



Utilization of Deep Convolutional Neural Networks for Accurate Chest X-Ray Diagnosis and Disease Detection

Mukesh Mann¹ · Rakesh P. Badoni² · Harsh Soni³ · Mohammed Al-Shehri⁴ · Aman Chandra Kaushik^{5,6} · Dong-Qing Wei⁵

Received: 29 September 2022 / Revised: 6 March 2023 / Accepted: 6 March 2023 / Published online: 26 March 2023
© International Association of Scientists in the Interdisciplinary Areas 2023

Abstract

Chest radiography is a widely used diagnostic imaging procedure in medical practice, which involves prompt reporting of future imaging tests and diagnosis of diseases in the images. In this study, a critical phase in the radiology workflow is automated using the three convolutional neural network (CNN) models, viz. DenseNet121, ResNet50, and EfficientNetB1 for fast and accurate detection of 14 class labels of thoracic pathology diseases based on chest radiography. These models were evaluated on an AUC score for normal versus abnormal chest radiographs using 112120 chest X-ray14 datasets containing various class labels of thoracic pathology diseases to predict the probability of individual diseases and warn clinicians of potential suspicious findings. With DenseNet121, the AUROC scores for hernia and emphysema were predicted as 0.9450 and 0.9120, respectively. Compared to the score values obtained for each class on the dataset, the DenseNet121 outperformed the other two models. This article also aims to develop an automated server to capture fourteen thoracic pathology disease results using a tensor processing unit (TPU). The results of this study demonstrate that our dataset can be used to train models with high diagnostic accuracy for predicting the likelihood of 14 different diseases in abnormal chest radiographs, enabling accurate and efficient discrimination between different types of chest radiographs. This has the potential to bring benefits to various stakeholders and improve patient care.

✉ Mukesh Mann
mukesh.maan@iiitsonapat.ac.in

✉ Rakesh P. Badoni
rakeshbadoni@gmail.com

✉ Dong-Qing Wei
dqwei@sjtu.edu.cn

Harsh Soni
harsh.11912082@iiitsonapat.ac.in

Mohammed Al-Shehri
mshehri@kku.edu.sa

Aman Chandra Kaushik
amanbioinfo@sjtu.edu.cn

² Department of Mathematics, École Centrale School of Engineering, Mahindra University, Hyderabad 500043, India

³ Department of Information Technology, Indian Institute of Information Technology, Sonapat, Haryana 131029, India

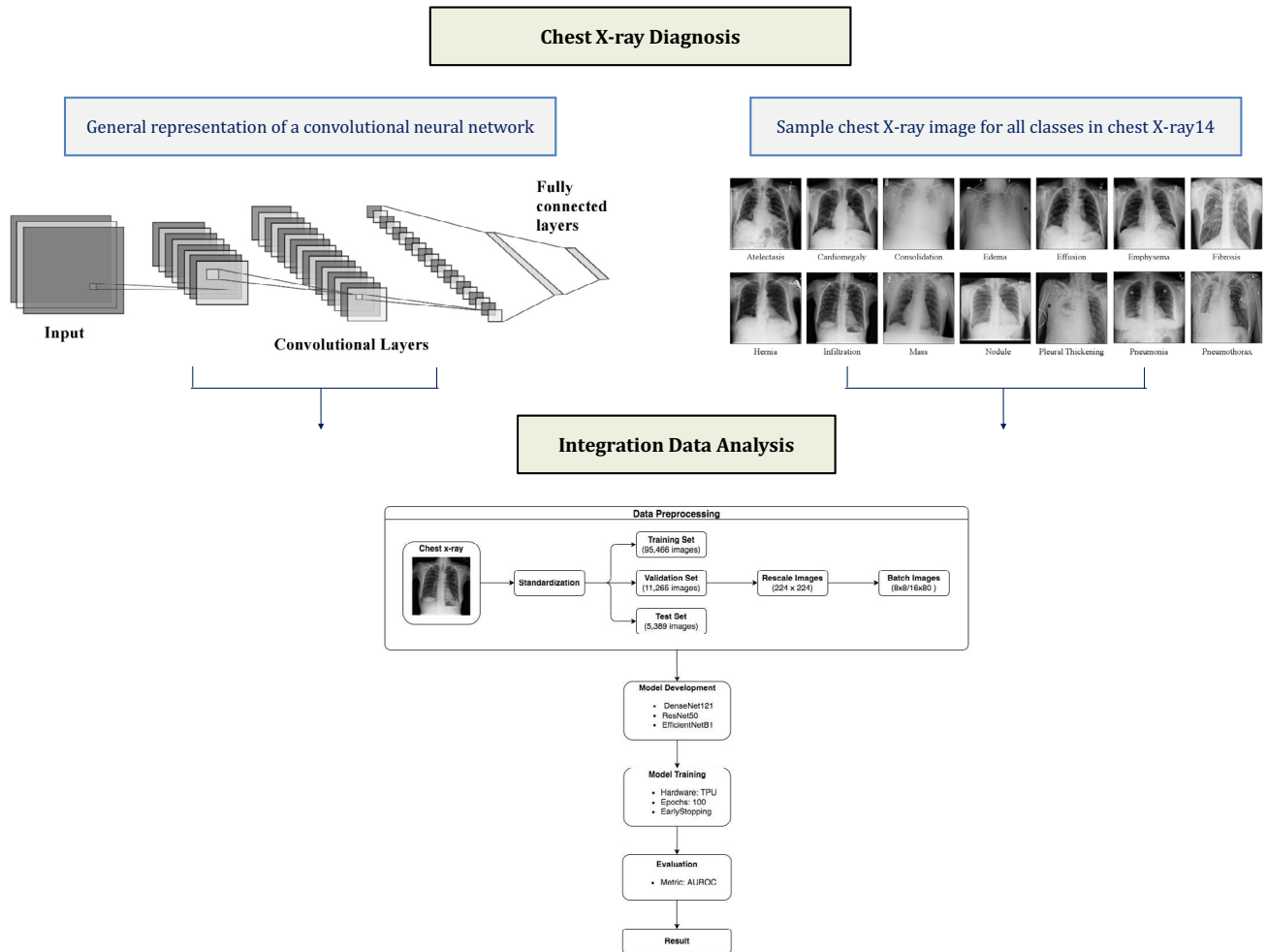
⁴ Department of Biology, Faculty of Science, King Khalid University, Abha, Saudi Arabia

⁵ State Key Laboratory of Microbial Metabolism, and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, 200030 Shanghai, China

⁶ School of Biomedical Informatics, University of Texas Health Science Centre at Houston, Houston, TX, USA

¹ Department of Computer Science and Engineering, Indian Institute of Information Technology, Sonapat, Haryana 131029, India

Graphical Abstract



Keywords Chest X-ray diagnosis · Chest X-ray14 · Deep ConvNet diagnosis · TPU training

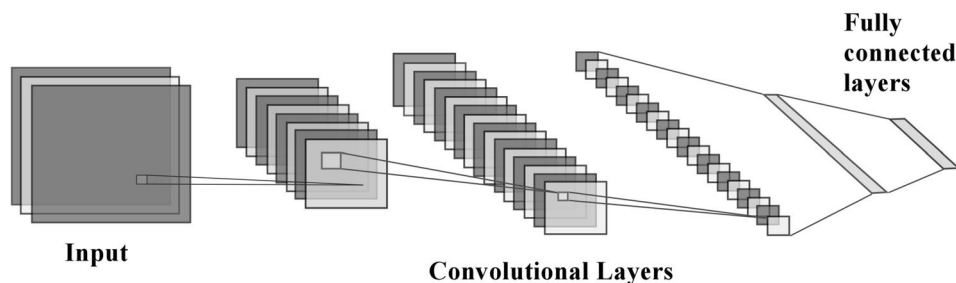
1 Introduction

Convolutional Neural Networks, also known as CNNs or ConvNets, have become a widely used method in the field of computer vision. Following the success of AlexNet [1] in ILSVRC 2012, CNNs have been extensively applied to various Cardiovascular (CV) tasks using deep learning. As time has passed, the architecture of CNNs has become increasingly complex and deep, leading to the development of deep ConvNets. With improvements in network architecture and computer hardware, we have been able to train these deep ConvNets, which have shown significant performance improvements in object detection, recognition, and image segmentation tasks. A general

representation of a convolutional neural network is provided in Fig. 1. The availability of large datasets and the progress in deep learning have made it possible for models to achieve human-level performance in many fields. Medical image analysis, such as the detection and segmentation of radiology images, has also yielded promising results using deep learning.

The detection of many diseases can be achieved through the use of chest radiology, which is the best and most commonly used clinical imaging tool. Every year, more than two billion chest radiology imaging tests are conducted [2]. This imaging tool is crucial for identifying various thoracic diseases, which are the leading cause of mortality worldwide. It would be exponentially beneficial if computer systems could

Fig. 1 General representation of a convolutional neural network



interpret chest X-rays with the same efficiency as a practicing radiologist [3, 4]. Over the last few years [5–9], the diagnosis of chest radiology images has received increased attention, and several algorithms have been developed for pulmonary tuberculosis classification [10–13] and pneumonia detection [3, 14, 15]. During the pandemic, deep learning has also found its usage in COVID-19 detection [16–21].

Our proposed study aims to identify multiple pathology diseases by re-implementing the CheXNet model and constructing several additional models with nearly identical hyperparameters. These models will be compared side-by-side. Additionally, we aim to train these models on a Tensor Processing Unit (TPU) to reduce training time. Deep learning has made significant advancements in the field of medicine due to the availability of vast datasets, enabling the development of models that surpass the performance of medical professionals. For instance, pneumonia detection [3, 14, 15], skin cancer classification [22–24], and lung cancer screening [25–27] have all benefited from deep learning. CheXNet [3], an algorithm that can detect pneumonia from chest X-rays, performs better than practicing radiologists. CNN models, such as CheXNeXt [4], can identify various pathologies diseases with a performance similar to that of practicing board-certified radiologists using frontal-view chest X-rays.

In recent years, various life-threatening diseases have been detected and diagnosed using deep learning techniques by a number of researchers [28–32]. Baltruschat et al. [28] conducted a study comparing multiple deep-learning approaches to classify chest X-Ray images with multiple labels. They analyzed various methods for using CNNs to classify X-ray images from the Chest X-ray14 dataset and found that fine-tuned networks using the ImageNet dataset produced satisfactory results. However, the most effective model was specifically trained using only X-ray images and incorporated non-image data. A systematic survey of deep learning techniques for the analysis of COVID-19 and their usability for detecting Omicron has been provided by [32]. The COVID-19 pandemic has caused a shift towards utilizing deep learning methods for analyzing and identifying infected areas in radiology images. These techniques can be divided into classification, segmentation, and multi-stage

approaches used for COVID-19 diagnosis at both the image and region levels. Khan et al. [33] introduced a new method called deep hybrid learning (DHL) and deep boosted hybrid learning (DBHL) for accurately detecting COVID-19 in chest X-ray images. The DBHL technique involves using data augmentation, transfer learning-based fine-tuning, deep features boosting, and hybrid learning to improve the performance of the COVID-RENet models (COVID-RENet-1 and COVID-RENet-2). In their experiments, the DBHL framework outperformed other well-established CNN models.

Stirenko et al. [11] conducted a study on the use of deep learning-based computer-aided diagnosis (CADx) to predict the presence of tuberculosis by statistically analyzing 2D chest X-ray images. They demonstrated the effectiveness of deep CNN for CADx of tuberculosis, particularly through techniques like lung segmentation and data augmentation, both lossless and lossy, on a small and unbalanced dataset. Rahman et al. [14] presented a study that aimed to develop an automated method for identifying bacterial and viral pneumonia by analyzing digital X-ray images. They gave a comprehensive overview of current approaches to detecting pneumonia and described the specific techniques used in their research. Alakus and Turkoglu [17] created clinical predictive models using deep learning and laboratory data to forecast which patients were likely to contract COVID-19. They evaluated the models' effectiveness using various performance metrics like precision, F1-score, recall, AUC, and accuracy with data from 600 patients and 18 laboratory findings and validated them through ten-fold cross-validation and train-test split methods. The results showed that the predictive models accurately identified patients with COVID-19. Dey et al. [15] designed a Deep-Learning System (DLS) for diagnosing lung abnormalities by using chest X-ray images. They tested the system with conventional and filtered chest radiographs and conducted an initial evaluation using a SoftMax classifier. The outcomes indicated that the VGG19 method provided higher classification accuracy than other methods.

Khan et al. [34] have proposed a new diagnostic system that employs deep CNNs to detect and analyze COVID-19 infections by identifying minor irregularities. This system

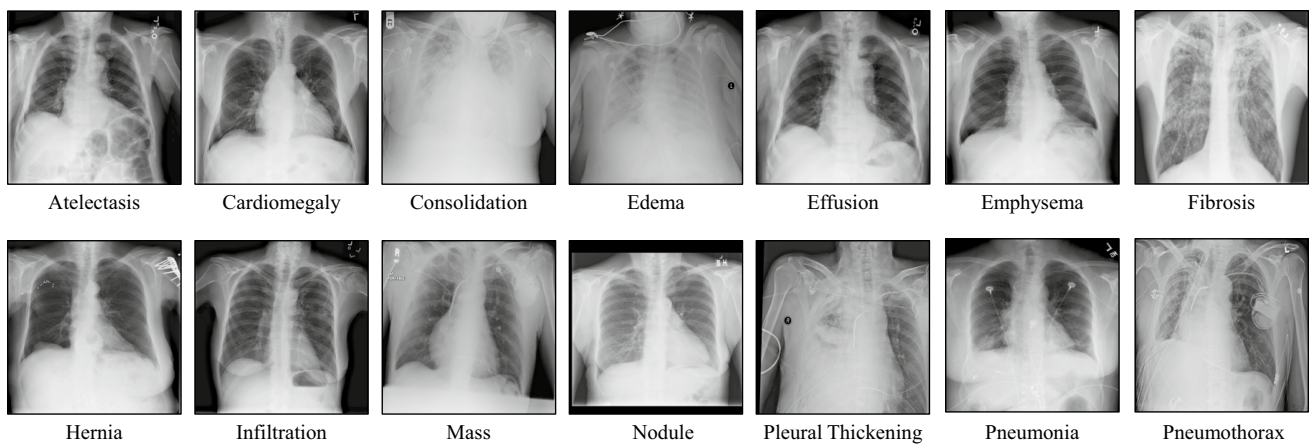


Fig. 2 Sample chest X-ray image for all classes in the chest X-ray14

comprises two phases. In the first phase, a new CNN named SB-STM-BRNet is utilized to identify COVID-19 infections in lung CT images. This is achieved using a Squeezed and Boosted (SB) channel and a Split-Transform-Merge (STM) block with dilated convolutions. In the second phase, the COVID-CB-RESeg CNN is employed to detect and analyze COVID-19-infected areas in the images. This CNN incorporates region-homogeneity and heterogeneity operations in each encoder-decoder block and auxiliary channels in the boosted-decoder to learn about low illumination and the boundaries of the infected regions. The proposed diagnostic system has shown promising results in identifying COVID-19 infections. Additionally, Khan et al. [35] have introduced a new CNN architecture called STM-RENet, which utilizes a split-transform-merge approach to analyze X-ray images and identify radiographic patterns associated with COVID-19 infection. This block-based CNN includes a new convolutional block named STM, which separately and jointly performs region and edge-based operations. By combining these operations with convolutional techniques, STM-RENet can analyze the homogeneity of regions, intensity inhomogeneity, and features that define boundaries. The authors have also presented an improved version of STM-RENet called CB-STM-RENet, which utilizes channel boosting and learns textural variations to enhance its performance. When evaluated on three datasets, CB-STM-RENet demonstrated significantly superior results than conventional CNNs.

A major limitation of previous research in the field of deep CNNs for chest X-ray diagnosis is that many studies have only examined their ability to perform binary classification tasks. These tasks involve detecting the presence or absence of a specific disease. There is a need for more research on the ability of these models to simultaneously detect and classify multiple diseases or conditions in a single image. Moreover, there are two additional issues in previous studies.

First, many studies have used small and potentially biased datasets, which negatively impacts the generalizability and accuracy of these models. Therefore, the need for more extensive and diverse datasets is essential. Second, there has been a lack of research on the ability of deep learning models to diagnose rare or less common diseases in chest X-rays accurately. Finally, previous studies have relied on simple accuracy metrics, which is inadequate for evaluating the performance of these models. More robust evaluation methods, such as sensitivity, specificity, and area under the curve, are required to understand these models' performance better.

2 Materials and Methods

2.1 Dataset

A significant amount of research has been done using the Chest X-ray14 dataset [3, 7–9, 36–38]. The dataset has been collected and made openly available by the National Institute of Health. It consists of 112120 frontal view chest X-ray images of 30805 unique patients. When loaded, these images are single-channel gray-scale images and need to be converted to 3-channel images to allow our pre-trained model to process them. Each image in the dataset is annotated with up to 14 different thoracic pathology labels. Figure 2 shows a sample for each of these diseases from the dataset itself. Table 1 lists all the diseases in chest X-ray14.

Wang et al. [39] used NLP to text-mine disease classifications from the related radiological reports to label the images. The labels are expected to have an accuracy greater than 90% [40]. The dataset also consists of images labeled as No Finding, which simply indicates that the NLP system

Table 1 List of all diseases in chest X-ray14

Pathology	Total positives
Atelectasis	11559
Cardiomegaly	2776
Effusion	13317
Infiltration	19894
Mass	5782
Nodule	6331
Pneumonia	1431
Pneumothorax	5302
Consolidation	4667
Edema	2303
Emphysema	2516
Fibrosis	1686
Pleural Thickening	3385
Hernia	227

was unable to find any diseases for that particular image. It does not necessarily imply a healthy chest X-ray image. To be able to train on a TPU, the entire dataset was converted to tfrecord and made publicly available at NIH Chest X-ray TFRecords.¹

2.2 Dataset Distribution

Table 1 provides a breakdown of the total positive labels for each of the 14 pathology diseases. To train and evaluate the models, the entire dataset was divided into three sets: a training set, a validation set, and a test set, in accordance with the recommendations of [41–44]. It is important to note that the dataset was not simply split into three parts. This is because the dataset contains follow-up images for each patient, and a direct split could result in data leaks, leading to misleading results. Instead, the dataset was first grouped by patient IDs, treating each patient as a separate entity, then split into the following three sets.

- Training dataset: 85% of the total groups.
- Validation dataset: 10% of the total groups.
- Test dataset: 5% of the total group.

Hence, the distribution of the dataset used is given as follows.

- Training dataset: 95466 examples.
- Validation dataset: 11265 examples.
- Test dataset: 5389 examples.

2.3 Dataset Preprocessing

The dataset needs to be preprocessed before building and training the model. The X-ray images in the dataset were preprocessed. For the training set, each image was standardized by subtracting the mean and standard deviation of that image from each pixel of that image.

$$\hat{X}_j^{[i]} = \frac{X_j^{[i]} - \bar{X}^{[i]}}{\sigma^{[i]}}$$

Here, i refers to the i^{th} image in the training set, j refers to the j^{th} pixel in the i^{th} image, and \hat{X} represents the standardized image. Similarly, \bar{X} represents the mean of an image, and σ represents the standard deviation of an image.

For standardization of validation and test set, the mean and standard deviation for each channel were calculated using a single batch of the training set, which was hypothesized to represent the mean and standard deviation of the entire training set. Then the individual images in the validation and test were standardized (feature-wise) using the above-stated formula. After standardizing, the images in the dataset were re-scaled to 224×224 ; this was done to remain consistent with the pre-trained models. The models used were pre-trained on the ImageNet dataset with an input size of 224×224 per image. After re-scaling, the images were batched, with each batch containing 16×8 (for DenseNet and ResNet50 models) and 8×8 (for EfficientNetB1 model) examples. The batch size was kept large to utilize TPU efficiently. Figure 3 shows the generalized workflow taken while developing models for this paper.

3 Model Development

3.1 Transfer Learning

Transfer learning refers to the idea of taking a model trained on a different task and using this pre-trained model for a downstream task. Figure 4 briefly describes the idea behind transfer learning. This paper uses a handful of models trained on the ImageNet [45] dataset, trains, or fine-tunes them for the task of chest X-ray diagnosis using transfer learning. All model variants are trained on a TPU, and their AUROC scores are recorded on the held-out test set.

3.2 DenseNet121

The first model used as the backbone for this task was the DenseNet121 [46]. Densely Connected Convolutional

¹ <https://www.kaggle.com/harshsoni/nih-chest-xray-tfrecords>

Fig. 3 Generalized workflow for the process

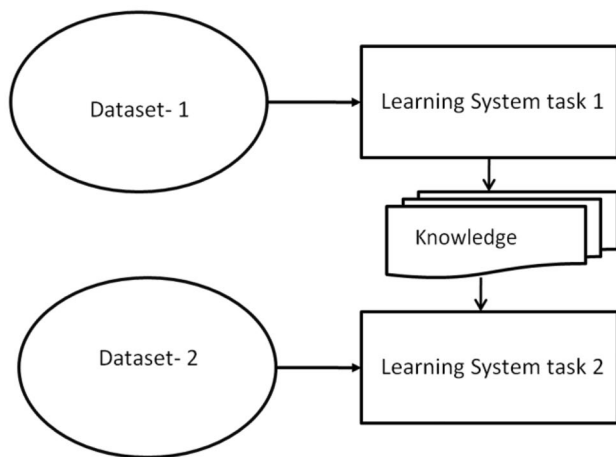
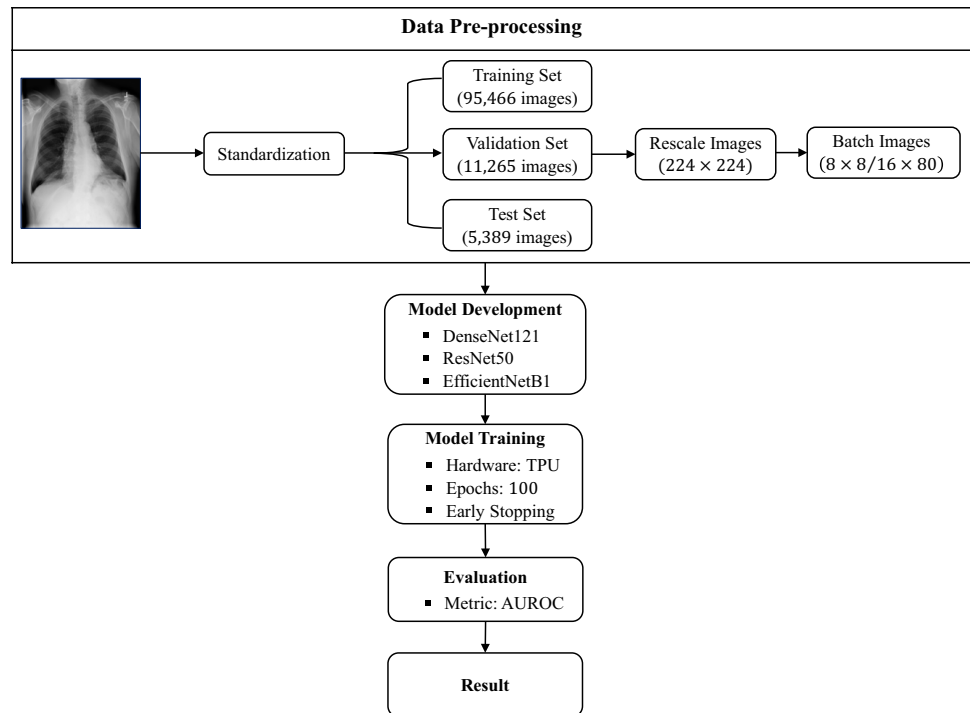


Fig. 4 Transfer learning process

Networks, or simply DenseNet, is another way to keep increasing the depth of a deep convolutional network. The problem in deep convnets arises when they become so deep that the gradients vanish on their way back. Huang et al. [46] designed an architecture to ensure maximum gradient flow during back-propagation to resolve this problem. DenseNet exploits the potential of the network through feature reuse. The architecture for DenseNets is exhibited in Fig. 6(a). This paper uses a DenseNet121, a 121-layered convolutional neural network model. This model derives inspiration from the

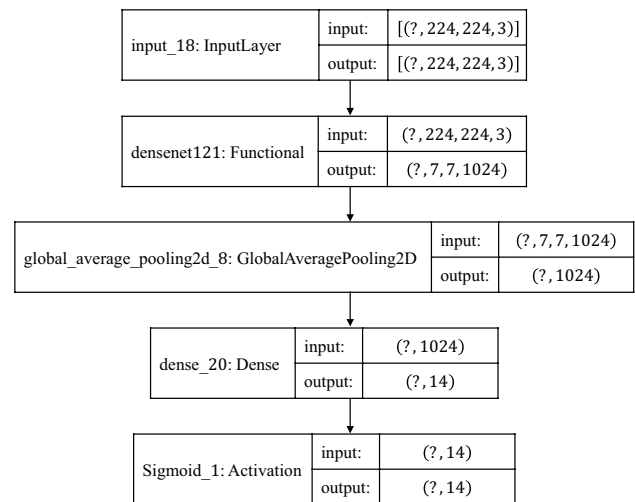


Fig. 5 Model architecture using DenseNet121 as a backbone

CheXNet model and is a re-implementation of the same. To use transfer learning on the DenseNet121, the final dense layer of the pre-trained model was replaced by a Dense layer with 14 units, and a sigmoid activation was applied to it. The resultant model architecture is shown in Fig. 5. In addition, the training parameters for the DenseNet121 backbone are given in Table 2.

(a)

Layers	Output Size	DenseNet – 121	DenseNet – 169	DenseNet – 201	DenseNet – 264
Convolution	112 × 112	7 × 7 conv, stride 2			
Pooling	56 × 56	3 × 3 max pool, stride 2			
Dense Block (1)	56 × 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56 × 56	1 × 1 conv			
	28 × 28	2 × 2 average pool, stride 2			
Dense Block (2)	28 × 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28 × 28	1 × 1 conv			
	14 × 14	2 × 2 average pool, stride 2			
Dense Block (3)	14 × 14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14 × 14	1 × 1 conv			
	7 × 7	2 × 2 average pool, stride 2			
Dense Block (4)	7 × 7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1 × 1	7 × 7 global average pool			
		1000D fully connected, softmax			

(b)

Layer Name	Output Size	18 – Layer	34 – Layer	50 – Layer	101 – Layer	152 – Layer
conv1	112 × 112	7 × 7, 64, stride 2				
conv2_x	56 × 56	3 × 3 max pool, stride 2				
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28 × 28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14 × 14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7 × 7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1 × 1	average pool, 1000-d fc, softmax				
FLOPs		1.8 × 10 ⁹	3.6 × 10 ⁹	3.8 × 10 ⁹	7.6 × 10 ⁹	11.3 × 10 ⁹

(c)

Stage <i>i</i>	Operator $\hat{\mathcal{F}}_i$	Resolution $\hat{\mathbf{H}}_i \times \hat{\mathbf{W}}_i$	# Channels $\hat{\mathbf{C}}_i$	# Layers $\hat{\mathbf{L}}_i$
1	Conv3x3	224 × 224	32	1
2	MBCConv1, k3x3	112 × 112	16	1
3	MBCConv6, k3x3	112 × 112	24	2
4	MBCConv6, k5x5	56 × 56	40	2
5	MBCConv6, k3x3	28 × 28	80	3
6	MBCConv6, k5x5	14 × 14	112	3
7	MBCConv6, k5x5	14 × 14	192	4
8	MBCConv6, k3x3	7 × 7	320	1
9	Conv1x1 & Pooling & FC	7 × 7	1280	1

Fig. 6 Model architecture on ImageNet for a DenseNet [46]; b ResNet [47]; and c EfficientNet-B0 baseline network [48]

Table 2 Training parameters for DenseNet121 backbone

Training parameter	Value
Input shape	(224, 224)
Output shape	(14,)
Batch size	16 × 8 (16 batches per TPU core)
Callbacks	Model checkpoint Reduce LR on plateau Early stopping

3.3 ResNet50

The second model used as the backbone for the diagnosis task was the ResNet50 [47]. ResNets are an exciting class of models and have served as the state-of-the-art model for various tasks. A deep neural network architecture tends to give a more significant error as compared to a comparatively shallow neural network. He et al. [47] overcame this problem by introducing a deep residual learning framework and skip or residual connections. Figure 6(b) exposes the architectures of different ResNets.

In this work, ResNet50V2 uses 50 layered deep convolutional neural networks with multiple residual connections. The pre-trained model was modified to apply transfer learning for this diagnosis task. For this task, the final fully connected layers of the pre-trained ResNet50V2 model were replaced by an average pooling, followed by a series of dense, relu, and dropout layers. A final dense layer with 14 units followed by a sigmoid was utilized for the output. The resultant model architecture is shown in Fig. 7. Also, the training parameters for the ResNet50V2 backbone are depicted in Table 3.

3.4 EfficientNetB1

The third and last model used as the backbone for this task was EfficientNetB1 [48]. EfficientNets are a class of efficiently designed models to optimize the model’s performance while having a considerably low amount of trainable parameters. Tan and Le [48] came up with a better way of scaling the network, which they call compound scaling, in which they selected an efficient scaling for all – width, depth, and image resolution. The baseline network architecture of EfficientNetB0 is shown in Fig. 6(c).

To use transfer learning on the EfficientNetB1, the final Dense layer of the pre-trained model was replaced by a Dense layer with 14 units, and a sigmoid activation was applied to it. The resultant model architecture is shown in Fig. 8. Also, the training parameters for the EfficientNetB1 backbone are presented in Table 4.

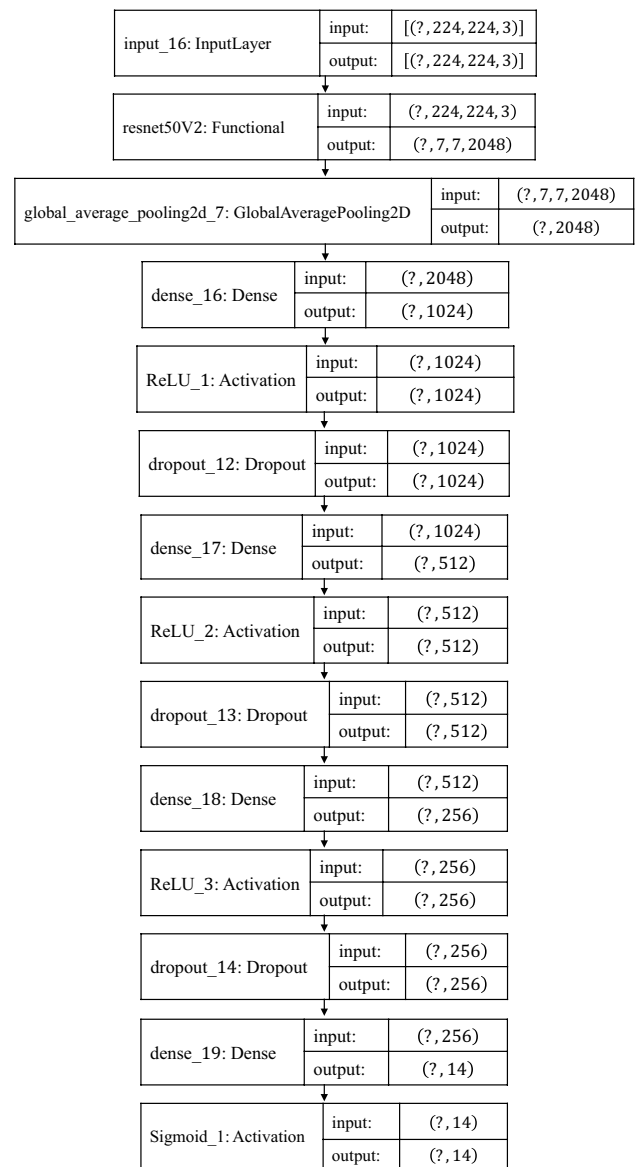


Fig. 7 Model architecture using ResNet50V2 as the backbone

Table 3 Training parameters for ResNet50V2 backbone

Training parameter	Value
Input shape	(224, 224)
Output shape	(14,)
Batch size	16 × 8 (16 batches per TPU core)
Callbacks	Model checkpoint Reduce LR on plateau Early stopping

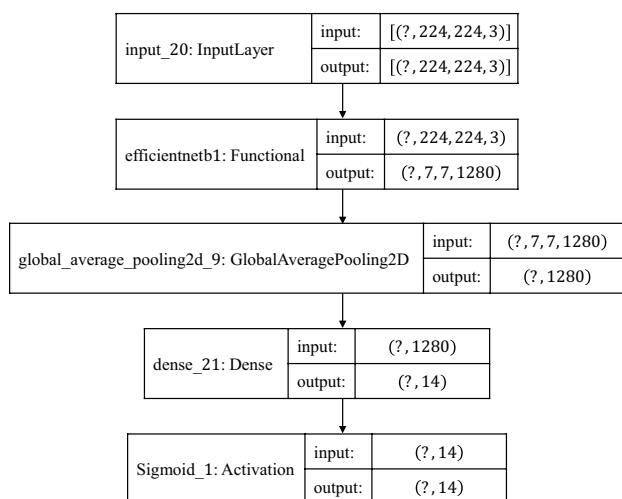


Fig. 8 Model architecture using EfficientNetB1 as a backbone

Table 4 Training parameters for EfficientNetB1 backbone

Training parameter	Value
Input shape	(224, 224)
Output shape	(14,)
Batch size	8 × 8 (8 batches per TPU core)
Callbacks	Model checkpoint Reduce LR on plateau Early stopping

4 Model Training

The models were individually trained on a TPU using various batch sizes while maintaining the same hyper-parameters. The models were initialized with parameters from a network that was pre-trained on ImageNet [25] before commencing training. The last layers were replaced with a dense layer having 14 units followed by a sigmoid layer for obtaining the predicted probabilities of all 14 pathology diseases, as discussed in the previous sections. The images were resized to 224 × 224 pixels before inputting them. Prior to feeding the network with the image, each image in the training set was subject to random horizontal flips and random rotations up to 10 degrees.

4.1 Loss Function

The data in the dataset were imbalanced. To account for this imbalance, instead of simply using a binary cross-entropy loss function, a weighted binary cross-entropy loss is minimized, as suggested in [3].

$$J = \sum_{c=1}^{14} L(y_c, \hat{y}_c)$$

$$L(y_c, \hat{y}_c) = \frac{1}{N} \sum_{i=1}^N [-w_{p_c} \times y_{c_i} \times \log(\hat{y}_{c_i}) - w_{n_c} \times (1 - y_{c_i}) \times \log(1 - \hat{y}_{c_i})]$$

Here, c refers to one of the classes of labels, i refers to the i^{th} example in the training set, y refers to true label, and \hat{y} refers to the predicted label or probability. Similarly, w_{p_c} and w_{n_c} are defined as follows.

$$w_{p_c} = \frac{\text{Total negative examples in class } c}{\text{Total examples}}$$

$$w_{n_c} = \frac{\text{Total positive examples in class } c}{\text{Total examples}}$$

The model was trained on a TPU V3.8 on Kaggle for 100 epochs. The above-mentioned custom loss, binary accuracy, and AUROC score were monitored during training. [49] optimizer was used for training. The learning rate was reduced by a factor of 10 if no improvements in validation loss were seen for two continuous epochs. Early stopping was used with the patience of 10 to prevent over-fitting of the model and prevent wastage of computing time. The end-to-end open-source deep learning framework Tensor Flow² was used to train and evaluate the models.

5 Results and Performance

The overall workflow for disease prediction in sample X-rays is graphically produced, as shown in Fig. 9.

In this study, the metric used for comparing the models is the AUROC score and curve. In this section, the results of the models are discussed and compared. Additionally, all the models are compared to the CheXNet model [3]. AUROC (Area Under the Receiver Operator Curve) is a performance metric that evaluates classification models for various threshold values. ROC is a probability curve, and AUC represents the degree of separability. The higher the AUC value, the better a model can differentiate between the positive and negative classes. An AUROC curve plots the true positive rate against the false-positive rate.

The true positive rate, also referred to as sensitivity, measures the proportion of positive examples in the dataset that the model accurately identified as positive. In other words, it represents the fraction of total positive examples correctly predicted as positive by the model, i.e.,

² <http://tensorflow.org>.



Fig. 9 Proposed CXD server, where **a** Front view of the CXD where boxes are given for sequences and blue button representing submission option; **b** Uploading Sample Chest X-ray; **c** Predicted results of

given sample using Deep Learning approaches; and **d** Depict previous history along with confidence scores generated from embedded Deep learning approaches

$$TPR = \frac{\text{True positive}}{\text{True positives} + \text{false negatives}}$$

The false positive rate, also given as 1-sensitivity, is the fraction of total negative examples in the dataset the model incorrectly predicted as positive, i.e.,

$$FPR = \frac{\text{False positive}}{\text{True positives} + \text{false negatives}}$$

The AUROC scores for different classes for all three model variants have been listed in Table 5, and the ROC curves have been illustrated in Fig. 10. The experimental results demonstrated that the model constructed using DenseNet121 outperformed the other two models. Therefore, our experimental investigations are in line with several studies [50–52] previously published in the literature. Some possible reasons for this superior performance could be:

Table 5 AUROC scores for different model variants

Pathology	CheXNet	DenseNet121	ResNet50	EfficientNetB1
Atelectasis	0.8094	0.8200	0.7630	0.7750
Cardiomegaly	0.9248	0.9120	0.7410	0.8840
Effusion	0.8638	0.8830	0.8600	0.8460
Infiltration	0.7345	0.7290	0.6660	0.6450
Mass	0.8676	0.8210	0.7480	0.7610
Nodule	0.7802	0.7200	0.5720	0.6240
Pneumonia	0.7680	0.7430	0.6940	0.7130
Pneumothorax	0.8887	0.8670	0.7790	0.8100
Consolidation	0.7901	0.8240	0.8030	0.7960
Edema	0.8878	0.8900	0.8710	0.8380
Emphysema	0.9371	0.9120	0.7600	0.8410
Fibrosis	0.8047	0.8080	0.6870	0.7310
Pleural Thickening	0.8062	0.7470	0.7150	0.7120
Hernia	0.9164	0.9450	0.6970	0.8490

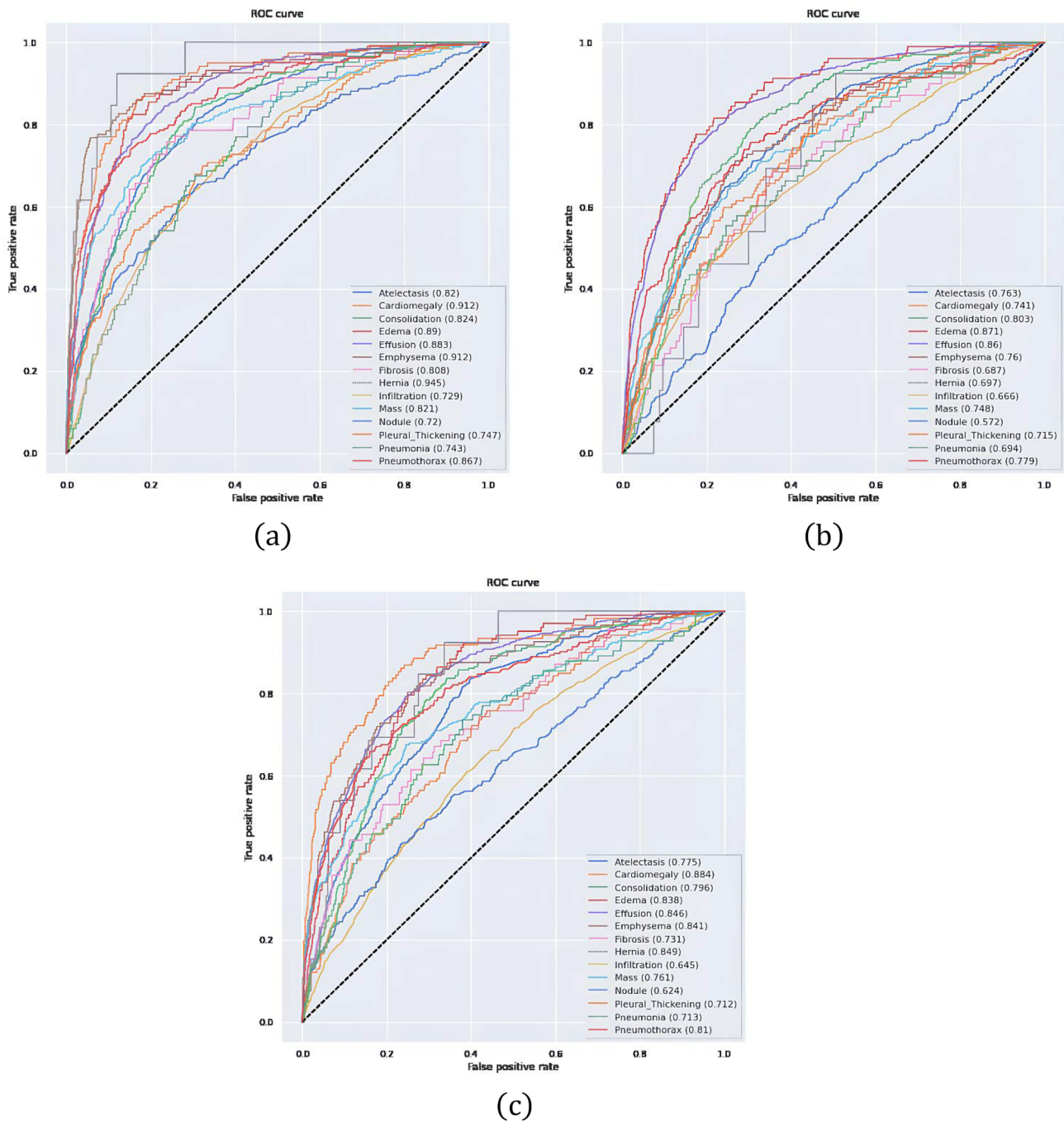


Fig. 10 **a** AUROC curve for DenseNet121 backbone; **b** AUROC curve for ResNet50 backbone; and **c** AUROC curve for EfficientNetB1 backbone

- The network architecture of DenseNet121 is more complex, enabling the model to extract more features from the data and potentially improve its performance.
 - The DenseNet121 model utilizes a more closely connected pattern between its layers, which can aid in decreasing the number of parameters within the model and avoiding overfitting.
 - DenseNet121 incorporates batch normalization and skip connections to enhance convergence and performance.
- Further, we employed a pre-trained DenseNet121 model and modified its fully connected layers as previously described. Subsequently, we assessed the model's

performance without any further training and found that the ROC values were approximately 0.5, indicating that its predictions were akin to random guessing. Figure 11(a) illustrates the results graphically. Similarly, we froze the DenseNet121 model and only trained the global average pooling (GAP) and final softmax layers, using the DenseNet121 model as the backbone, without modifying its parameters. The outcomes of this method are demonstrated in Fig. 11(b). Essentially, we maintained the original DenseNet121 model and only updated the final layers that were added to it.

Furthermore, DenseNet121 was utilized as an unalterable feature extractor, and a more intricate fully connected layer was trained on it. Only the fully connected layers were modified during the training process, and the rest of the DenseNet121 architecture remained unchanged. Figure 12 illustrates the new layers appended to the fully connected network. The ROC curve achieved from this approach is displayed in Fig. 13.

We have computed the lower and upper confidence intervals for ResNet50, EfficientNetB1, and DenseNet121 to further analyze these models. A confidence interval is a range of values that are likely to include the true value of a population parameter. The lower and upper confidence intervals for these models indicate the potential range of performance when applied to a specific task or dataset. For instance, if the lower confidence interval for a model’s accuracy is 95%, there is a 95% chance that the true accuracy

of the model will be at least that value. These intervals help to understand the uncertainty surrounding a model’s performance and to compare the performance of different models. Based on the assumption of a normal distribution, we have provided tables below that show the minimum and maximum estimated prevalence of Atelectasis disease with a 95% confidence level for ResNet50, EfficientNetB1, and DenseNet121 models. These tables, labeled as Tables 6, 7, 8, 9, 10,11, indicate the minimum estimated prevalence of Atelectasis disease with a 95% confidence level based on the assumption of a normal distribution for the respective models. In these tables, TPR represents the true positive rate, and FPR represents the false positive rate.

Next, we have thoroughly analyzed the lower and upper bounds of the precision-recall (PR) curves for the three models being considered. Nevertheless, it is important to note that the ROC curve is also a suitable alternative for evaluating a classifier’s performance, particularly for datasets that have imbalanced classes. This is because, unlike the PR curve, which only considers the true positive rate, the ROC curve considers both the true positive rate and the false positive rate. Therefore, we have included both the PR and ROC curves to provide a more comprehensive evaluation of these models, which are depicted in Figs. 14, 15, and 16.

The aforementioned tables, given as Tables 6, 7, 8, 9, 10,11, also display the F1 score, also known as the F-measure or F-score. It is a metric that combines

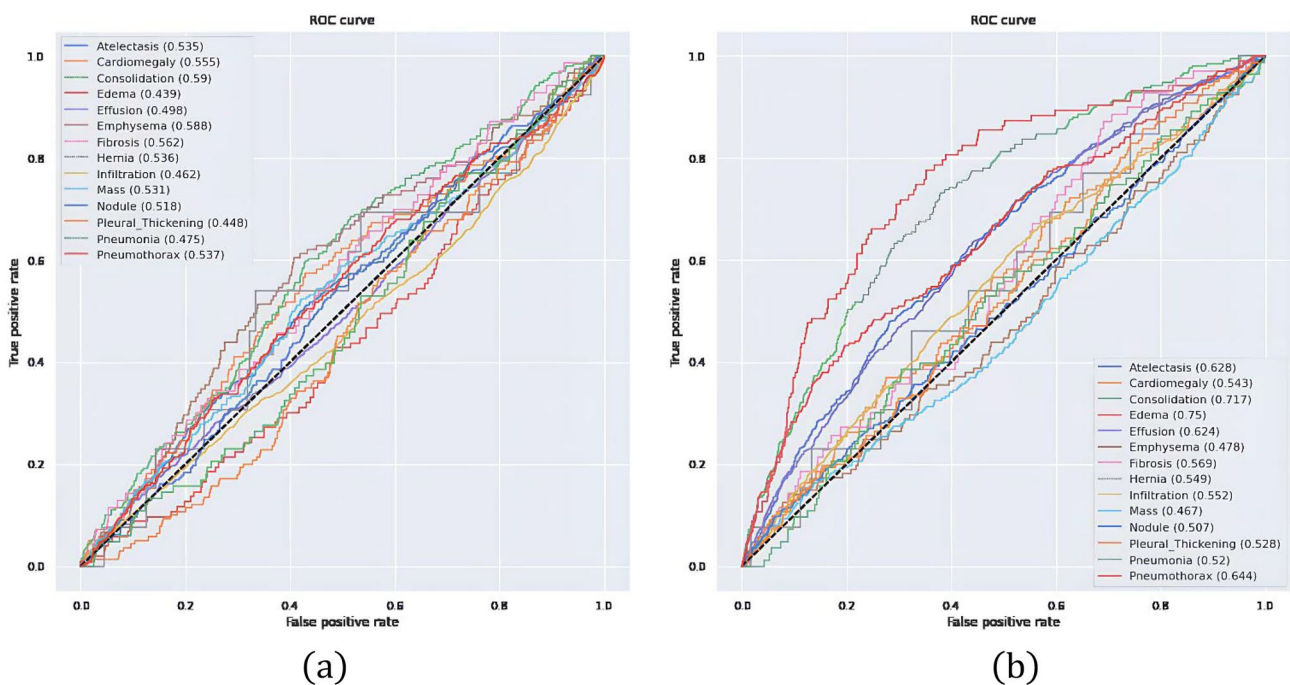
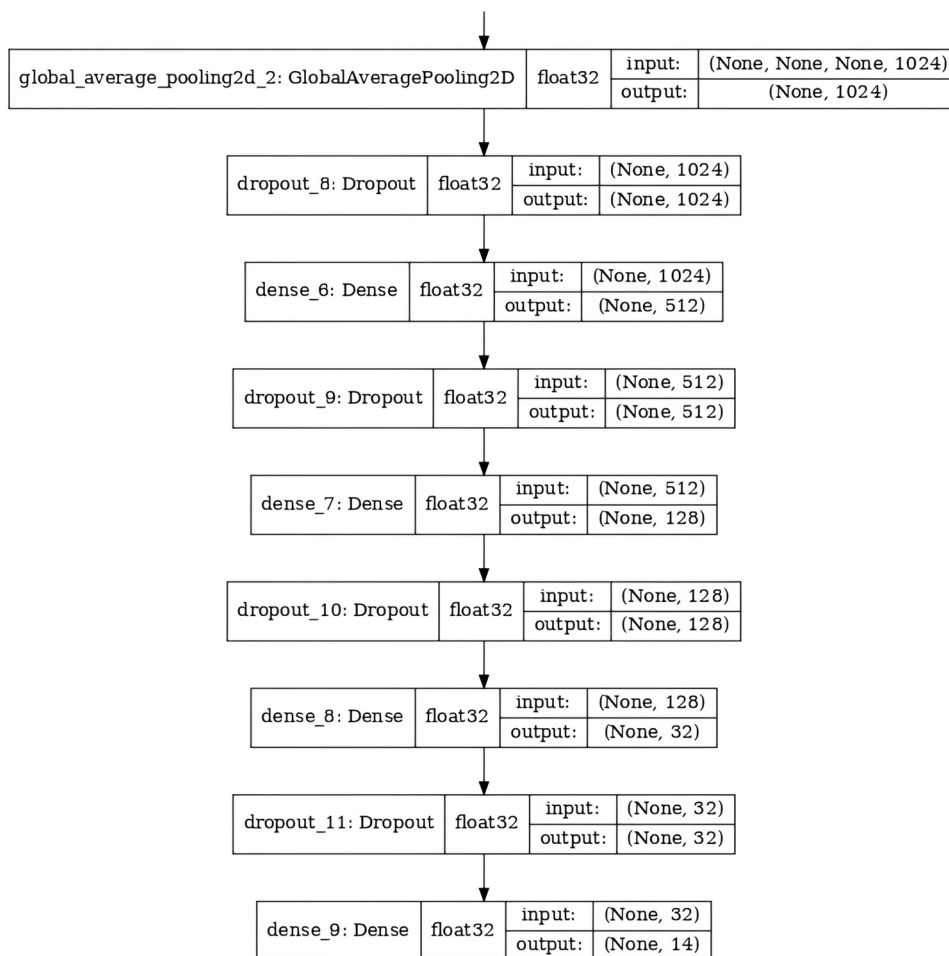


Fig. 11 **a** ROC curve for DenseNet121 without training model; and **b** ROC curve for DenseNet121 model after training only the GAP and final softmax layers

Fig. 12 The architecture of extra added layers to the fully connected network of DenseNet121



precision and recall into a single score, commonly used in classification tasks. The score is calculated as the harmonic mean of precision and recall. Precision is the number of true positive predictions divided by the total number of positive predictions, and recall is the number of true positive predictions divided by the total number of actual positive samples. The F1 score is valuable for evaluating the performance of classification models because it provides a balance between precision and recall and allows for the comparison of models with different precision and recall values.

6 Conclusions and Future Scopes

Pneumonia is a major cause of human fatalities worldwide. According to the Centers for Disease Control and Prevention,³ over one million adults in the US are hospitalized due to pneumonia, and around 50,000 die from the disease each

year. India has over 10 million cases of pneumonia each year. Although chest X-rays are the most effective means of diagnosing pneumonia [53], medical imaging has constraints in terms of access to expertise in some areas [54]. Additionally, chest radiographs may also be utilized to diagnose other illnesses.

In addition, even expert radiologists are limited by various human factors [38, 55–58]. Therefore, the creation of detection systems could greatly benefit humanity. As a result of the difficulty in training these three models on a CPU, this study considers using a TPU. The CXD server has an improved interface that is more efficient and has been developed using a large chest X-ray dataset up until January 2021. This extensive and accurate data is being constantly utilized to enhance our proposed CXD server, ensuring the quality of our work.

The objective of this study was to automate an essential stage of the radiology process by utilizing three convolutional neural networks (CNNs), namely DenseNet121, ResNet50, and EfficientNetB1, to precisely detect 14 types of thoracic pathology diseases from chest radiography images. A total of 112,120 chest X-ray datasets containing

³ <https://www.cdc.gov/features/pneumonia/index.html>.

Fig. 13 ROC curve for extra added layers to the fully connected network of DenseNet121

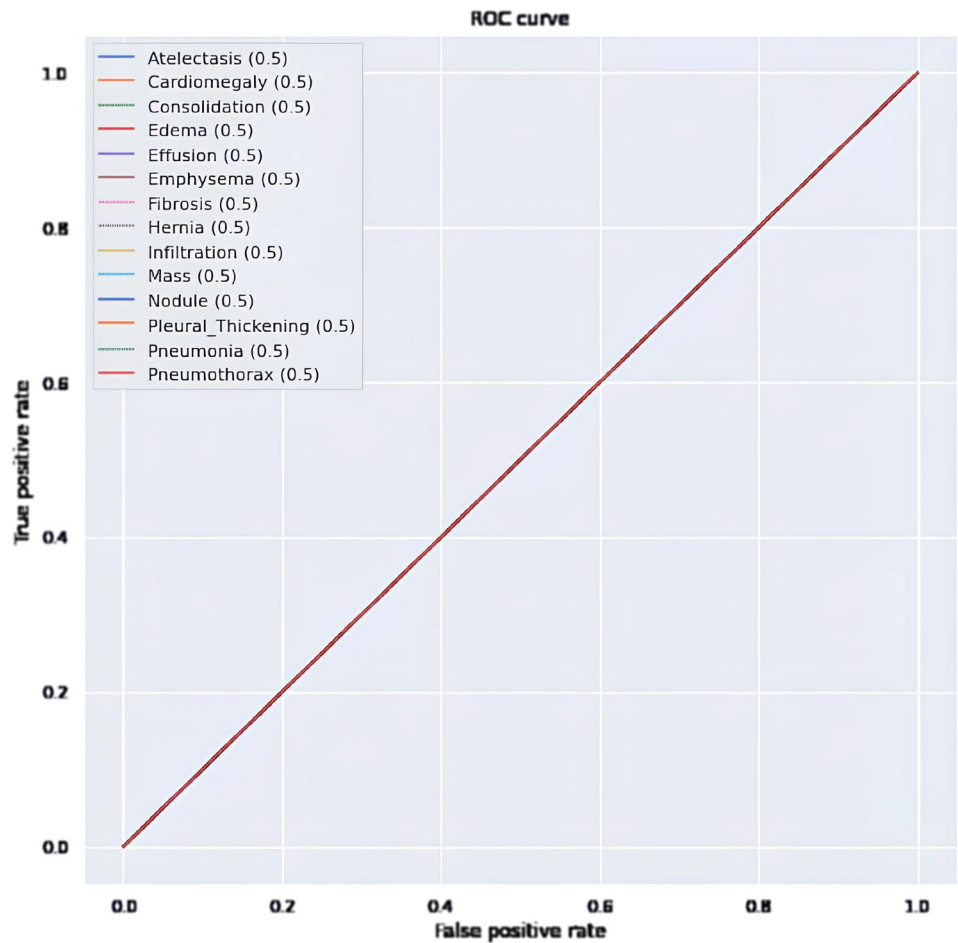


Table 6 The minimum estimated prevalence of Atelectasis disease with a 95% confidence level based on the assumption of a normal distribution for ResNet50

Threshold	TPR	FPR	Accuracy	Precision	Recall	F1 score
0.1	0.805903354	0.015975	0.895843	0.972882	0.805903	0.88118
0.2	0.841285476	0.029508	0.895843	0.966523	0.841285	0.899306
0.3	0.865006785	0.047695	0.895843	0.959846	0.865007	0.909793
0.4	0.882387525	0.068971	0.895843	0.954492	0.882388	0.916916
0.5	0.899576808	0.101926	0.895843	0.949523	0.899577	0.923817
0.6	0.918142515	0.163129	0.895843	0.942756	0.918143	0.93027
0.7	0.943963684	0.3045	0.895843	0.928266	0.943964	0.936041
0.8	0.972310777	0.600504	0.895843	0.912073	0.972311	0.941086
0.9	0.999218124	0.990938	0.895843	0.896111	0.999218	0.943991

various thoracic pathology diseases were utilized to evaluate the performance of these models based on their ability to predict the likelihood of individual diseases and alert

clinicians to potentially abnormal findings. The results indicated that the DenseNet121 model outperformed the other two models in terms of the score values achieved for each

Table 7 The maximum estimated prevalence of Atelectasis disease with a 95% confidence level based on the assumption of a normal distribution for ResNet50

Threshold	TPR	FPR	Accuracy	Precision	Recall	F1 score
0.1	0.82011924	0.020878	0.906718	0.978497	0.820119	0.892725
0.2	0.854381302	0.035998	0.906718	0.972778	0.854381	0.910013
0.3	0.877223562	0.05577	0.906718	0.966704	0.877224	0.919967
0.4	0.893879821	0.078499	0.906718	0.961793	0.89388	0.926705
0.5	0.91027128	0.113223	0.906718	0.95721	0.910271	0.933211
0.6	0.927862328	0.176824	0.906718	0.950935	0.927862	0.939273
0.7	0.952058015	0.321407	0.906718	0.937392	0.952058	0.944676
0.8	0.97798659	0.618292	0.906718	0.922126	0.977987	0.949381
0.9	0.999957983	0.994081	0.906718	0.906973	0.999958	0.952083

Table 8 The minimum estimated prevalence of Atelectasis disease with a 95% confidence level based on the assumption of a normal distribution for EfficientNetB1

Threshold	TPR	FPR	Accuracy	Precision	Recall	F1 score
0.1	0.69087922	0.00361946	0.89584263	0.98295658	0.69087922	0.81089795
0.2	0.82653939	0.02196173	0.89584263	0.97044309	0.82653939	0.89240883
0.3	0.88065577	0.06462379	0.89584263	0.95866499	0.88065577	0.91787436
0.4	0.91498338	0.14903047	0.89584263	0.94340845	0.91498338	0.92895791
0.5	0.93846433	0.27433506	0.89584263	0.93229964	0.93846433	0.9353707
0.6	0.95526986	0.44986595	0.89584263	0.92421408	0.95526986	0.93945243
0.7	0.97303584	0.65328083	0.89584263	0.91259259	0.97303584	0.94169824
0.8	0.98557233	0.84551181	0.89584263	0.90482787	0.98557233	0.94315702
0.9	0.99518978	0.97055598	0.89584263	0.89868163	0.99518978	0.94388506

Table 9 The maximum estimated prevalence of Atelectasis disease with a 95% confidence level based on the assumption of a normal distribution for EfficientNetB1

Threshold	TPR	FPR	Accuracy	Precision	Recall	F1 score
0.1	0.70759942	0.00616315	0.90671814	0.98736494	0.70759942	0.8249679
0.2	0.84012727	0.02763145	0.90671814	0.97631232	0.84012727	0.90344688
0.3	0.8922237	0.07388037	0.90671814	0.96562332	0.8922237	0.92760923
0.4	0.92487865	0.16224773	0.90671814	0.95154207	0.92487865	0.93804209
0.5	0.94693812	0.29075071	0.90671814	0.94117529	0.94693812	0.94404906
0.6	0.96250887	0.46803446	0.90671814	0.93358409	0.96250887	0.94785958
0.7	0.97863467	0.67052869	0.90671814	0.92261785	0.97863467	0.94995146
0.8	0.98960864	0.85845931	0.90671814	0.9152598	0.98960864	0.95130836
0.9	0.99740433	0.97641372	0.90671814	0.90941992	0.99740433	0.95198497

Table 10 The minimum estimated prevalence of Atelectasis disease with a 95% confidence level based on the assumption of a normal distribution for DenseNet121

Threshold	TPR	FPR	Accuracy	Precision	Recall	F1 score
0.1	0.84306717	0.03027407	0.89584263	0.96648739	0.84306717	0.90031113
0.2	0.87086545	0.05159917	0.89584263	0.96172846	0.87086545	0.91387505
0.3	0.88815469	0.07713721	0.89584263	0.95646389	0.88815469	0.92093865
0.4	0.90111604	0.10671467	0.89584263	0.95189192	0.90111604	0.92574521
0.5	0.91569139	0.15556943	0.89584263	0.94509502	0.91569139	0.93013811
0.6	0.9282404	0.21364821	0.89584263	0.93823539	0.9282404	0.93320833
0.7	0.93842546	0.28161416	0.89584263	0.93317418	0.93842546	0.93579162
0.8	0.94847708	0.37734796	0.89584263	0.92822659	0.94847708	0.93822927
0.9	0.96335002	0.56196969	0.89584263	0.9196265	0.96335002	0.94091021

Table 11 The maximum estimated prevalence of Atelectasis disease with a 95% confidence level based on the assumption of a normal distribution for DenseNet121

Threshold	TPR	FPR	Accuracy	Precision	Recall	F1 score
0.1	0.85610088	0.03684002	0.90671814	0.97274542	0.85610088	0.91096946
0.2	0.88284634	0.05996685	0.90671814	0.96842214	0.88284634	0.92383087
0.3	0.89938912	0.08714851	0.90671814	0.96360561	0.89938912	0.93049949
0.4	0.91173454	0.11823421	0.90671814	0.95939841	0.91173454	0.93502422
0.5	0.92554771	0.16901291	0.90671814	0.95310858	0.92554771	0.93914936
0.6	0.93736856	0.22878152	0.90671814	0.94672453	0.93736856	0.94202619
0.7	0.94690186	0.29815644	0.90671814	0.94199421	0.94690186	0.94444251
0.8	0.9562432	0.39509976	0.90671814	0.93735556	0.9562432	0.94671882
0.9	0.96989882	0.58001505	0.90671814	0.92926239	0.96989882	0.94921784

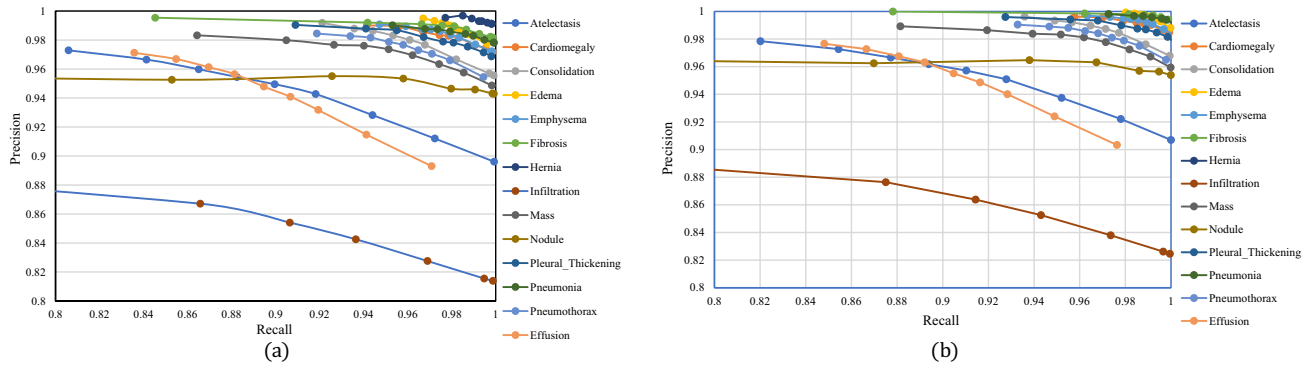


Fig. 14 PR curve for **a** lower bound, and **b** upper bound at 95% confidence interval for ResNet50 model

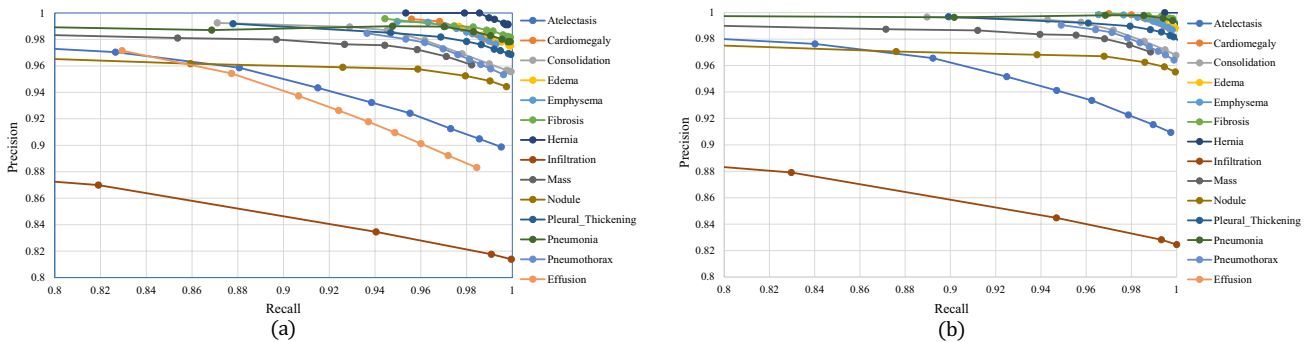


Fig. 15 PR curve for **a** lower bound, and **b** upper bound at 95% confidence interval for EfficientNetB1 model

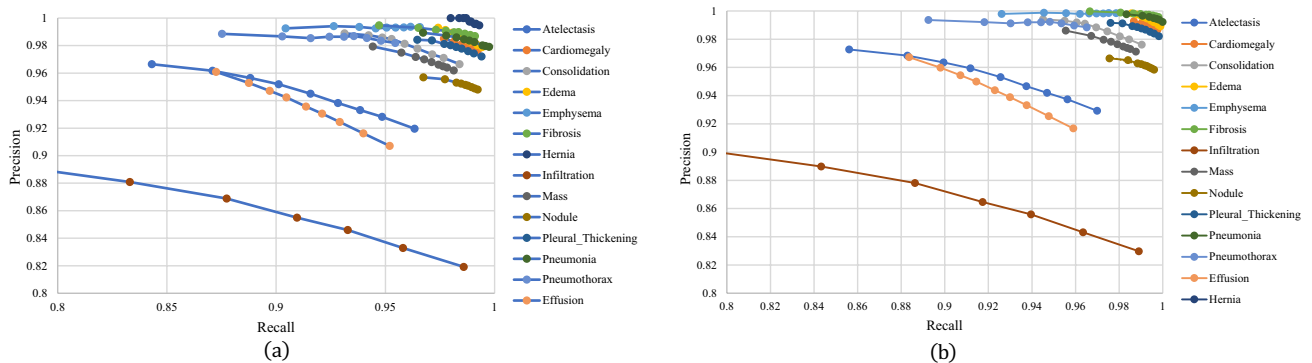


Fig. 16 PR curve for **a** lower bound, and **b** upper bound at 95% confidence interval for DenseNet121 model

class on the dataset. Furthermore, the performance of these models was compared to that of the ChexNet model.

Our future plans involve expanding our CNN training by incorporating extra data and assessing different architectures for diagnosing other thoracic pathology diseases. We are convinced that a computer-aided diagnostic tool of this kind could enhance the effectiveness and precision of diagnosing thoracic pathology diseases, including pandemics like COVID-19 and Swine Flu, significantly. This tool could prove to be especially useful during a pandemic when the demand for prevention and treatment often surpasses the available resources.

Appendix

The proposed server “CXD” is accessible at: <https://drive.google.com/file/d/1gKJNFfJc2FQoDo4IGz10wdTenbcAhC73/view?usp=sharing>.

NIH dataset (tfrecords) can be accessible at: <https://www.kaggle.com/harshsoni/nih-chest-xray-tfrecords>.

Kaggle TPU documentations can be accessible at: <https://www.kaggle.com/docs/tpu>.

The source code and sample data can be accessible at: <https://github.com/harsh020/cxd>.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12539-023-00562-2>.

Acknowledgements Dong-Qing Wei is funded by the National Science Foundation of China (Grant numbers 32070662, 61832019, 32030063), the Science and Technology Commission of Shanghai Municipality (Grant number 19430750600), the SJTU JiRLMDS Joint Research Fund, the Joint Research Funds for Medical and Engineering and Scientific Research at Shanghai Jiao Tong University (YG2021ZD02), and the PCL Major Key Project (PCL2021A13). The computations were partially performed at the Pengcheng Lab and the Center for High-Performance Computing at Shanghai Jiao Tong University.

Declarations

Conflict of interest The corresponding author states, on behalf of all authors, that there is no conflict of interest.

References

- Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. *Commun ACM* 60(6):84–90. <https://doi.org/10.1145/3065386>
- Raouf S, Feigin D, Sung A, Raouf S, Irugulpati L, Rosenow EC III (2012) Interpretation of plain chest roentgenogram. *Chest* 141(2):545–558. <https://doi.org/10.1378/chest.10-1302>
- Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz C, Shpanskaya K *et al.* (2017) Chexnet: radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint [arXiv:1711.05225](https://arxiv.org/abs/1711.05225). <https://doi.org/10.48550/arXiv.1711.05225>
- Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz CP *et al.* (2018) Deep learning for chest radiograph diagnosis: a retrospective comparison of the chexnet algorithm to practicing radiologists. *PLoS Med* 15(11):1002686. <https://doi.org/10.1371/journal.pmed.1002686>
- Cicero M, Bilbily A, Colak E, Dowdell T, Gray B, Perampaladas K, Barfett J (2017) Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs. *Investig Radiol* 52(5):281–287. <https://doi.org/10.1097/RLI.0000000000000341>
- Bar Y, Diamant I, Wolf L, Lieberman S, Konen E, Greenspan H (2015) Chest pathology detection using deep learning with non-medical training. In: 2015 IEEE 12th International Symposium on Biomedical Imaging, pp. 294–297. <https://doi.org/10.1109/ISBI.2015.7163871>. IEEE
- Guendel S, Grbic S, Georgescu B, Liu S, Maier A, Comaniciu D (2018) Learning to recognize abnormalities in chest x-rays with location-aware dense networks. In: Iberoamerican Congress on Pattern Recognition, pp. 757–765. https://doi.org/10.1007/978-3-030-13469-3_88. Springer
- Yan C, Yao J, Li R, Xu Z, Huang J (2018) Weakly supervised deep learning for thoracic disease classification and localization on chest x-rays. In: Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, pp. 103–110. <https://doi.org/10.1145/3233547.3233573>
- Wang H, Xia Y (2018) Chestnet: a deep neural network for classification of thoracic diseases on chest radiography. arXiv preprint [arXiv:1807.03058](https://arxiv.org/abs/1807.03058). <https://doi.org/10.48550/arXiv.1807.03058>
- Lakhani P, Sundaram B (2017) Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology* 284(2):574–582. <https://doi.org/10.1148/radiol.2017162326>
- Stirenko S, Kochura Y, Alienin O, Rokovyi O, Gordienko Y, Gang P, Zeng W (2018) Chest x-ray analysis of tuberculosis by deep learning with segmentation and augmentation. In: 2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO), pp. 422–428. <https://doi.org/10.1109/ELNANO.2018.8477564>. IEEE
- Maduskar P, Muyoyeta M, Ayles H, Hogeweg L, Peters-Bax L, vanGinneken B (2013) Detection of tuberculosis using digital chest radiography: automated reading vs. interpretation by clinical officers. *Int J Tuberc Lung Dis* 17(12):1613–1620. <https://doi.org/10.5588/ijtld.13.0325>
- Hwang S, Kim H-E, Jeong J, Kim H-J (2016) A novel approach for tuberculosis screening based on deep convolutional neural networks. In: Medical Imaging 2016: Computer-aided Diagnosis, vol. 9785, pp. 750–757. <https://doi.org/10.1117/12.2216198>. SPIE
- Rahman T, Chowdhury ME, Khandakar A, Islam KR, Islam KF, Mahbub ZB, Kadir MA, Kashem S (2020) Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest x-ray. *Appl Sci* 10(9):3233. <https://doi.org/10.3390/app10093233>
- Dey N, Zhang Y-D, Rajinikanth V, Pugalenth R, Raja NSM (2021) Customized vgg19 architecture for pneumonia detection in chest x-rays. *Pattern Recognit Lett* 143:67–74. <https://doi.org/10.1016/j.patrec.2020.12.010>
- Wang L, Lin ZQ, Wong A (2020) Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Sci Rep* 10(1):1–12. <https://doi.org/10.1038/s41598-020-76550-z>

17. Alakus TB, Turkoglu I (2020) Comparison of deep learning approaches to predict covid-19 infection. *Chaos Solit Fractals* 140:110120. <https://doi.org/10.1016/j.chaos.2020.110120>
18. Basu S, Mitra S, Saha N (2020) Deep learning for screening covid-19 using chest x-ray images. In: 2020 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 2521–2527. <https://doi.org/10.1109/SSCI47803.2020.9308571>. IEEE
19. Jamshidi M, Lalbakhsh A, Talla J, Peroutka Z, Hadjilooei F, Lalbakhsh P, Jamshidi M, La Spada L, Mirmozafari M, Dehghani M et al (2020) Artificial intelligence and covid-19: deep learning approaches for diagnosis and treatment. *IEEE Access* 8:109581–109595. <https://doi.org/10.1109/ACCESS.2020.3001973>
20. Zhang J, Xie Y, Li Y, Shen C, Xia Y (2020) Covid-19 screening on chest x-ray images using deep learning based anomaly detection. *arXiv preprint arXiv:2003.1233827*. <https://doi.org/10.48550/arXiv.2003.12338>
21. Wang S, Zha Y, Li W, Wu Q, Li X, Niu M, Wang M, Qiu X, Li H, Yu H et al (2020) A fully automatic deep learning system for covid-19 diagnostic and prognostic analysis. *Eur Respir J*. <https://doi.org/10.1183/13993003.00775-2020>
22. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542(7639):115–118. <https://doi.org/10.1038/nature21056>
23. Rezvantalab A, Safigholi H, Karimijeshni S (2018) Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms. *arXiv preprint arXiv:1810.10348*. <https://doi.org/10.48550/arXiv.1810.10348>
24. Hosny KM, Kassem MA, Foad MM (2018) Skin cancer classification using deep learning and transfer learning. In: 2018 9th Cairo International Biomedical Engineering Conference (CIBEC), pp. 90–93. <https://doi.org/10.1109/CIBEC.2018.8641762>. IEEE
25. Manning DJ, Ethell S, Donovan T (2004) Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph. *Br J Radiol* 77(915):231–235. <https://doi.org/10.1259/bjr/28883951>
26. Ciompi F, Chung K, Van Riel SJ, Setio AAA, Gerke PK, Jacobs C, Scholten ET, Schaefer-Prokop C, Wille MM, Marchiano A et al (2017) Towards automatic pulmonary nodule management in lung cancer screening with deep learning. *Sci Rep* 7(1):1–11. <https://doi.org/10.1038/srep46479>
27. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, Tse D, Etemadi M, Ye W, Corrado G et al (2019) End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med* 25(6):954–961. <https://doi.org/10.1038/s41591-019-0447-x>
28. Baltruschat IM, Nickisch H, Grass M, Knopp T, Saalbach A (2019) Comparison of deep learning approaches for multi-label chest x-ray classification. *Sci Rep* 9(1):1–10. <https://doi.org/10.1038/s41598-019-42294-8>
29. Ayan E, Ünver HM (2019) Diagnosis of pneumonia from chest x-ray images using deep learning. In: 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), pp. 1–5. <https://doi.org/10.1109/EBBT.2019.8741582>. IEEE
30. Khan SH, Sohail A, Khan A, Lee YS (2020) Classification and region analysis of covid-19 infection using lung ct images and deep convolutional neural networks. *arXiv preprint arXiv:2009.08864*. <https://doi.org/10.48550/arXiv.2009.08864>
31. Khan SH, Sohail A, Zafar MM, Khan A (2021) Coronavirus disease analysis using chest x-ray images and a novel deep convolutional neural network. *Photodiagnosis Photodyn Ther* 35:102473. <https://doi.org/10.1016/j.pdpdt.2021.102473>
32. Khan A, Khan SH, Saif M, Batool A, Sohail A, Waleed Khan M (2023) A survey of deep learning techniques for the analysis of covid-19 and their usability for detecting omicron. *J Exp Theor Artif Intell*. <https://doi.org/10.1080/0952813X.2023.2165724>
33. Khan SH, Sohail A, Khan A, Hassan M, Lee YS, Alam J, Basit A, Zubair S (2021) Covid-19 detection in chest x-ray images using deep boosted hybrid learning. *Comput Biol Med* 137:104816. <https://doi.org/10.1016/j.combiomed.2021.104816>
34. Khan SH (2022) Covid-19 detection and analysis from lung ct images using novel channel boosted cnns. *arXiv preprint arXiv:2209.10963*. <https://doi.org/10.48550/arXiv.2209.10963>
35. Khan SH, Sohail A, Khan A, Lee Y-S (2022) Covid-19 detection in chest x-ray images using a new channel boosted cnn. *Diagnostics* 12(2):267. <https://doi.org/10.3390/diagnostics12020267>
36. Kumar P, Grewal M, Srivastava MM (2018) Boosted cascaded convnets for multilabel classification of thoracic diseases in chest radiographs. In: International Conference Image Analysis and Recognition, pp. 546–552. https://doi.org/10.1007/978-3-319-93000-8_62. Springer
37. Yao L, Poblenz E, Dagunts D, Covington B, Bernard D, Lyman K (2017) Learning to diagnose from scratch by exploiting dependencies among labels. *arXiv preprint arXiv:1710.10501*. <https://doi.org/10.48550/arXiv.1710.10501>
38. Pesce E, Withey SJ, Ypsilantis P-P, Bakewell R, Goh V, Montana G (2019) Learning to detect chest radiographs containing pulmonary lesions using visual attention networks. *Med Image Anal* 53:26–38. <https://doi.org/10.1016/j.media.2018.12.007>
39. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM (2017) Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition, pp. 3462–3471. <https://doi.org/10.1109/CVPR.2017.369>
40. Summers R (2019) NIH chest X-ray dataset of 14 common thorax disease categories. *figshare* <https://nihcc.app.box.com/v/ChestXray-NIHCC/file/220660789610>
41. Chen Y, Zhang Y, Huang Z, Luo Z, Chen J (2021) Celebhair: A new large-scale dataset for hairstyle recommendation based on celeba. In: Knowledge Science, Engineering and Management: 14th International Conference, KSEM 2021, Tokyo, Japan, August 14–16, 2021, Proceedings, Part III, pp. 323–336. https://doi.org/10.1007/978-3-030-82153-1_27. Springer
42. Shrimali S (2021) Plantifyai: a novel convolutional neural network based mobile application for efficient crop disease detection and treatment. *Procedia Comput Sci* 191:469–474. <https://doi.org/10.1016/j.procs.2021.07.059>
43. Kenter T, Jones L, Hewlett D (2018) Byte-level machine reading across morphologically varied languages. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32. <https://doi.org/10.1609/aaai.v32i1.12050>
44. Mariani S, Rendu Q, Urbani M, Sbarufatti C (2021) Causal dilated convolutional neural networks for automatic inspection of ultrasonic signals in non-destructive evaluation and structural health monitoring. *Mech Syst Signal Process* 157:107748. <https://doi.org/10.1016/j.ymsp.2021.107748>
45. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L (2009) Imagenet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. <https://doi.org/10.1109/CVPR.2009.5206848>. IEEE
46. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2261–2269. <https://doi.org/10.1109/CVPR.2017.243>. IEEE Computer Society
47. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on

- Computer Vision and Pattern Recognition, pp. 770–778. <https://doi.org/10.1109/CVPR.2016.90>
48. Tan M, Le Q (2019) Efficientnet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning, pp. 6105–6114. <https://doi.org/10.48550/arXiv.1905.11946>. PMLR
 49. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980). <https://doi.org/10.48550/arXiv.1412.6980>
 50. Shazia A, Xuan TZ, Chuah JH, Usman J, Qian P, Lai KW (2021) A comparative study of multiple neural network for detection of covid-19 on chest x-ray. EURASIP J Adv Signal Process 2021(1):1–16. <https://doi.org/10.1186/s13634-021-00755-1>
 51. Hasan N, Bao Y, Shawon A, Huang Y (2021) Densenet convolutional neural networks application for predicting covid-19 using CT image. SN Comput Sci 2(5):389. <https://doi.org/10.1007/s42979-021-00782-7>
 52. Ogundokun RO, Maskeliūnas R, Misra S, Damasevicius R (2022) A novel deep transfer learning approach based on depth-wise separable CNN for human posture detection. Information 13(11):520. <https://doi.org/10.3390/info13110520>
 53. Organization WH et al. (2001) Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children. Technical report, World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/66956/WHO_V_and_B_01.35.pdf
 54. Welling RD, Azene EM, Kalia V, Pongpirul K, Starikovskiy A, Sydnor R, Lungren MP, Johnson B, Kimble C, Wiktorek S et al (2011) White paper report of the 2010 rad-aid conference on international radiology for developing countries: identifying sustainable strategies for imaging services in the developing world. J Am Coll Radiol 8(8):556–562. <https://doi.org/10.1016/j.jacr.2011.01.011>
 55. Fitzgerald R (2001) Error in radiology. Clin Radiol 56(12):938–946. <https://doi.org/10.1053/crad.2001.0858>
 56. Donovan T, Litchfield D (2013) Looking for cancer: expertise related differences in searching and decision making. Appl Cogn Psychol 27(1):43–49. <https://doi.org/10.1002/acp.2869>
 57. Bass JC, Chiles C (1990) Visual skill. Correlation with detection of solitary pulmonary nodules. Investig Radiol 25(9):994–998. <https://doi.org/10.1097/00004424-199009000-00006>
 58. Carmody DP, Nodine CF, Kundel HL (1980) An analysis of perceptual and cognitive factors in radiographic interpretation. Perception 9(3):339–344. <https://doi.org/10.1068/p090339>
- Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.