Normal	0.97	1.00	1.00	1.00	0.98	1.00	–	–
Pneumonia	1.00	0.99	0.95	0.99	0.98	0.99	–	–
COVID-19	0.99	0.99	0.93	1.00	0.96	0.99	–	–
Average	0.98	0.99	0.96	0.99	0.97	0.99	97.42	99.55

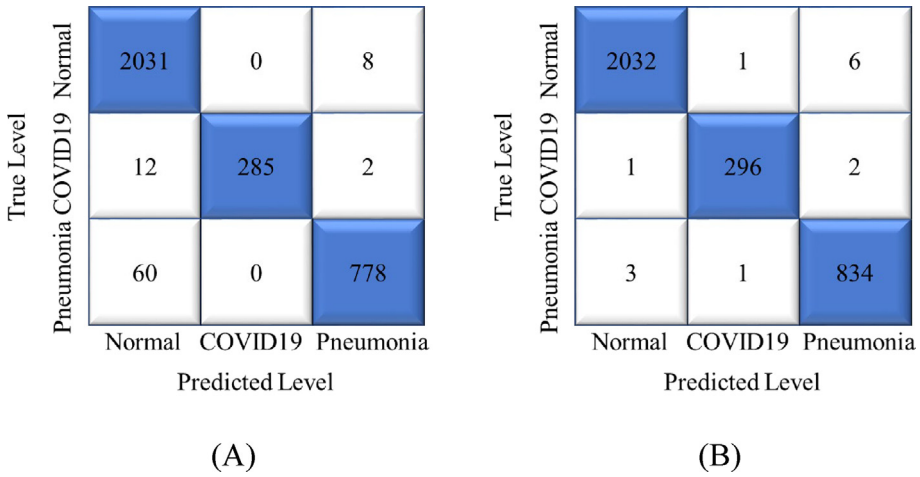


Fig. 13 – CMs for Scheme 4: (A) CNN-ELM and (B) CNN-PCC-ELM.

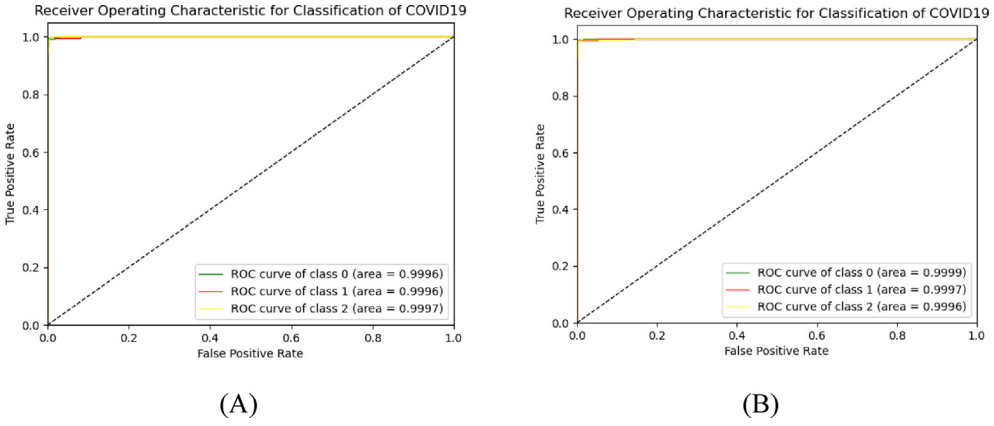


Fig. 14 – ROC curves for Scheme 4: (A) CNN-ELM and (B) CNN-PCC-ELM.

Table 12 – Classification performance results of Scheme 5.

Type of lung diseases	*Precision*		*Recall*		*F1-score*		*Accuracy (%)*	
	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM
Normal	0.99	0.99	0.98	0.99	0.98	0.99	–	–
COVID-19	0.97	0.98	0.96	0.98	0.96	0.98	–	–
Average	0.99	0.99	0.97	0.99	0.98	0.99	96.66	98.82

CNN-ELM model used 512 features while the CNN-PCC-ELM model used 195, 60, 72, 59, 45, and 29 features for Schemes 1, 2, 3, 4, 5, and 6, respectively, which were selected by eliminating the redundant features resulting in better classification performance. Hence, the results demonstrated that PCC removed the duplicate and irrelevant features efficiently and this was reflected in the improved classification performance with less complexity of the CNN-PCC-ELM model.

4. Performance comparison and discussions

4.1. *Comparison of CNN-PCC-ELM with ResNet50-PCC-ELM and VGG19-PCC-ELM*

In this section, a 5-fold CV was applied to the CNN-PCC-ELM (CPE) model to determine the effect on the outcomes. The classification performance of each fold was compared with

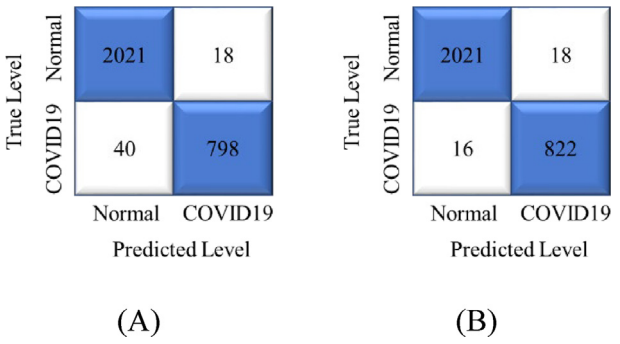


Fig. 15 – CMs for Scheme 5: (A) CNN-ELM and (B) CNN-PCC-ELM.

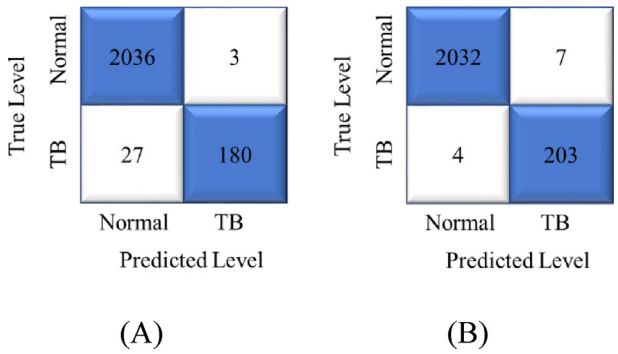


Fig. 17 – CMs for Scheme 6 (A) CNN-ELM and (B) using CNN-PCC-ELM.

ResNet50-PCC-ELM (RPE) and VGG19-PCC-ELM (VPE) as shown in Table 14.

The highest accuracy of 75.69% and AUC of 91.9% were obtained from Fold 5, when employing RPE and the highest accuracy of 89.91% and AUC of 97.9% were obtained from the same fold utilizing VPE. On the other hand, the proposed CPE attained an optimistic accuracy of 98.12% (almost 9% higher than the VPE model) and AUC of 99.7% (almost 2% higher than the VPE) from the same fold. Following PCC, the RPE and VPE contained just nine and three redundant characteristics, respectively. Thus, the RPE and VPE models required 503 and 509 features, respectively, whereas the CPE model required only 195 features. As a result, the CPE model (445 s) needed shorter processing time than the VPE (609 s) and RPE (686 s) models. Since pre-trained ResNet 50 and VGG 19 were not trained for specific CXR images hence their features extraction capabilities for the CXR images were limited. Whereas the CXR images were specifically employed to train the proposed CNN model that resulted in extracting discriminant features leading to an improved the performance. As shown in Fig. 20, the suggested CPE model was adaptable and flexible in terms of classification performance and processing time.

4.2. *Comparison with SOTA models*

This section compared different schemes with the existing SOTA methods as shown in Table 15. Scheme 1 was a new merged dataset; hence, it could not be compared with other SOTA methods. CNN-PCC-ELM model performed better with this new merged dataset compared with different TL models as evidenced in Section 3.3 and Section 4.1. For scheme 2, the best performances were achieved by Rashid et al. in classifying four types of diseases (normal-COVID-19-bacterial-viral pneumonia) with an accuracy, recall, and precision of 90.13%, 90.13%, and 90.13%, respectively [86]. On the contrary, the proposed framework achieved an accuracy of 96.57%, almost 6% higher than their model, and a recall of almost 7% (97%) higher, demonstrating the model's robustness for scheme 2.

For Scheme 3, there was not much work performed to classify bacterial pneumonia from the viral pneumonia. The authors [25] published a study to detect the multivariate pneumonia where the 99.01 % AUC was achieved from the model and this value was greater than that of prior studies. However, the accuracy (97.34%) in this work was slightly lower than the most recent work and the previous model (98.32%) [25]. The discriminating ability of the proposed model (AUC of 99.33%) was remarkable compared to the prior works demonstrating its robustness in the detection of lung diseases. The main contribution could be differentiated in a way that in contrast to two types of pneumonia in the previous work, a maximum seven types of lung diseases including COVID-19 were considered with a large amount of data in the datasets in this work.

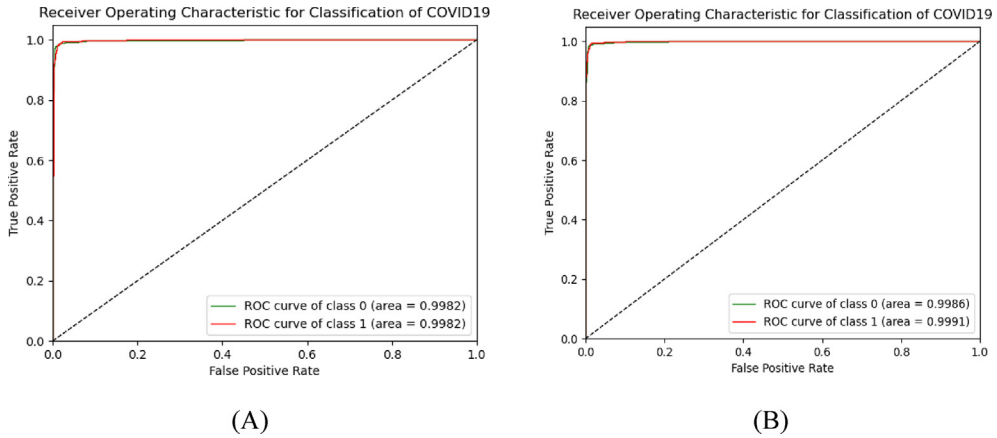


Fig. 16 – ROC curves for Scheme 5: (A) CNN-ELM and (B) CNN-PCC-ELM.

Table 13 – Performance Classification results of Scheme 6.

Type of lung diseases	*Precision*		*Recall*		*F1-score*		*Accuracy (%)*	
	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM	CNN-ELM	CNN-PCC-ELM
Normal	0.99	1.00	0.99	1.00	0.99	1.00	–	–
Tuberculosis	0.97	0.97	0.90	0.98	0.94	0.97	–	–
Average	0.99	1.00	0.98	1.00	0.98	1.00	98.13	99.51

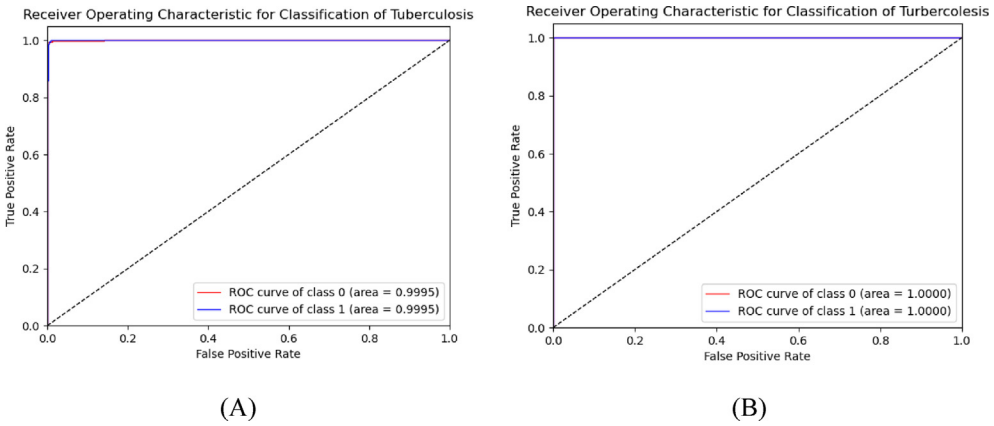


Fig. 18 – ROC curves for Scheme 6: (A) CNN-ELM and (B) CNN-PCC-ELM.

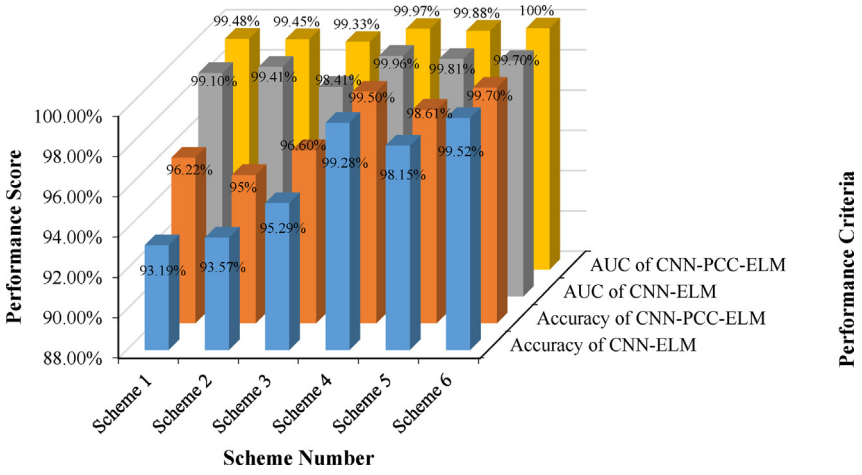


Fig. 19 – Performance comparison graph between CNN-ELM and CNN-PCC-ELM models for different schemes.

For Scheme 4, several studies were carried out to detect COVID-19 from pneumonia patients. Only the latest works were used for comparison in the table. The proposed model achieved a 100% precision and AUC of 99.97%, surpassing the seventeen SOTA methods for the three-class classification (normal-pneumonia-COVID-19). Besides this, Patro et al. developed a custom SCovNet based on CNN, which achieved the highest accuracy of 97.99% among the other SOTA models. In contrast, the proposed CNN-PCC-ELM achieved a satisfactory accuracy of 99.55%, almost 2% higher than their model [72]. Azam et al. had the best precision and recall results out of all the SOTA models, coming in at 99.02% and 98.26%, respectively [77]. The proposed CNN-PCC-ELM outweighs their precision (100%) and recall (99%).

Scheme 5 has been one of the most important focuses for researchers in the last three years. For the classification of COVID-19 from the regular patients with the help of CXR images, Joshi et al. achieved an optimistic accuracy of 99.81% but the recall was 98.45% which was slightly lower than the proposed model [82]. The proposed model performed a promising result with an optimistic AUC and recall of 99.89% and 99%, respectively which outperformed the existing nineteen SOTA models. Tuberculosis, one of the most dangerous lung-related diseases, was also detected accurately in

Table 14 – Performance comparison of CNN-PCC-ELM with ResNet50-PCC-ELM and VGG19-PCC-ELM.

	Precision			Recall			Accuracy			AUC		
	RPE	VPE	CPE	RPE	VPE	CPE	RPE	VPE	CPE	RPE	VPE	CPE
Fold1	0.66	0.87	**0.96**	0.57	0.80	**0.96**	75.26	89.19	**97.78**	0.909	0.980	**0.996**
Fold2	0.65	0.88	**0.96**	0.57	0.81	**0.95**	74.91	89.82	**97.36**	0.915	0.979	**0.995**
Fold3	0.65	0.87	**0.96**	0.57	0.80	**0.96**	74.55	89.80	**97.83**	0.911	0.981	**0.996**
Fold4	0.65	0.86	**0.96**	0.57	0.80	**0.96**	74.84	88.62	**97.72**	0.911	0.981	**0.997**
Fold5	0.66	0.87	**0.97**	0.58	0.81	**0.96**	75.69	89.91	**98.12**	0.919	0.979	**0.997**

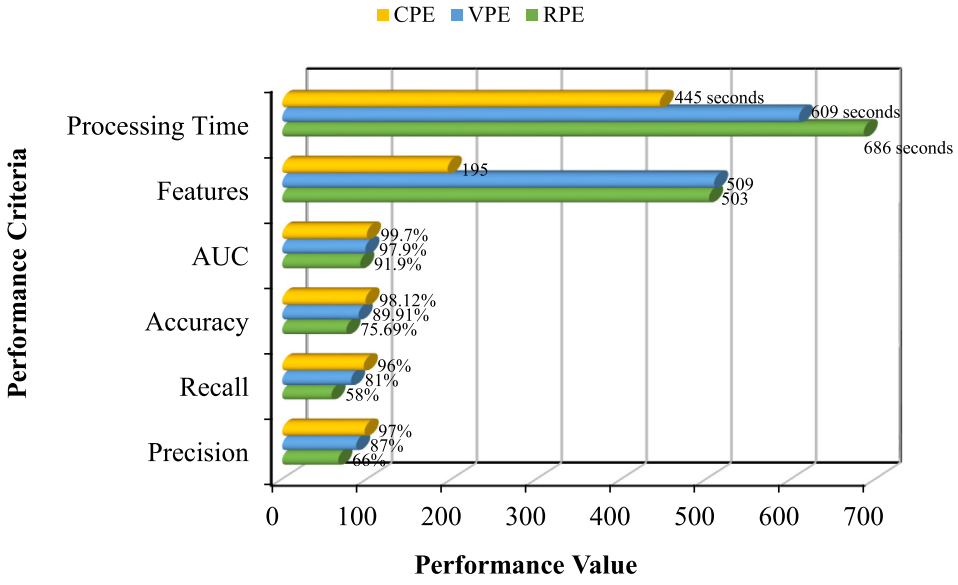


Fig. 20 – Classification performance comparison between CNN-PCC-ELM and the ResNet50-PCC-ELM, VGG19-PCC-ELM models for Scheme 1.

this study. The proposed model detected the tuberculosis disease from the regular patient with a precision and recall of 99% greater than the nine latest methods reported. For Scheme 6, the CNN-PCC-ELM model achieved an accuracy of 99.51% and an AUC of 100%, which were again much better than other SOTA models.

Additionally, it should be noted that although the majority of SOTA models required high-quality images (227×227; 224×224) to identify certain disorders, the proposed model employed low-resolution images (124×124) to detect the discriminating features [13,32,58,73,78,79,81]. As the number of pixels in the input CXR images are decreased, the complexity of the model is proportionately decreased as well.

As a result of the above discussion, it can be concluded that the suggested CNN-PCC-ELM model outperformed the selected 53 SOTA models available in the literature for various schemes that ensured efficiency and accuracy in diagnosing numerous lung illnesses with small number of features and low-resolution input CXR images.

Table 15 depicts that as the number of classes decreased, the model's performance improved. However, the accuracy of Scheme 5 was worse than the Scheme 6 even though both schemes fell in the category of binary class. This could be due to the differences in disease varieties and diverse features between the schemes. Similar arguments can also be reasoned for the difference in accuracy between Scheme 3 and Scheme 4. The number of redundant features increased from higher to the lower-class schemes (Table 5) and the number of redundant features can fluctuate depending on the disease types. As a result, training models for different schemes or datasets with the same number of features was not a wise move.

Using a number of deep TL models, the majority of researchers showed a favorable outcome in the detection of lung diseases in the past several years. However, the TL models required longer processing times because of their larger number of parameters and layers. The main goal of this study was to develop a lightweight CNN model with fewer parameters and layers, which would reduce the processing time compared to the conventional TL models. Therefore, the relevant features were extracted using a basic CNN model and unnecessary features were removed by employing PCC. The combination of CNN's feature extraction, PCC's most notable feature selection and ELM model's classification capabilities outperformed the earlier SOTA models and maintained high accuracy even in the presence of imbalanced datasets as well as low quality CXR images. Therefore, it can be argued that the proposed framework will ensure processing efficiency and correctness in faster detecting the lung diseases.

This study also solved another challenging task of detecting lung diseases from the subtle appearance of illness signs

Table 15 – Performance comparison with previous studies.

Scheme No	*Reference*	*No. of Classes*	*Precision*	*Recall*	*Accuracy*	*AUC*
Scheme 2	[24]	4	89.84%	89.94%	89.6%	–
	[26]		–	79.79%	87.07%	–
	[71]		90%	87.75%	89.89%	–
	[82]		80.92%	85.66%	76.46%	–
	[86]		90.13%	90.13%	90.13%	–
	CNN-PCC-ELM		**96%**	**97%**	**96.57%**	**99.45%**
Scheme 3	[25]	3	**99%**	**98%**	**98.32%**	99.01%
	[67]		93.7%	93.2%	93.3%	95.00%
	CNN-PCC-ELM		95%	94%	97.34%	**99.33%**
Scheme 4	[27]	3	–	–	93.41%	–
	[12]		–	–	72.6%	–
	[22]		–	86.7%	92.53%	
	[14]		97.95%	97.94%	97.94%	–
	[23]		92.70%	92.70%	95.99%	–
	[68]		93.33%	93.33%	93.3%	–
	[69]		96.33%	93%	97.97%	–
	[8]		94.8%	–	94.4%	94.4%
	[72]		98.05%	98.02%	97.99%	99.41%
	[76]		95%	95%	97.41%	96%
	[77]		99.02%	98.26%	97.12%	–
	[78]		96.67%	93.33%	96%	–
	[75]		95.86%	97.62%	97.51%	–
	[82]		94.27%	96.40%	95.66%	–
	[83]		–	–	88.52%	–
	[86]		96.45%	96.41%	96.41%	–
	[89]		–	97.62%	97.27%	99.60%
	CNN-PCC-ELM		**100%**	**99%**	**99.55%**	**99.97%**
Scheme 5	[5]	2	–	–	96%	–
	[7]		–	–	80.9%	–
	[8]		96%	–	96%	96%
	[9]		95.63%	–	95.78%	98.21%
	[22]		–	92.64%	96%	–
	[12]		–	–	89.6%	–
	[24]		93%	98.2%	89.6%	–
	[23]		–	93.92%	98.39%	96.48%
	[13]		97%	98%	98%	–
	[17]		94%	76%	89.47%	–
	[72]		–	98.81%	98.67%	99.10%
	[76]		95%	95%	98.06%	95%
	[82]		98.45%	98.45%	**99.81%**	–
	[10]		98.00%	98.00%	98.00%	–
	[84]		88.00%	94.00%	98.00%	–
	[85]		**99.18%**	98.37%	98.78%	–
	[87]		98%	99%	99%	–
	[88]		96%	91.14%	93.24%	96.86%
	[90]		–	93.62%	93.4%	–
	CNN-PCC-ELM		99%	**99%**	98.82%	**99.88%**
Scheme 6	[30]	2	–	91.94%	90.23%	–
	[70]		–	97.3%	98.7%	99%
	[71]		98.3%	100%	97.72%	100%
	[40]		–	–	97.59%	99%
	[33]		99.42%	99.40%	99.40%	99%
	[34]		98.57%	98.56%	98.60%	–
	[42]		–	–	84.7%	92.60%
	[74]		95.67%	95.10%	95.10%	–
	[75]		99.18%	99.16%	99.16%	–
	[80]		–	98.41%	97.23%	–
	CNN-PCC-ELM		**100%**	**100%**	**99.51%**	**100%**

in the CXR images, which may be difficult for the radiologists to distinguish. Sometimes these modest radiographic characteristics of various diseases such as TB, pneumonia, opacity, and others also mislead the classifier, reducing the system's diagnostic performance, as highlighted in previous studies [27]. This study considered eight lung-related diseases

(Scheme 1) and the proposed model distinguished the lung diseases accurately from the extracted features while attaining an optimistic classification performance (96.22%) in the case of multi-class classifications. From the above discussion, it can be summarized that the proposed framework may assist the radiologists in detecting multiple lung-related diseases from the CXR images accurately and confidently freeing up valuable time of medical doctors to engage with other high priority tasks.

In future, more lung-related diseases will be collected and used to further develop the proposed lightweight model, which can be applied to an embedded system that automatically detects different lung-related diseases from the CXR images. This can aid the medical practitioners in quickly detecting the diseases and providing appropriate treatment to the patients into real-world clinical care.

5. Conclusion

This study used an extensive dataset composed of 23,690 CXR images from seven types of lung diseases, including 4,192 CXR images of COVID-19 patients, to detect these diseases using ML and DL models. A lightweight CNN model with only three layers and fewer parameters has been used to extract 512 features, and PCC has efficiently reduced unnecessary redundant features. Finally, a simple ELM with a single hidden layer has been used to classify the life-threatening diseases. The proposed CNN-PCC-ELM has successfully detected the multi-class and classified the COVID-19 disease from the lung diseases with high classification performance and reduced complexity, parameters, layers, and time. For all the schemes, the proposed CNN-PCC-ELM outperformed several SOTA methods with a high accuracy of 96.22% and an AUC of 99.483% for eight class classification. At the same time, the proposed model achieved an optimistic AUC of 99.45%, 99.33%, 99.97%, 99.88%, and 100% for the Schemes 2, 3, 4, 5, and 6 which outperformed the most recent SOTA models. Again, most studies utilized transfer learning methods for recognizing COVID-19, TB, and pneumonia, which required pre-training and also required particular resolution images (like 224×224, 331×331). On the contrary, the proposed model did not require any pre-training, and with small-resolution images (124×124), the model correctly detects different types of lung-related diseases. Finally, comprehensive experiments demonstrate that the CNN-PCC-ELM model can accurately diagnose several well-known lung-related diseases with a lower computational overhead than the TL models, which can help radiologists and other doctors save patients' lives.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Md. Nahiduzzaman: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing – review & editing, Writing – original draft, Project administration. **Md Omaer Faruq Goni:** Investigation, Formal analysis, Writing – original draft, Writing – review & editing. **Md. Robiul Islam:** Supervision, Formal analysis, Validation, Writing – review & editing. **Abu Sayeed:** Supervision, Validation, Writing – review & editing. **Md. Shamim Anower:** Funding acquisition, Supervision, Validation, Writing – review & editing. **Mominul Ahsan:** Formal analysis, Supervision, Validation, Visualization, Writing – review & editing. **Julfikar Haider:** Formal analysis, Supervision, Visualization, Validation, Writing – original draft, Writing – review & editing. **Marcin Kowalski:** Formal analysis, Supervision, Validation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to thank Rajshahi University of Engineering and Technology (RUET) for supporting to conduct the research.

Appendix

Appendices.

Area Under Curve	AUC
Area Under the Receiver Operating Characteristics	AUROC
Chest X-Ray	CXR
Chest X-Ray8	CXR8
CNN-PCC-ELM	CPE
Computer Aided Diagnosis	CAD
Confusion Matrix	CM
Contrast Limited Adaptive Histogram Equalization	CLAHE
Convolutional Layers	CL
Convolutional Neural Network	CNN
Correlation Coefficients	CC
Cross-Validation	CV
Decompose, Transfer, and Compose	DETRAC
Deep Learning	DL
Extreme Learning Machine	ELM
Fully Connected	FC
Machine Learning	ML
Montgomery	MT
Neural Network	NN
Pearson Correlation Coefficient	PCC

REFERENCES:

[1] Ruuskanen O, Lahti E, Jennings LC, Murdoch DR. Viral pneumonia. Lancet 2011;377(9773):1264–75.
[2] World Health Organization, "Who coronavirus (covid-19) dashboard," 2021. Accessed on: May, 22, 2022. [Online]. Available at: https://COVID19.who.int/table
[3] World Health Organization, "Global tuberculosis report, 2021". Accessed on: May, 22, 2022. [Online]. Available at:

https://www.who.int/teams/global-tuberculosis-programme/tb-reports
[4] Kamel SI, Levin DC, Parker L, Rao VM. Utilization trends in noncardiac thoracic imaging, 2002–2014. J Am Coll Radiol 2017;14(3):337–42.
[5] Loey M, El-Sappagh S, Mirjalili S. Bayesian-based optimized deep learning model to detect covid-19 patients using chest x-ray image data. Comput Biol Med 2022;142 105213.
[6] Haghanifar A, Majdabadi MM, Choi Y, Deivalakshmi S, Ko S. Covid-cxnet: Detecting covid-19 in frontal chest x-ray images using deep learning. Multimed Tools Appl 2022;81 (21):30615–45.
[7] Ieracitano C, Mammone N, Versaci M, Varone G, Ali A-R, Armentano A, et al. A fuzzy-enhanced deep learning approach for early detection of covid-19 pneumonia from portable chest x-ray images. Neurocomputing 2022;481:202–15.
[8] Agrawal T, Choudhary P. Focuscovid: automated covid-19 detection using deep learning with chest x-ray images. Evol Syst 2022;13(4):519–33.
[9] Gayathri J, Abraham B, Sujarani M, Nair MS. A computer-aided diagnosis system for the classification of covid-19 and non-covid-19 pneumonia on chest x-ray images by integrating cnn with sparse autoencoder and feed forward neural network. Comput Biol Med 2022;141 105134.
[10] Kassania SH, Kassanib PH, Wesolowskic MJ, Schneidera KA, Detersa R. Automatic detection of coronavirus disease (covid-19) in x-ray and ct images: a machine learning based approach. Biocybern Biomed Eng 2021;41(3):867–79.
[11] Yousefi B, Kawakita S, Amini A, Akbari H, Advani SM, Akhloufi M, et al. Impartially validated multiple deep-chain models to detect covid-19 in chest x-ray using latent space radiomics. J Clin Med 2021;10(14):3100.
[12] Akter S, Shamrat F, Chakraborty S, Karim A, Azam S. Covid-19 detection using deep learning algorithm on chest x-ray images. Biology 2021;10(11):1174.
[13] Chowdhury MEH, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub ZB, et al. "Can ai help in screening viral and covid-19 pneumonia?". IEEE Access 2020;8:132665–76.
[14] Horry MJ, Chakraborty S, Paul M, Ulhaq A, Pradhan B, Saha M, et al. "Covid-19 detection through transfer learning using multimodal imaging data", IEEE. Access 2020;8:149808–24.
[15] Rasheed J, Hameed AA, Djeddi C, Jamil A, Al-Turjman F. A machine learning-based framework for diagnosis of covid-19 from chest x-ray images. Interdisciplinary Sciences: Computational Life Sciences 2021;13(1):103–17.
[16] Panwar H, Gupta P, 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 110190.
[17] Minaee S, Kafieh R, Sonka M, Yazdani S, Soufi GJ. Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning. Med Image Anal 2020;65 101794.
[18] Afshar P, Heidarian S, Naderkhani F, Oikonomou A, Plataniotis KN, Mohammadi A. Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images. Pattern Recogn Lett 2020;138:638–43.
[19] Abbas A, Abdelsamea MM, Gaber MM. Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network. Appl Intell 2021;51(2):854–64.
[20] Arias-Londono JD, Gomez-Garcia JA, Moro-Velazquez L, Godino-Llorente JI. "Artificial intelligence applied to chest x-ray images for the automatic detection of covid-19. a thoughtful evaluation approach", IEEE. Access 2020;8:226811–27.
[21] Alam N-A-A, Ahsan M, Based MA, Haider J, Kowalski M. Covid-19 detection from chest x-ray images using feature fusion and deep learning. Sensors 2021;21(4):1480.
[22] Pandit MK, Banday SA, Naaz R, Chishti MA. Automatic detection of covid-19 from chest radiographs using deep learning. Radiography 2021;27(2):483–9.
[23] Sekeroglu B, Ozsahin I. <? COVID19?> detection of covid-19 from chest x-ray images using convolutional neural networks. SLAS TECHNOLOGY: Translating Life Sciences Innovation 2020;25(6):553–65.
[24] Khan AI, Shah JL, Bhat MM. Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images. Comput Methods Programs Biomed 2020;196 105581.
[25] Nahiduzzaman Md, Goni MOF, Anower MS, Islam MR, Ahsan M, Haider J, et al. "A novel method for multivariant pneumonia classification based on hybrid CNN-PCA based feature extraction using extreme learning machine with CXR images", IEEE. Access 2021;9:147512–26.
[26] Yamaç M, Ahishali M, Degerli A, Kiranyaz S, Chowdhury ME, Gabbouj M. Convolutional sparse support estimator-based covid-19 recognition from x-ray images. IEEE Trans Neural Networks Learn Syst 2021;32(5):1810–20.
[27] Chandra TB, Verma K, Singh BK, Jain D, Netam SS. Coronavirus disease (covid-19) detection in chest x-ray images using majority voting-based classifier ensemble. Expert Syst Appl 2021;165 113909.
[28] Nahiduzzaman M, Islam MR, Hassan R. ChestX-Ray6: Prediction of multiple diseases including COVID-19 from chest X-ray images using convolutional neural network. Expert Syst Appl 2023;211 118576.
[29] Que Q, Tang Z, Wang R, Zeng Z, Wang J, Chua M, et al. 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE 2018;2018:612–5.
[30] Islam MR, Nahiduzzaman M. Complex features extraction with deep learning model for the detection of COVID19 from CT scan images using ensemble based machine learning approach. Expert Syst Appl 2022;195 116554.
[31] Serte S, Serener A. Early pleural effusion detection from respiratory diseases including covid-19 via deep learning. IEEE 2020;2020:1–4.
[32] Sahlol AT, Abd Elaziz M, Tariq Jamal A, Damaševičius R, Farouk Hassan O. A novel method for detection of tuberculosis in chest radiographs using artificial ecosystem-based optimisation of deep neural network features. Symmetry 2020;12(7):1146.
[33] Chandra TB, Verma K, Singh BK, Jain D, Netam SS. Automatic detection of tuberculosis related abnormalities in chest x-ray images using hierarchical feature extraction scheme. Expert Syst Appl 2020;158 113514.
[34] Rahman T, Khandakar A, Kadir MA, Islam KR, Islam KF, Mazhar R, et al. "Reliable tuberculosis detection using chest x-ray with deep learning, segmentation and visualization", IEEE. Access 2020;8:191586–601.
[35] Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225; 2017.
[36] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 4700–8.
[37] Van der Maaten L, Hinton G. Visualizing data using t-sne. J Mach Learn Res 2008;9(11).
[38] Rahman M, Cao Y, Sun X, Li B, Hao Y. Deep pre-trained networks as a feature extractor with xgboost to detect tuberculosis from chest x-ray. Comput Electr Eng 2021;93 107252.

[39] Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, et al. Xgboost: extreme gradient boosting. R package version 0.4-2 2015;1(4):1–4.
[40] Ayaz M, Shaukat F, Raja G. Ensemble learning based automatic detection of tuberculosis in chest x-ray images using hybrid feature descriptors. Phys Eng Sci Med 2021;44 (1):183–94.
[41] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (roc) curve. Radiology 1982;143(1):29–36.
[42] Lopes U, Valiati JF. Pre-trained convolutional neural networks as feature extractors for tuberculosis detection. Comput Biol Med 2017;89:135–43.
[43] "Chest x-ray (pneumonia, covid-19, tuberculosis)," 2021. Accessed on: December, 20, 2021. [Online]. Available at: https://www.kaggle.com/datasets/jtiptj/chest-xray-pneumoniacovid19tuberculosis
[44] ml-workgroup. "covid-19-image-repository," 2019. Accessed on: December, 20, 2021. [Online]. Available at: https://github.com/ml-workgroup/covid-19-image-repository/tree/master/png
[45] Joseph Paul Cohen. "covid-chestxray-dataset," 2019. Accessed on: December, 20, 2021. [Online]. Available: https://github.com/ieee8023/covid-chestxray-dataset/tree/master/images
[46] Arman Haghanifar. "Covid-cxnet," 2019. Accessed on: December, 20, 2021. [Online]. Available at: https://github.com/armiro/COVID-CXNet/tree/master/chest_xray_images/covid19
[47] BIMCV Medical Imaging Databank of the Valencia Region. "Bimcv-COVID-19," 2020. Accessed on: December 20, 2021. [Online]. Available: https://bimcv.cipf.es/bimcv-projects/bimcv-%20covid19/
[48] Ingus T. "Neo xrays," 2019. Accessed on: December, 20, 2021. [Online]. Available: https://www.kaggle.com/ingusterbets/neo-xrays
[49] Scott Mader K. "Pulmonary chest x-ray abnormalities," 2017. Accessed on: December, 20, 2021. [Online]. Available: https://www.kaggle.com/kmader/pulmonary-chest-xray-abnormalities
[50] Jaeger S, Candemir S, Antani S, Wáng Y-X-J, Lu P-X, Thoma G. Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. Quant Imaging Med Surg 2014;4 (6):475.
[51] TB Portal Tuberculosis Chest X-ray dataset for Belarus. Accessed on: December, 20, 2021. [Online] Available at: https://academictorrents.com/details/509f986b456b6fce04c15f9d1de22cd4ccb2c4b7
[52] Paul Mooney. "Chest x-ray images (pneumonia)," 2017. Accessed on: December, 20, 2021. [Online]. Available: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
[53] Radiological Society of North America. "Rsna pneumonia detection challenge," 2018. Accessed on: December, 20, 2021. [Online]. Available: https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data
[54] Agarap AF. Deep learning using rectified linear units (relu). arXiv preprint arXiv:1803.08375; 2018.
[55] Kovács G, Tóth L, Van Compernolle D, Ganapathy S. Increasing the robustness of cnn acoustic models using autoregressive moving average spectrogram features and channel dropout. Pattern Recogn Lett 2017;100:44–50.
[56] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980; 2014.
[57] Kira K, Rendell LA. A practical approach to feature selection. In: Machine learning proceedings. Elsevier; 1992. p. 249–56.
[58] Benesty J, Chen J, Huang Y, Cohen I. Pearson correlation coefficient. In: Noise reduction in speech processing. Springer; 2009. p. 1–4.
[59] Farrell J, Saloner G. Standardization, compatibility, and innovation. Rand J Econ 1985:70–83.
[60] Nahiduzzaman M, Nayeem MJ, Ahmed MT, Zaman MSU. Prediction of heart disease using multi-layer perceptron neural network and support vector machine. IEEE 2019;2019:1–6.
[61] Huang G-B, Zhu Q-Y, Siew C-K. Extreme learning machine: theory and applications. Neurocomputing 2006;70(1–3):489–501.
[62] Nahiduzzaman Md, Islam MR, Islam SMR, Goni MOF, Anower MS, Kwak K-S. Hybrid CNN-SVD based prominent feature extraction and selection for grading diabetic retinopathy using extreme learning machine algorithm. IEEE Access 2021;9:152261–74.
[63] Nahiduzzaman M, Islam MR, Goni MOF, Anower MS, Ahsan M, Haider J, et al. Diabetic retinopathy identification using parallel convolutional neural network based feature extractor and ELM classifier. Expert Syst Appl 2023;217 119557.
[64] Goni MOF, Mondal MNI, Islam SR, Nahiduzzaman M, Islam MR, Anower MS, et al. Diagnosis of malaria using double hidden layer extreme learning machine algorithm with CNN feature extraction and parasite inflator. IEEE Access 2023;11:4117–30.
[65] Swets JA. Measuring the accuracy of diagnostic systems. Science 1988;240(4857):1285–93.
[66] Powers DM. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. arXiv preprint arXiv:2010.16061; 2020.
[67] Rahman T, Chowdhury ME, Khandakar A, Islam KR, Islam KF, Mahbub ZB, et al. Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray. Appl Sci 2020;10(9):3233.
[68] Wang L, Lin ZQ, Wong A. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. Sci Rep 2020;10(1):1–12.
[69] Jain R, Gupta M, Taneja S, Hemanth DJ. Deep learning-based detection and analysis of covid-19 on chest x-ray images. Appl Intell 2021;51(3):1690–700.
[70] Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology 2017;284(2):574–82.
[71] Duong LT, Le NH, Tran TB, Ngo VM, Nguyen PT. Detection of tuberculosis from chest X-ray images: boosting the performance with vision transformer and transfer learning. Expert Syst Appl 2021;184:115519.
[72] Patro KK, Prakash AJ, Hammad M, Tadeusiewicz R, Pławiak P. SCovNet: a skip connection-based feature union deep learning technique with statistical approach analysis for the detection of COVID-19. Biocybern Biomed Eng 2023;43 (1):352–68.
[73] Yuan J, Wu F, Li Y, Li J, Huang G, Huang Q. DPDH-CapNet: a novel lightweight capsule network with non-routing for COVID-19 diagnosis using X-ray images. J Digit Imaging 2023:1–13.
[74] Iqbal A, Usman M, Ahmed Z. Tuberculosis chest X-ray detection using CNN-based hybrid segmentation and classification approach. Biomed Signal Process Control 2023;84 104667.
[75] Indumathi V, Siva R. An efficient lung disease classification from X-ray images using hybrid Mask-RCNN and BiDLSTM. Biomed Signal Process Control 2023;81 104340.
[76] George GS, Mishra PR, Sinha P, Prusty MR. COVID-19 detection on chest X-ray images using Homomorphic

Transformation and VGG inspired deep convolutional neural network. Biocybern Biomed Eng 2023;43(1):1–16.
[77] Azam S, Rafid ARH, Montaha S, Karim A, Jonkman M, De Boer F. Automated detection of broncho-arterial pairs using CT scans employing different approaches to classify lung diseases. Biomedicines 2023;11(1):133.
[78] Akbulut Y. Automated pneumonia based lung diseases classification with robust technique based on a customized deep learning approach. Diagnostics 2023;13(2):260.
[79] Shamrat FJM, Azam S, Karim A, Ahmed K, Bui FM, De Boer F. High-precision multiclass classification of lung disease through customized MobileNetV2 from chest X-ray images. Comput Biol Med 2023;155 106646.
[80] Zhang G, Luo L, Zhang L, Liu Z. Research progress of respiratory disease and idiopathic pulmonary fibrosis based on artificial intelligence. Diagnostics 2023;13(3):357.
[81] Rajagopal R, Karthick R, Meenalochini P, Kalaichelvi T. Deep convolutional spiking neural network optimized with arithmetic optimization algorithm for lung disease detection using chest X-ray images. Biomed Signal Process Control 2023;79 104197.
[82] Joshi RC, Yadav S, Pathak VK, Malhotra HS, Khokhar HVS, Parihar A, et al. A deep learning-based COVID-19 automatic diagnostic framework using chest X-ray images. Biocybern Biomed Eng 2021;41(1):239–54.
[83] Mishra NK, Singh P, Joshi SD. Automated detection of COVID-19 from CT scan using convolutional neural network. Biocybern Biomed Eng 2021;41(2):572–88.
[84] Munusamy H, Muthukumar KJ, Gnanaprakasam S, Shanmugakani TR, Sekar A. FractalCovNet architecture for COVID-19 chest X-ray image classification and CT-scan image segmentation. Biocybern Biomed Eng 2021;41 (3):1025–38.
[85] Murugan R, Goel T, Mirjalili S, Chakrabartty DK. WOANet: whale optimized deep neural network for the classification of COVID-19 from radiography images. Biocybern Biomed Eng 2021;41(4):1702–18.
[86] Rashid N, Hossain MAF, Ali M, Islam Sukanya M, Mahmud T, Fattah SA. AutoCovNet: Unsupervised feature learning using autoencoder and feature merging for detection of COVID-19 from chest X-ray images. Biocybern Biomed Eng 2021;41 (4):1685–701.
[87] Fang L, Wang X. COVID-RDNet: a novel coronavirus pneumonia classification model using the mixed dataset by CT and X-rays images. Biocybern Biomed Eng 2022;42 (3):977–94.
[88] Meng J, Tan Z, Yu Y, Wang P, Liu S. TL-med: a two-stage transfer learning recognition model for medical images of COVID-19. Biocybern Biomed Eng 2022;42(3):842–55.
[89] Gour M, Jain S. Automated COVID-19 detection from X-ray and CT images with stacked ensemble convolutional neural network. Biocybern Biomed Eng 2022;42(1):27–41.
[90] Patel RK, Kashyap M. Automated diagnosis of COVID stages from lung CT images using statistical features in 2-dimensional flexible analytic wavelet transform. Biocybern Biomed Eng 2022;42(3):829–41.