



Automated pneumonia detection on chest X-ray images: A deep learning approach with different optimizers and transfer learning architectures

Adhiyaman Manickam^a, Jianmin Jiang^a, Yu Zhou^a, Abhinav Sagar^b, Rajkumar Soundrapandian^{c,*}, R. Dinesh Jackson Samuel^e

^a Research Institute for Future Media Computing, College of Computer Science & Software Engineering, and Guangdong Laboratory of Artificial Intelligence & Digital Economy (SZ), Shenzhen University, China

^b School of Mechanical Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

^c School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

^e Faculty of Technology, Design and Environment, Visual Artificial Intelligence Lab, Oxford Brookes University, Oxford, United Kingdom

ARTICLE INFO

Keywords:

Pneumonia Detection
Deep Learning
Transfer Learning
ResNet 50
Inception V3
Adam Optimizer
Stochastic Gradient Descent Optimizer
U-Net
Convolutional Neural Networks
Accuracy

ABSTRACT

Pneumonia is a disease that leads to the death of individuals within a short period since the flow of fluid in the lungs. Hence, initial diagnosis and drugs are very important to avoid the progress of the disease. This paper proposes a novel deep learning approach for automatic detection of pneumonia using deep transfer learning to simplify the detection process with improved accuracy. This work was aimed to preprocess the input chest X-ray images to identify the presence of pneumonia using U-Net architecture based segmentation and classifies the pneumonia as normal and abnormal (Bacteria, viral) using pre-trained on ImageNet dataset models such as ResNet50, InceptionV3, InceptionResNetV2. Besides, to extract the efficient features and improve accuracy of pre-trained models two optimizers, namely, Adam and Stochastic Gradient Descent (SGD) used and its performances are analyzed with batch sizes of 16 and 32. Based on the values obtained, the performances of undertaken pre-trained models are analyzed and compared with other Convolutional Neural Network (CNN) models such as DenseNet-169+SVM, VGG16, RetinaNet + Mask RCNN, VGG16 and Xception, Fully connected RCNN, etc using various measures. From the results observed that the proposed ResNet50 model work achieved 93.06% accuracy, 88.97 % precision rate, 96.78% Recall rate and 92.71% F1-score rate, which than is higher than the other models aforementioned.

1. Introduction

In today's world, Healthcare research has immense importance and plays pivotal role in improving the living standards of peoples across the world in terms of addressing and identifying the correct diagnosis of conditions and diseases of patients. Problems faced by the Medical professionals to correctly diagnosing a condition has made the researchers and academicians to work on healthcare research and its related issues in improving the living standards of both the doctor and the patient i.e. avoiding unnecessary treatments, medications, malpractice lawsuits and etc [1].

Though, mankind suffers with various diseases, Pneumonia is considered as "one of the common acute respiratory lung disease" and among the top diseases which cause most of the deaths all over the

world. The existing Medical technology's failure to detect pneumonia at its earlier stage has taken at least of 1.4 millions of children's lives, and around 1 million adults were hospitalized where the death count was at the range of 50,000 every year in the USA alone. However, detecting pneumonia is not only depending on the medical technologies in practice, but also on the availability of expert radiologists [2]. In addition, the World Health Organization (WHO) stated that around 4 million premature deaths occur annually and over 150 million peoples get infected with pneumonia due to the non-availability of medical resources and expert professionals across the world [3,40].

Chest X-ray analysis is the most common analysis method used in diagnosing pneumonia disease and one of the most challenging tasks faced by the researchers is categorizing pneumonia's subtypes: Viral Pneumonia (VP) and Bacterial Pneumonia (BP) where VP requires

* Corresponding author.

E-mail addresses: adhiyaman.m@szu.edu.cn (A. Manickam), jianmin.jiang@szu.edu.cn (J. Jiang), yu.zhou@szu.edu.cn (Y. Zhou), abhinav.sagar2016@vitstudent.ac.in (A. Sagar), rajkumars@vit.ac.in (R. Soundrapandian), rsamuel@brookes.ac.uk (R. Dinesh Jackson Samuel).

Table 1

summary of the pneumonia detection methods for chest X-ray images and CT images.

| Reference | Model | Dataset/Type of Images | No. of Images | Evaluation measures | Accuracy (%) |
|-----------|----------------------------------|---|----------------------|--|--------------|
| [6] | DCNN | OpenI Dataset (Chest X-ray images) | 108,948 | Precision, Recall, F1-score, Accuracy | 99.76 |
| [7] | Random Sets of DCNN | CT images | 983 | Sensitivity, False Positive | – |
| [8] | 3D DCNN | LUNA16 Dataset (CT images) | 1018 | Free response receiver operating characteristic, Competition performance metric, Sensitivity | – |
| [9] | Mask-RCNN (ResNet50 + ResNet101) | NIH CXR14 dataset (X-ray images) | 25,684 | Intersection over union, Mean threshold value, Mean score | – |
| [10] | CMixNet | LUNA16 and LIDC-IDRI datasets (CT images) | 1906 | Sensitivity, Specificity, Accuracy | 94.17 |
| [11] | 3D screening CNN | LIDC dataset (CT images) | 1018 | Sensitivity, False positive rate | – |
| [12] | Neuro-heuristic approach | Montgomery country, Project and Shenzhen dataset (X-ray images) | 862 | Sensitivity, Specificity, Accuracy, Misclassification error | 79.06 |
| [13] | Inception V3 | Optical coherence tomography dataset (X-ray images) | 207,130 | Sensitivity, Specificity, Accuracy | 96 |
| [14] | Deep Siamese network | Kaggle dataset (X-ray images) | 5528 | Area under curve | – |
| [15] | DCNN | TB Chest X-ray images | 1007 | Sensitivity, Specificity, Area under curve | – |
| [16] | Different Pre-trained CNN | Kaggle dataset (X-ray images) | 5247 | Sensitivity, Specificity, Accuracy, F1-score, Area under curve | 93.3 |
| [17] | CheXNet | X-ray images | 112,120 | Area under curve | – |
| [18] | Dual CNN | MIMIC-CXR dataset (X-ray images) | 473,064 | Area under curve | – |
| [19] | CNN | ILD dataset (CT images) | 14,696 image patches | Accuracy | 85.61 |
| [20] | CNN | X-ray images | – | Area under curve | – |
| [21] | Different pre-trained model | Kaggle dataset (X-ray images) | 5840 | Accuracy | 92.31 |
| [22] | MAN + SVM | X-ray images | 2000 | Accuracy, Sensitivity, Specificity, F1-score | 96.80 |

supportive care and different medication process for recovery whereas BP requires antibiotics treatment for recovery. Deep learning methodologies are used for identifying and recognizing the type of pneumonia disease whereas automated detection helps in reducing time and improving accuracy. In some cases, it is difficult to identify the correct region of interest from the radiographic image, yet researchers have observed that diagnostic accuracy of automated method would reach the human level in analyzing and segmenting an image [4,5].

Traditionally, classifying and detecting pneumonia has been done by trained radiologists. It is very common for radiologists to make a mistake as to the naked eye there are cases where pneumonia could be missed which are also known as false negatives. In other cases, radiologists could diagnose someone as having pneumonia even if he doesn't have it in real life. Both the cases should be avoided which is where convolutional neural networks can help. It is evident from the aforementioned facts that the development of cheap and accurate pneumonia diagnostics is required. Hence, in this paper, deep learning approach along with transfer learning was developed to simplify the pneumonia detection process.

The novelty of this paper are as follows:

- To classify the pneumonia patient by considering the Chest X-ray images, in this paper proposed the method using pre-trained model (State-of-the-art) deep CNN descriptor. The CNN descriptor is used to extract deep convolutional feature for discriminate between normal and pneumonia cases.
- The input chest X-ray images are pre-processed to identify the presence of pneumonia using U-Net architecture based segmentation method.
- Pre-processing method is applied to the dataset to reduce bias toward the model performance.
- To overcome the over-fitting problem in deep learning due to limited number of training images, in this paper transfer learning based deep model proposed.
- Prepared a dataset of around 5300 Chest X-ray images for pneumonia detection from unbalanced public (kaggle) dataset.
- A customized optimization function used to help the physicians to identify the pneumonia with high accuracy and this process takes a short time.

- Evaluating the performance of the proposed model compared to existing model.

The research paper is organized as follows: **Section 1** introduces the importance of this topic. **Section 2** explores literature work related to proposed model. **Section 3** gives the short information of the dataset used in this paper. **Section 4** discusses the methodology of the paper. This section also presents the dataset used to train and test the three models. **Section 5** presents the experimental result and performance analysis, and Finally, **Section 6** concludes this paper.

2. Related literature work

Many significant research papers have been published to classify the chest x-ray images into different classes such as normal, pneumonia bacterial, pneumonia virus, etc. Xiaosong et al. [6] proposed a ChestX-ray8 dataset with 108,948 frontal view X-ray images of around 32,717 patients where Deep Convolutional Neural Network (DCNN) was employed in analyzing and recognizing pneumonia infection and its subtypes. The end results of this study were promising and are highly accurate for the proposed ChestX-ray8 dataset. In addition, the ChestX-ray8 dataset can be extended with new additional disease labels of patients. Holger et al. [7] showed how deep CNN can be used to identify lymph node(s): Lymph nodes are small structures that work as filters for harmful substances; sudden swelling of these nodes may indicate cancer, lung infection and other deadly diseases. This study has produced greater results for all images irrespective of image quality. Hwejin et al. [8] addressed the gradient vanishing problem by employing two 3D deep CNN, one with shortcut connections and other with dense connections. Both general and specific features of lung nodules were captured with these connections. However, the results were not accurate and suffered with time delay. Amit Kumar et al. [9] used Mask-RCNN which performed pulmonary image segmentation combined with image augmentation by utilizing both global and local features of radiographic imagery. Along with image augmentation, this study had used dropout and L2 regularization to identifying pneumonia infection. This study had produced much improved results. Nasrullah et al. [10] employed 3D –Customized mixed link network (CMixNet) architectures to learn features from input images along with U-Net like encoder and

Table 2
Dataset information.

| Dataset | No. of Images |
|---------------------|---------------|
| Bacterial Pneumonia | 2538 |
| Virus Pneumonia | 1345 |
| Normal Category | 1346 |
| Total | 5229 |

decoder. Then, these features were fed to lung nodule recognition with faster R-CNN, while classification was done by Gradient Boosting Machine (GBM). Aria et al. [11] aimed to achieve fast screening through generating candidate's suspicious regions using a 3D fully CNN. Extensive data augmentations were used to train the ensemble 3D CNN from the obtained positive and negative patches, and different threshold and data augmentation types were used on false positive patches to train classifiers. The final results were produced by averaging the second stage network's output to come up with final prediction on pneumonia presence. Qiao et al. [12] considered hue, saturation and brightness in the X-ray image as image descriptors using spatial distribution and diseased lung tissues were detected by combining ANN with Moth-Flame and Ant Lion Optimization algorithms. Daniel et al. [13] used CNN model based on inception V3 model to classify various kinds of pneumonia infection on pediatrics patient. Inception V3 model classifies the viral and bacterial pneumonia using extracted fixed features and use AUC as a measuring tool. Anuja kumar et al. [14] proposed a deep Siamese based neural network for automatic detection of pneumonia by analyzing the amount of white substance presence on both the right and the left chest of X-ray image. Kaggle dataset was used for the study to train and validate the deep Siamese neural network and the results presented with reduced bias and improved generalization. Paras et al. [15] considered AlexNet and GoogleNet pre-trained data models along with data augmentation and achieved the value of 0.94–0.95 AUC (Area Under the Curve). On another study, they achieved AUC value of 0.99 using modified two-network ensemble architecture. Tawsifur et al. [16] stated that the highest accuracy achieved by the aforementioned literature studies in identifying or classifying normal patients and

pneumonia infected patients by taking X-ray images as input, the values lied between 96.84% and 93.6% respectively.

Further, Pranav et al. [17] created a ChXNet model to detect pneumonia on ChestX-ray14 dataset. This model can detect all the 14 diseases in ChestX-ray14 and which is achieved good accuracy. Jonathan et al. [18] developed a dual convolutional neural network that takes input as frontal and lateral chest X-ray images for the detection of pneumonia. Marios et al. [19] created a CNN model which consists of five convolutional layer, LeakyReLU activations, average pooling layer and three dense layers. This model was developed to identify interstitial lung disease patterns. Another author Mark et al. [20] proposed a CNN model to detect and classify the abnormalities in frontal chest radiographs. Rachna et al. [21] created a pneumonia detection model using CNN and transfer learning. This model used the pre-trained models such as VGG16, VGG19, ResNet50 and Inception-v3. One more author Abhir et al. [22] built a modified AlexNet with support vector machine to detect lung abnormality from chest X-ray images and CT images.

From the literature, it is perceived that the many of the works are done with unbalanced dataset. In addition, many authors worked with small amount images and their results are in the acceptable level. Thus, still exists room to come up with much improved deep learning algorithms and methodologies which could generate greater results. The summary of the pneumonia detection methods for chest X-ray images and CT images is shown in Table 1.

3. Dataset

The proposed work was tested on publically available database. This database contains 5,232 chest x-ray images of RGB channels, which includes 3883 pneumonia images and 1346 normal images of size 227 × 227 and it details presented in Table 2. Chest X-ray images were selected from retrospective cohorts of pediatric patients of one to five years old from Guangzhou Women and Children's Medical Center, Guangzhou.

The dataset is organized into 3 folders for training, validation and testing respectively. It also contains subfolders for each image category i.e. pneumonia and normal images. Fig. 1 shows the sample normal and pneumonia images present in the dataset. As can be seen, to the naked

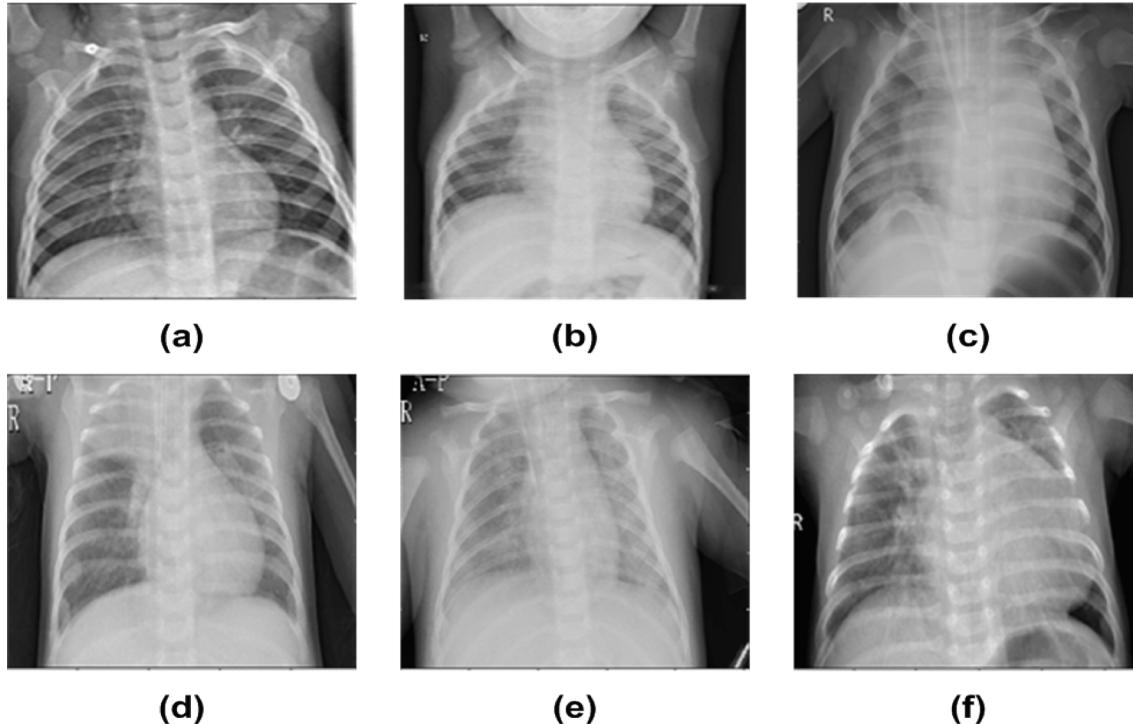


Fig. 1. Sample images with and without Pneumonia ((a)-(b) Normal images, (c)-(d) Bacterial pneumonia images, (e)-(f) Viral pneumonia images).

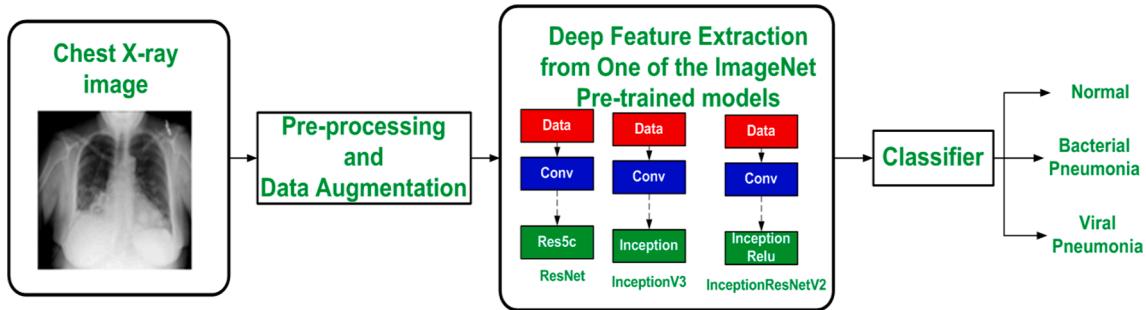


Fig. 2. The work flow of the proposed work.

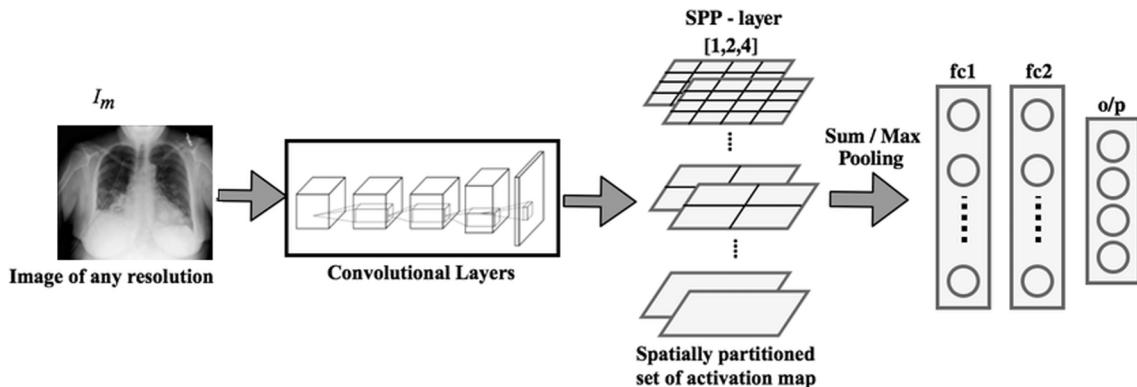


Fig. 3. Convolutional Neural Network Architecture.

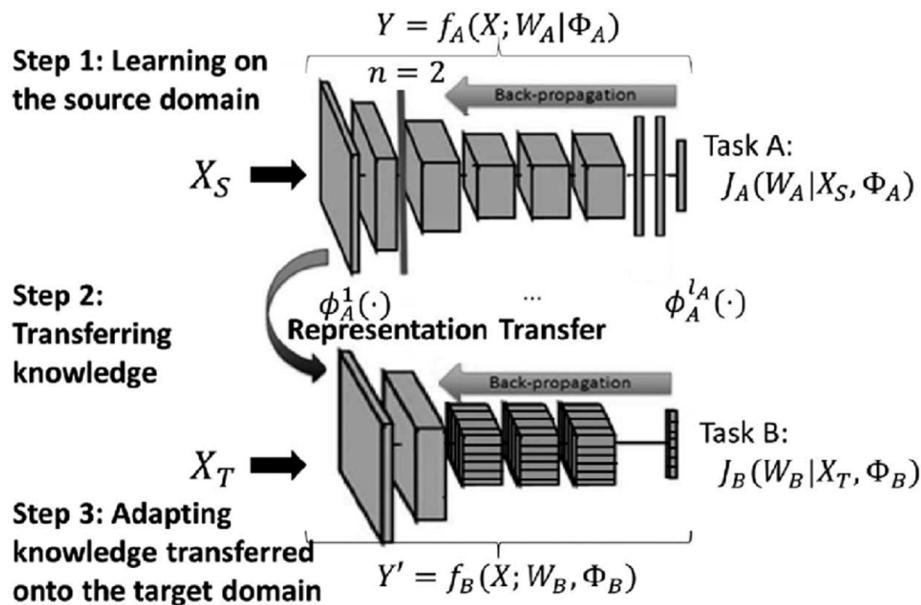


Fig. 4. General Architecture in Transfer Learning.

eye it is very difficult to tell the differences between these two images.

4. Methodology

In this section the theoretical background required for the proposed work explained. Next, CNN architecture and model architecture briefly explained. The different stages of the proposed work are illustrated in Fig. 2.

4.1. Theoretical background

This section of the paper discusses about the concepts of Convolutional Neural Networks (CNN) and Transfer Learning (TL).

4.1.1. Convolutional neural networks

In recent years, numerous researchers and studies have employed different Artificial Intelligence (AI) based solutions to address various medical problems and Convolutional Neural Networks (CNN) have been

successful in addressing a wide range of diseases such as brain tumor detection, breast cancer detection and disease classification in X-ray images, etc. [23]. The generic architecture of CNN is depicted in Fig. 3.

Convolutional neural networks assumes that an type of input to it is an image and modern deep learning models in computer vision prefers CNN as it detects low-level features in an image like edges and it also captures the spatial and temporal dependencies through the use of filters [24]. Further, CNN also reduces computation time by using a weight-sharing technique with lower number of parameters and each layer consists of kernels or filters. Pooling layers helps to extract dominant features that are positional and rotational invariant by reducing the spatial size of representation generated by previous kernels after convolution. These layers are included between two convolution layers in common and the most widely used pooling layer is max pooling layer which “separates input into squares of a given size, and outputs the maximum value of each square”. In addition, average of each square is calculated y average pooling layer. Both these methods are very effective in reducing the dimensionality and computation efforts. Considering the achievements and improvements of CNN in recognizing pneumonia, the performance of CNN is much better on larger datasets whereas CNN failed to produce expected results on smaller datasets unless proper care is taken [25].

4.1.2. Transfer learning

Many researchers have suggested that pre-trained models can be considered while addressing a problem related to computer vision which could yield better results instead of using the long process of training models from scratch. Transfer learning is a method of transferring learning from one predefined and trained model to some new domain by reusing the network layer weights and many studies have found that this method have produced better results in computer vision and other areas of research as well [26]. The steps involved in the transfer learning with mathematical notation is defined as

Step 1: Learning on the source domain : A source Domain $Y = f_A(X; W_A | \varphi_A)$ is defined by two parts; where X is a feature space; and marginal partial distribution f_A of A . Given a specific domain Y , task A is $J_A(W_A | X_S, \varphi_A)$; where J_A is predictive function.

Step 2: Transferring knowledge : In the study, there are three transfer learning knowledge utilized by the researchers are ConvNet as fixed feature extractor, Fine-tuning the ConvNet and Pretrained models and it can be represented as $\varphi_A^1(\cdot)$.

Step 3: Adapting knowledge transferred onto the target domain : Given a source Domain Y and it learning task J_A , a target domain Y' and learning task J_B , transfer learning aims to improve the learning.

Fig. 4 represents the general architecture of transfer learning working process where learning was done from a pre-trained source domain and the knowledge has been transferred through representation transfer to the targeting domain [27].

4.2. Pre-processing

To segment the pneumonia presence on the input chest X-ray images, U-Net architecture has been used in this paper, because U-Net architecture recently shown remarkable results on medical image segmentation applications. The U-Net architecture works with the convolution neural network model and form the pixel based segmentation than conventional models. Hence, this architecture can accurately segment the required target features effectively from the chest x-ray images.

The U-Net architecture consist of contracting path and an expansive path. The contracting path hold 4 encoding blocks and expansive path hold 4 decoding blocks. In contracting path, each encoding block consists of two 3×3 convolution layers followed by a ReLU activation function and a 2×2 max-pooling operation with stride 2 for down-sampling. During the downampling, the spatial information is reduced and number of feature channels are doubled. In expansive path each step

perform the upsampling on feature map and it is achieved by decoding blocks. . The decoding blocks consist of the transposed convolution layers for upsampling followed by 2×2 convolution, a concatenation with the corresponding feature map from the contracting path, and two 3×3 convolution layers, each followed by a ReLU function. At the last layer, 1×1 convolution is used to map the feature vector to the 2 number of class (lung or not lung) with softMax function [41]. In total network has U-Net architecture has 23 convolutional layers.

4.3. CNN architecture

CNN architecture is shown in Fig. 3. CNN consists of four layers for processing, namely, convolutional layer, pooling layer, flattening layer, and the fully-connected layer. The following subsection gives a comprehensive explanation of every layer.

4.3.1. Convolutional layer

The convolutional (CONV) layer is the heart building block of CNN that converts the image into matrix form on which convolution operation (performs a dot product) performed with a set of defined learnable filter parameters. In this paper, 3×3 filter/kernel/mask is used to extract the feature map from the input matrix to reduce the size of the image which eases the processing function. Though, the CONV operation might lead to loss of some information but retains all the integral part of the image portions as feature map. Multiple filters are applied into input matrix to obtain a layer of feature map. Further, pooling, flattening and fully-connected layers are applied on this layer to generate the complete CNN.

4.3.2. Activation function

Rectified Linear (ReLU) function is a category of linear function that is applied to the convolutional layer as an active function which returns one if the input is positive and returns zero, otherwise. In neural network models ReLU function works better than the other activation functions such as sigmoid or hyperbolic tangent activation function and rectifies vanishing gradient problem. The ReLU function is defined in Eq. (1).

$$T(a) = \max(0, a) \quad (1)$$

4.3.3. Pooling layer

The pooling layer makes the representations smaller and operates over each activation map to reduce the dimensions by using sub-sampling technique. The resulting sub-sampling pixels will not change the object and also reduces the computational complexity. The sub-sampling techniques used in this work, are max-pooling and average pooling. Max-pooling is a discrete process; the polling layer of 2×2 window is applied on the feature map and selects the highest value from the window covered on feature map of the image. Max-pooling provides the salient features in the image. Average pooling works similar to max-pooling but it considers the highest pixel value and calculates the average of the 2×2 window covered on the feature map of the image. Average pooling retains more amount of information from an image than the max-pooling.

4.3.4. Flattening and Fully-connected layer

The pooled feature map was taken as input to the flattening and fully-connected layer. Here, the pooled features are converted into a single column just to feed features into the neural network after which the flattening layer feature map fed into the fully connected layer. In a fully connected layer, input forward propagates calculating weights and makes a prediction in network. Depending on the prediction, we calculate a cost function that tells that how a network is performing.

4.3.5. Reducing Over-fitting

To reduce the over-fitting dropout, we used image augmentation in models ResNet50, InceptionV3 and InceptionResNetV2. Further, we

Table 3
Image Augmentation Parameters.

| Transformation | Parameters |
|-------------------|-------------------------------|
| Rotation | Angle between -20 and $+20$ |
| Horizontal Flip | Randomness = 50% |
| Vertical Flip | Randomness = 50% |
| Horizontal Shift | Up to 12 pixels |
| Vertical Shift | Up to 12 pixels |
| Gaussian Blurring | Up to $\sigma = 1$ |

used model checkpoint while training the neural network and early stopping when the validation loss started increasing. The following subsection explains the dropout, image augmentation, model checkpoint and early stopping techniques and its procedure.

4.3.5.1. Dropout. Dropouts are a very common regularization technique that is used for preventing neural networks from over-fitting. Other regularization methods like L1 and L2 reduces over-fitting by adding a penalty term to the cost function while dropouts modify the network itself for each iteration. For every iteration, it randomly drops neurons from the neural networks during the training which is equivalent to training an ensemble of neural networks for every iterations. The different networks will over-fit in different ways, so the net effect of dropout will be to reduce over-fitting.

4.3.5.2. Image augmentation techniques. The image augmentation algorithm was implemented using keras library in python. The scaling, cropping, rotation, width change (horizontal flip), height change (vertical flip), padding, rotation, affine transformation, brightness, contrast, saturation, hue and filling operations were performed for normal class images. The parameters used for the image augmentation is given in Table 3. The sample resultant augmentation image is shown in Fig. 5.

After the image augmentation, the number of normal samples in the dataset was increased from 1,346 to 4,266. The augmentation used to balance the images in the classes and this allow us to train all the images rather than picking random images for training, and it also increases the accuracy of the training and positively affects the classification results.

4.3.5.3. Model checkpoint. To obtain better results, more number of iterations needs to be performed during training, which could possibly

consume a lot of time. To tackle this issue, copy of the best performing model was saved till an epoch that improves the metrics ends. We adopted model checkpoint to achieve this goal in our work.

4.3.5.4. Early stopping. Sometimes, during training we can notice that the generalization gap (i.e. the difference between training and validation error) starts to increase, instead of decreasing. This is a symptom of over-fitting that can be solved in many ways (reducing model capacity, increasing training data, data augmentation, regularization, dropout, etc). Often a practical and efficient solution to address this problem is to stop training earlier when the generalization gap is getting worse.

4.3.6. Optimization techniques

Optimization algorithm plays a vital role in minimizing the error rate while training a machine learning model and the efficiency of an optimizer is determined by the following metrics in general: speed of convergence and generalization. The most commonly used optimizers are Adaptive Moment Estimation (Adam) and Stochastic Gradient Descent (SGD). Both of these optimizers is capable of covering one or the other metric when employed. Basically, Adam combines the advantages of two SGD extensions-Root Mean Square Propagation (RMSProp) and Adaptive Gradient Algorithm (AdaGrad) which computes individual adaptive learning rates for different parameters. Despite widespread popularity of Adam, recent research studies have proved that Adam can fail “to converge to an optimal solution under specific settings” [28].

Liangchen et al. [29] proposed an AdaBound, a novel optimizer that is well defined and structured, which can train machine learning models “as fast as Adam and as good as SGD”. It also converges to SGD in the infinite limit with faster convergence and better generalization. In this study, we used both the optimizers and analysed them separately.

4.4. Model architecture

Three models were trained and tested on the chest-x ray images. The descriptions of these three models are given below.

4.4.1. ResNet50 architecture

Computer vision and deep learning world sees Residual Network (ResNet) as a gift because it made possible to train ultra deep neural

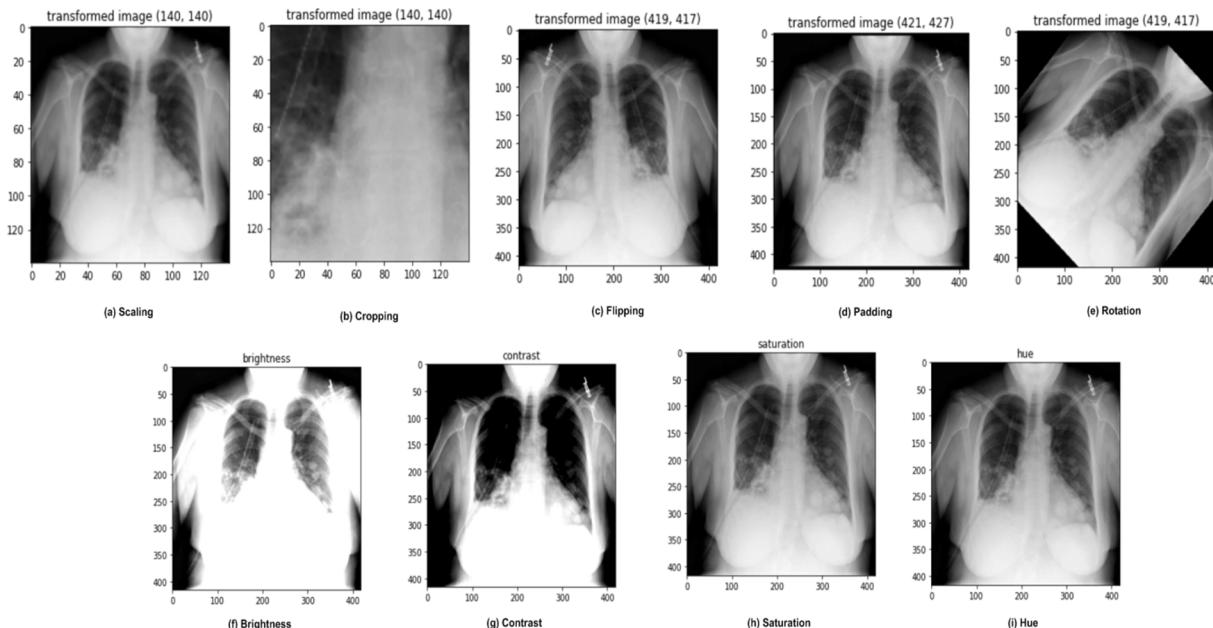


Fig. 5. Sample resultant augmentation image.

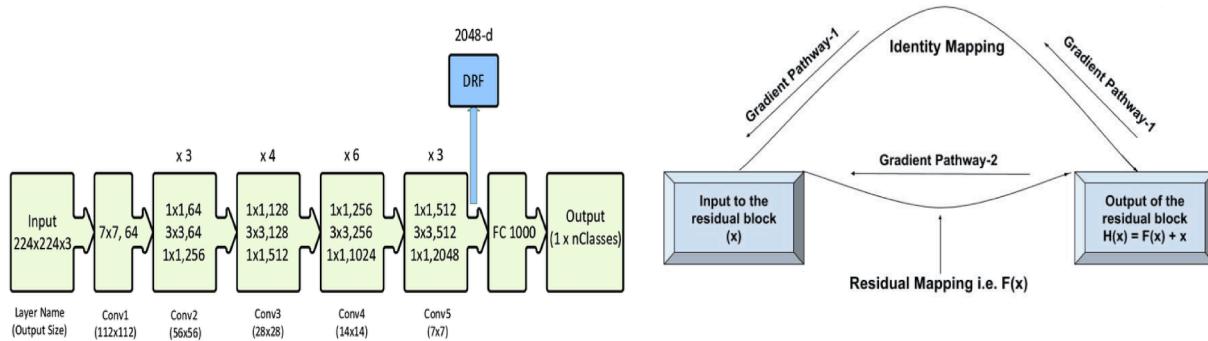


Fig. 6. RestNet50 architecture and Single Residual Block.

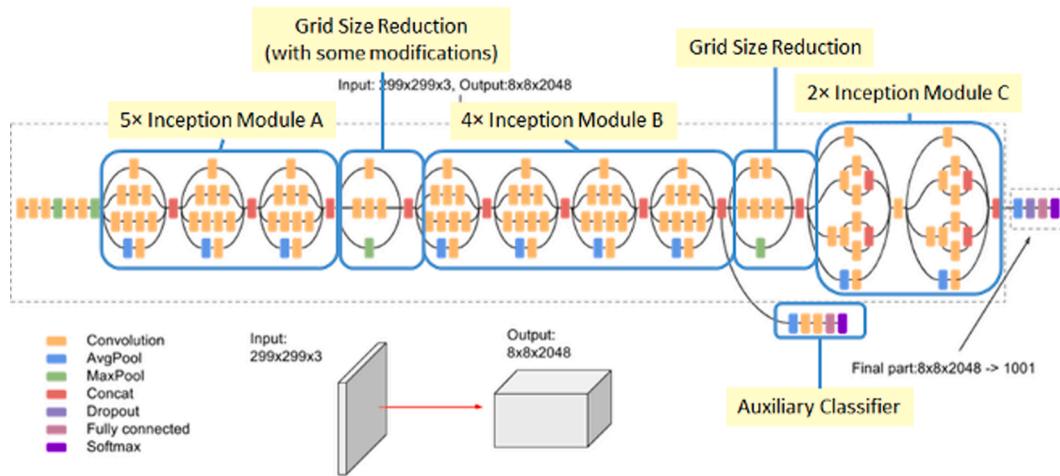


Fig. 7. InceptionV3 Architecture.

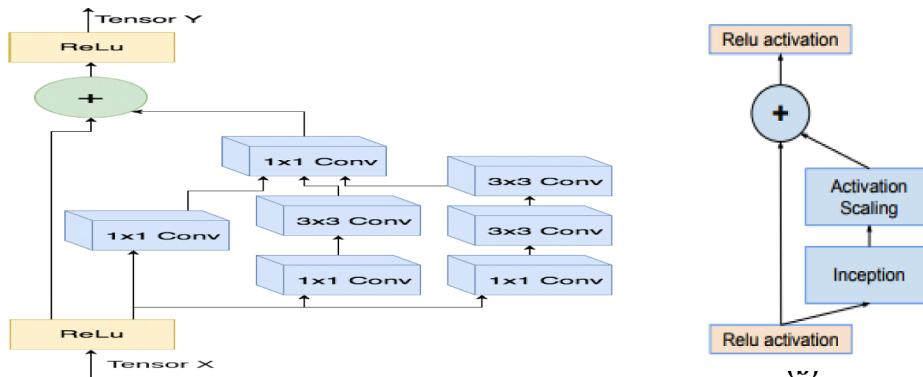


Fig. 8. InceptionResNet V2 Architecture (a) Residual Inception Block (b) Scaling of Residuals.

networks that can contain hundreds or thousands of layers and still offers better performance than AlexNet. The main advantage of using the ResNet is: short connections or skip connections, which allows the network to rectify the vanishing gradient and degrading accuracy [9]. In addition, reduces the training error and helps the network to converge faster. The architecture of ResNet50 is shown in Fig. 6.

ResNet consists of five stages where each stage contains convolution and identity block and each of which has 3 convolution layers. Moreover, it holds more than 23 million trainable parameters. To reduce errors and computational time in check, shortcut identity mapping was introduced. It explicitly let the layers fit a residual mapping and denoted that as $H(x)$ and it let the non linear layers fit another mapping $F(x) := H(x) - x$. Therefore, the original mapping becomes $H(x) := F(x) + xH(x) := F(x) + x$ as can be seen in Fig. 6 [30].

(x) + x as can be seen in Fig. 6 [30].

4.4.2. Inception V3 architecture

The inception V3 model works as a multi-level feature generator for increasing depth and width of the deep learning network by computing 1 X 1, 3 X 3 and 5 X 5 convolutions. In general, this model achieves great performance by using all kinds of kernels on the image and to also get results from all those kernels [38].

With 42 layers deep, the computation cost is only about 2.5 higher than that of GoogleNet, and much more efficient than that of VGGNet. Inception V3 is mainly focusing on reducing the number of connections/parameters with Factorization Convolutions without decreasing network efficiency. In Inception V3, Auxiliary Classifier is used as

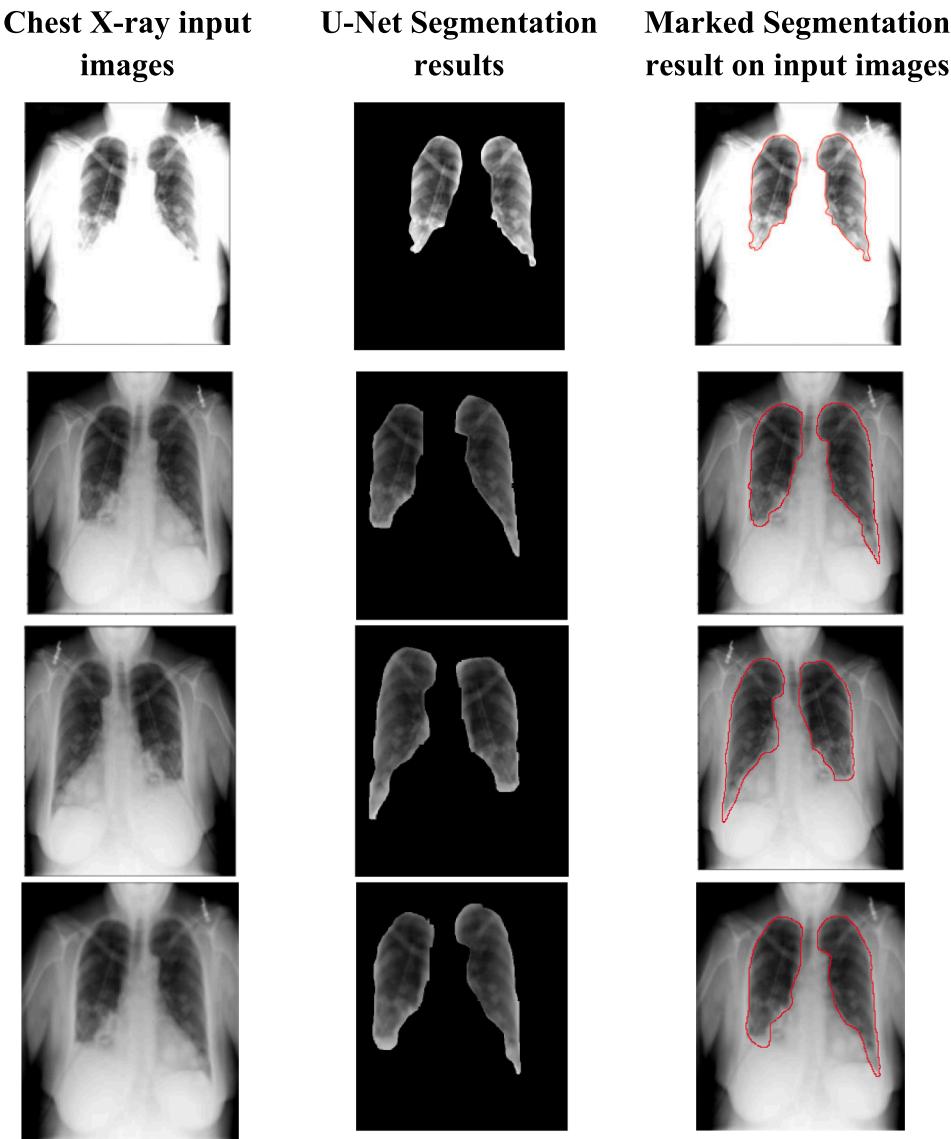


Fig. 9. Sample segmentation results: Chest x-ray images (Left), U-Net Architecture segmented results (Middle), Marked Segmentation result on input images (right).

Table 4
Comparison of segmentation methods using IOC measure.

| Method | IOC Score |
|-------------------------|-----------|
| U-Net | 0.93 |
| Active Shape Model [43] | 0.89 |
| Thresholding [44] | 0.74 |

Table 5
Parameter values of the proposed model used in this study.

| CNN architecture | Optimization | Momentum | Mini batch | Learning rate |
|--------------------|--------------|----------|------------|---------------|
| ResNet50 | SGD | 0.9 | 16 | 0.001 |
| Inception V3 | SGD | 0.9 | 16 | 0.001 |
| Inception ResNetV2 | SGD | 0.9 | 16 | 0.001 |

regularizer with efficient grid size reduction. Label smoothing is used for regularization to prevent the largest logit from becoming much larger than all others. These features make Inception V3 as less expensive and

efficient. The architecture of InceptionV3 is shown in Fig. 7.

4.4.3. InceptionResNetV2 architecture

Both the Inception and Residual network are SOTA architectures which have shown very good performance with relatively low computational cost. These two networks can be combined together to boost the performance up further and it is shown in Fig. 8. InceptionResNetV2 is a convolutional neural network that is trained on more than a million images from the ImageNet database and it is 164 layers deep and can classify images into 1000 object categories. Therefore, this network has learned rich feature representations for a wide range of images. The image input size of this network is 299 by 299 [39].

This architecture uses batch-normalization for scaling up the dimensionality of the filter bank and the same is applied only on the top of the traditional layers, but not on top of the summations. To prevent the network from dying early in the training, the scaling factors between 0.1 and 0.3 were picked in InceptionResNet V2 since neither by lowering the learning rate nor by adding an extra batch normalization would not help in prolonging the network life.

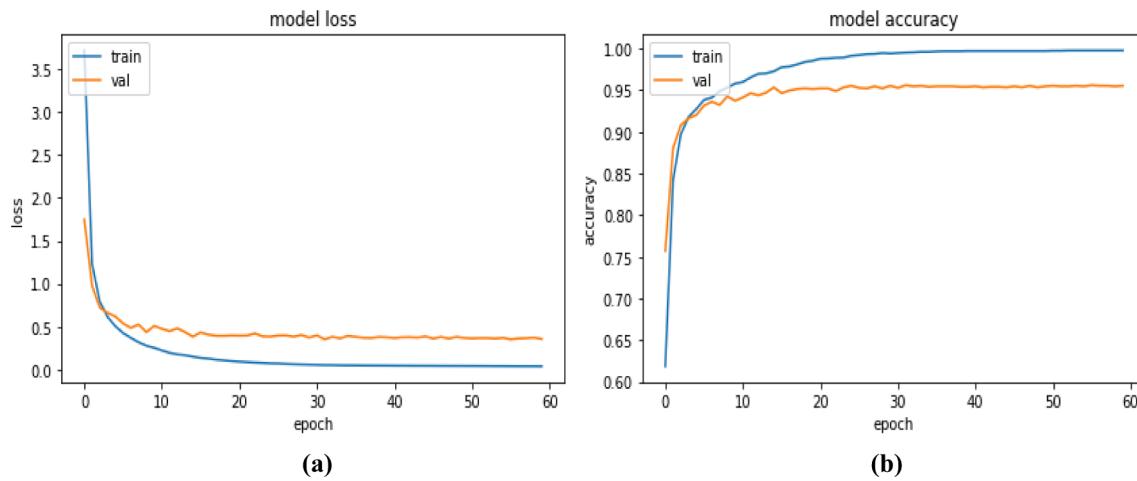


Fig. 10. (a) Cross entropy loss against epoch and (b) Accuracy against epoch.

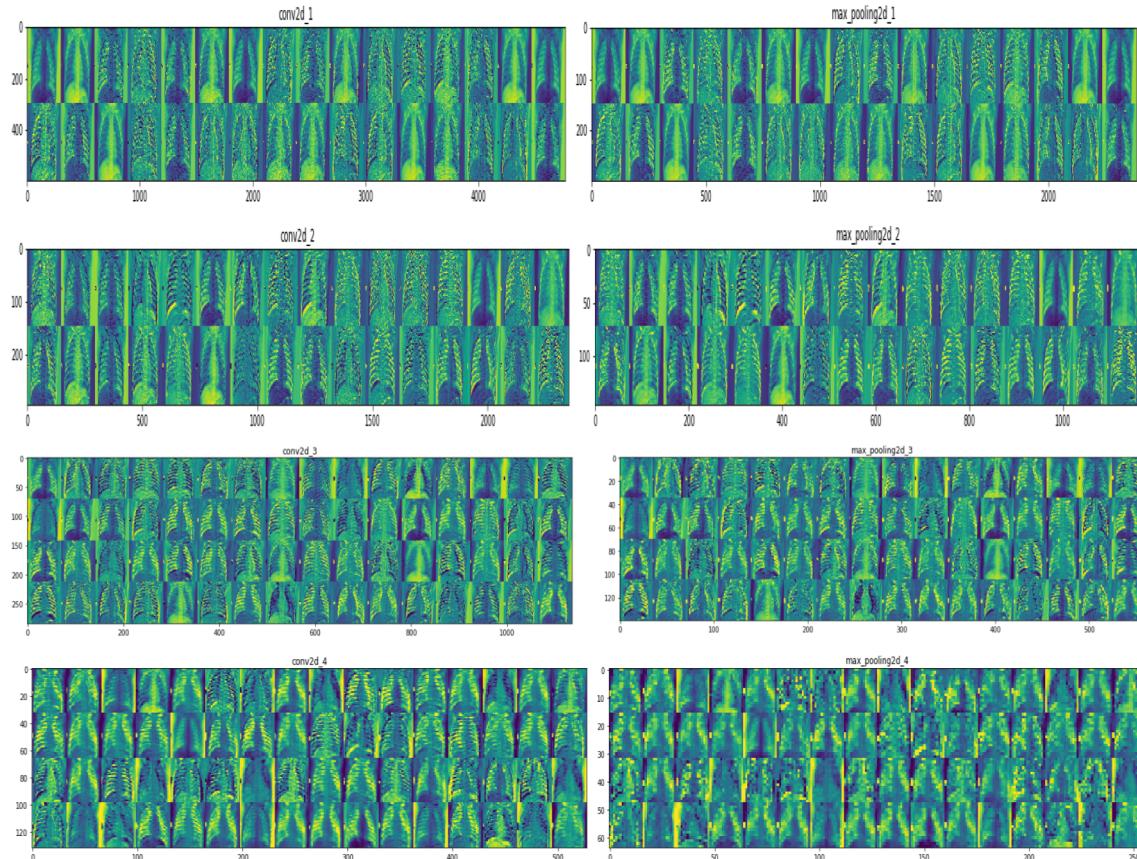


Fig. 11. Visualization of feature maps of first four convolutional layer and max pooling layer.

Table 6

Comparison of different learning rate result on three models using the evaluation metrics.

| Model | Learning Rate | Accuracy (%) | Precision(%) | Recall (%) | F1 score(%) | ROC-AUC |
|-------------------|---------------|--------------|--------------|------------|-------------|---------|
| ResNet50 | 0.001 | 92.74 | 88.35 | 95.82 | 91.93 | 0.92 |
| | 0.0001 | 93.06 | 88.97 | 96.78 | 92.71 | 0.93 |
| InceptionV3 | 0.001 | 92.15 | 88.36 | 91.87 | 90.08 | 0.90 |
| | 0.0001 | 92.67 | 88.70 | 92.70 | 90.65 | 0.91 |
| InceptionResNetV2 | 0.001 | 90.87 | 88.42 | 92.85 | 90.58 | 0.91 |
| | 0.0001 | 92.40 | 88.88 | 93.20 | 90.98 | 0.91 |

Table 7

Comparison of different batch size result on three models using the evaluation metrics.

| Model | Batch Size | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|-------------------|------------|--------------|---------------|------------|--------------|
| ResNet50 | 16 | 93.06 | 88.97 | 96.78 | 92.71 |
| | 32 | 91.25 | 87.25 | 95.36 | 91.12 |
| InceptionV3 | 16 | 92.67 | 88.70 | 92.70 | 90.65 |
| | 32 | 90.89 | 87.14 | 91.14 | 89.09 |
| InceptionResNetV2 | 16 | 92.40 | 88.88 | 93.20 | 90.98 |
| | 32 | 90.65 | 87.74 | 92.24 | 89.93 |

5. Experimental results

This section presents the experimental results of the lung segmentation and classification on the chest x-ray images. The work

experimentation was done on python environment with graphics card enabled 64-bit windows 10 operating system.

5.1. Lung segmentation results

The sample segmentation results obtained from the U-Net architecture are shown in Fig. 9. It contains four rows of the chest X-ray images, segmented results and bounding box imposed on the original input image. The first row of Fig. 9 shows the sample results of augmented normal case image. Similarly, the remaining rows are represents the sample results of pneumonia image. The segmented lung areas are marked with red color. It is observed from Fig. 9 that the lungs are accurately detected and remaining portions are completely suppressed during U-Net architecture based pre-processing. The performance of the segmentation method is done using measure Jaccard index. The Jaccard index measures the similarity between a limited number of sets. It is also called as Intersection Over Union (IOC). IOC is a statistic measure used

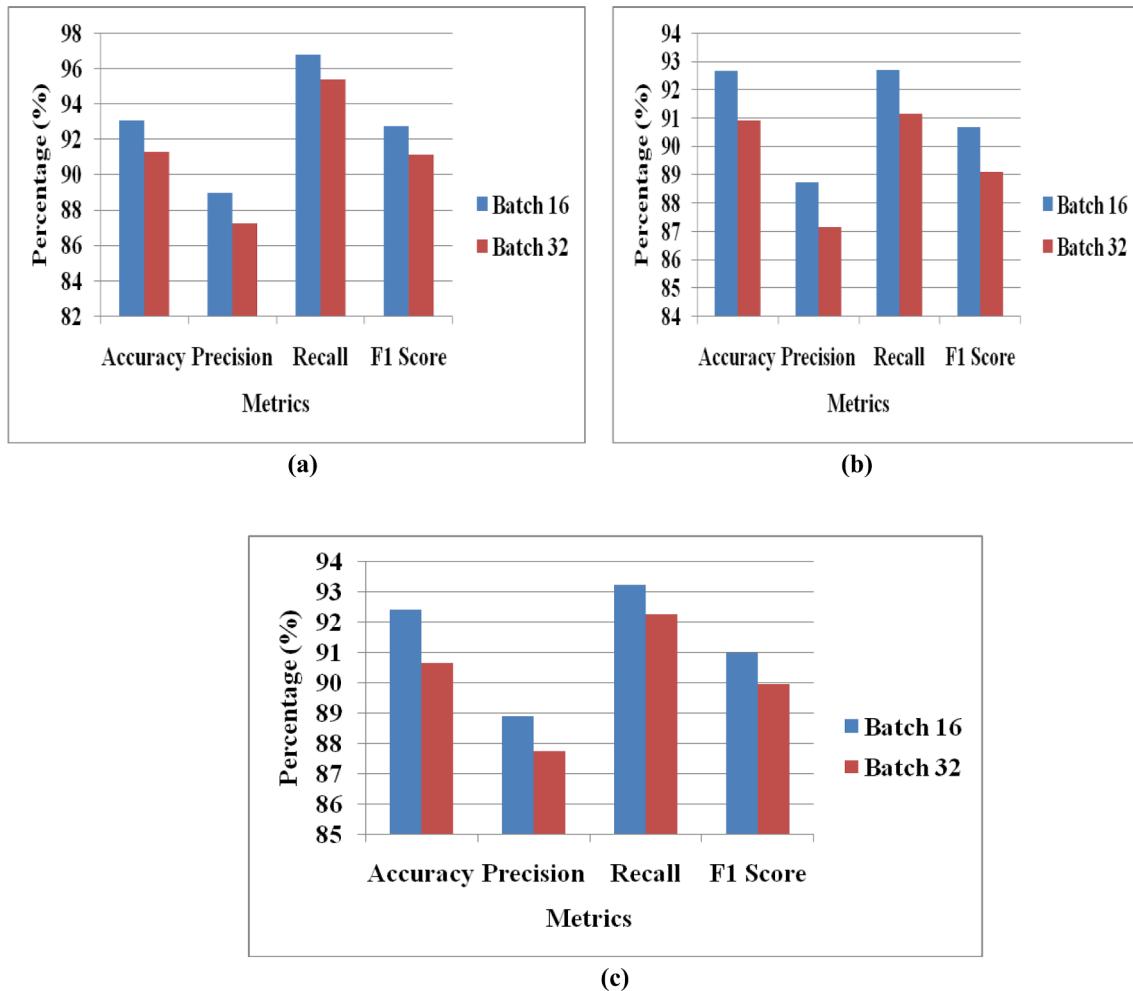


Fig. 12. Batch analysis (a) ResNet50 (b) Inception V3 and (c) InceptionResNet V2.

Table 8

Comparison of different optimizer result on three models using the evaluation metrics.

| Model | Optimizer | Accuracy(%) | Precision(%) | Recall(%) | F1 score(%) | ROC-AUC |
|-------------------|-----------|-------------|--------------|-----------|-------------|---------|
| ResNet50 | Adam | 91.03 | 89.76 | 96.67 | 93.09 | 0.93 |
| | SGD | 93.06 | 88.97 | 96.78 | 92.71 | 0.93 |
| InceptionV3 | Adam | 90.84 | 89.46 | 96.14 | 92.67 | 0.93 |
| | SGD | 92.67 | 88.70 | 92.70 | 90.65 | 0.91 |
| InceptionResNetV2 | Adam | 90.68 | 90.27 | 96.27 | 93.17 | 0.93 |
| | SGD | 92.40 | 88.88 | 93.20 | 90.98 | 0.91 |

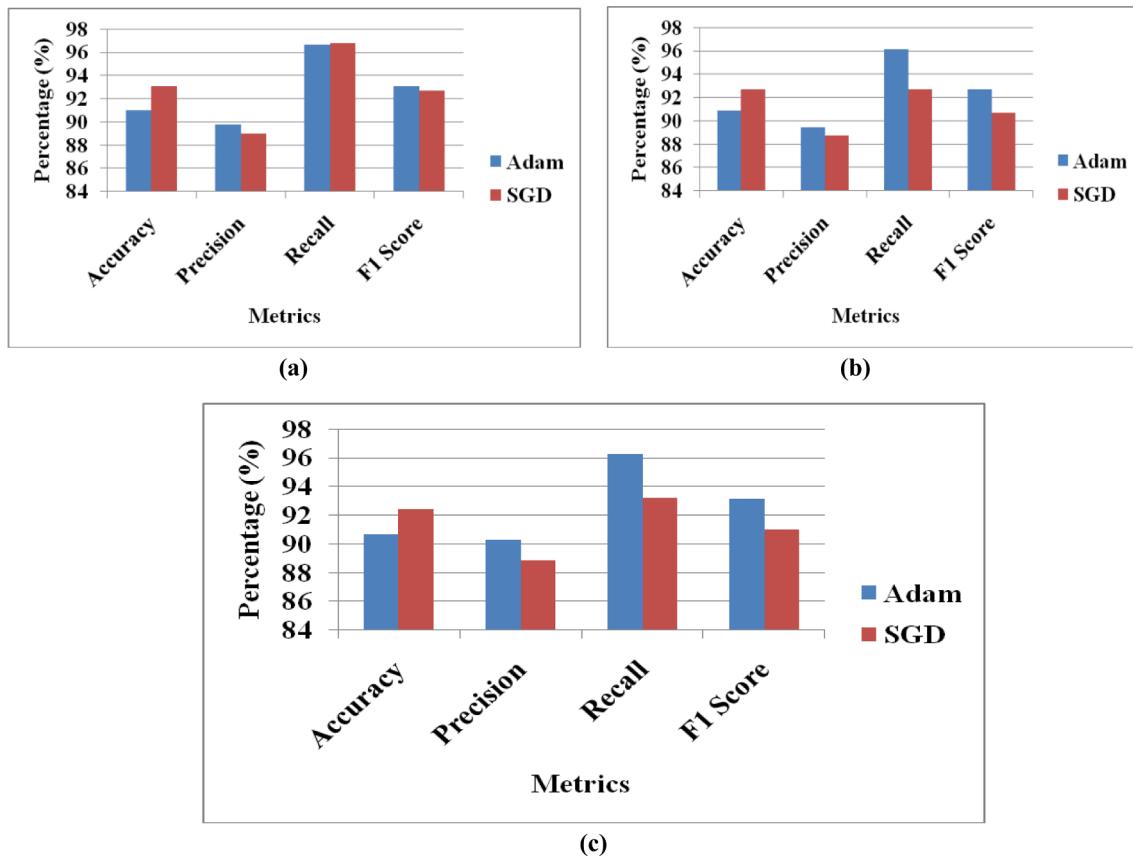


Fig. 13. Effect of varying optimizers (a) ResNet50 (b) InceptionV3 and (c) InceptionResNet V2.

Table 9
Comparison of proposed method using the evaluation metrics.

| Metrics | InceptionV3 | InceptionResNetV2 | ResNet50 |
|----------------|-------------|-------------------|----------|
| True Positive | 267 | 288 | 331 |
| True Negative | 428 | 405 | 367 |
| False Positive | 34 | 36 | 41 |
| False Negative | 21 | 21 | 11 |
| Accuracy(%) | 92.67 | 92.40 | 93.06 |
| Precision(%) | 88.70 | 88.88 | 88.97 |
| Recall(%) | 92.70 | 93.20 | 96.78 |
| F1 Score(%) | 90.65 | 90.98 | 92.71 |

for comparing the similarity and diversity measure between two sets A and B. It can be defined in Eq. (2).

$$IOC(A, B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (2)$$

Table 4 shows the comparison result of segmentation methods using IOC measure. From the **Table 4**, it is observed that the U-Net convolution neural network outperformed than the existing conventional lung segmentation methods.

5.2. Lung classification results

For the experimentation of lung classification, pneumonia images

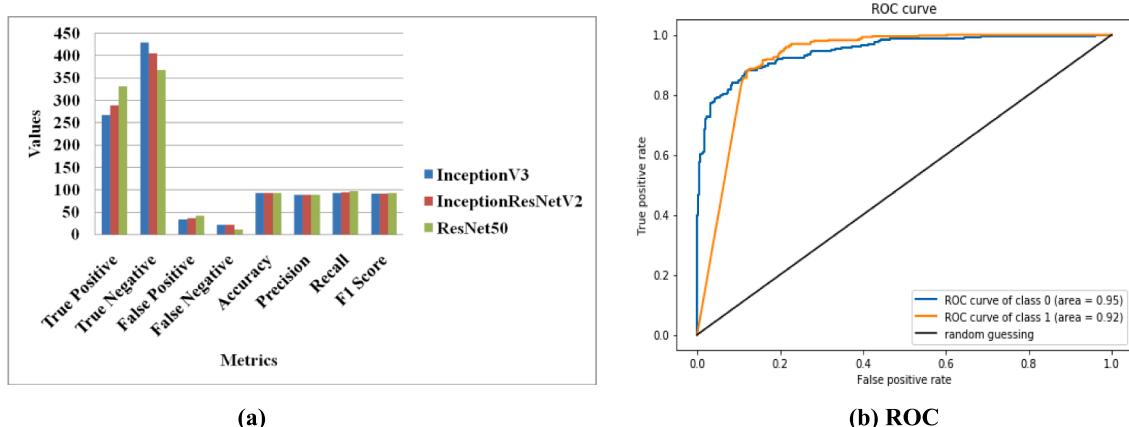


Fig. 14. (a) Performance analysis of proposed method using various evaluation metrics and (b) ROC.

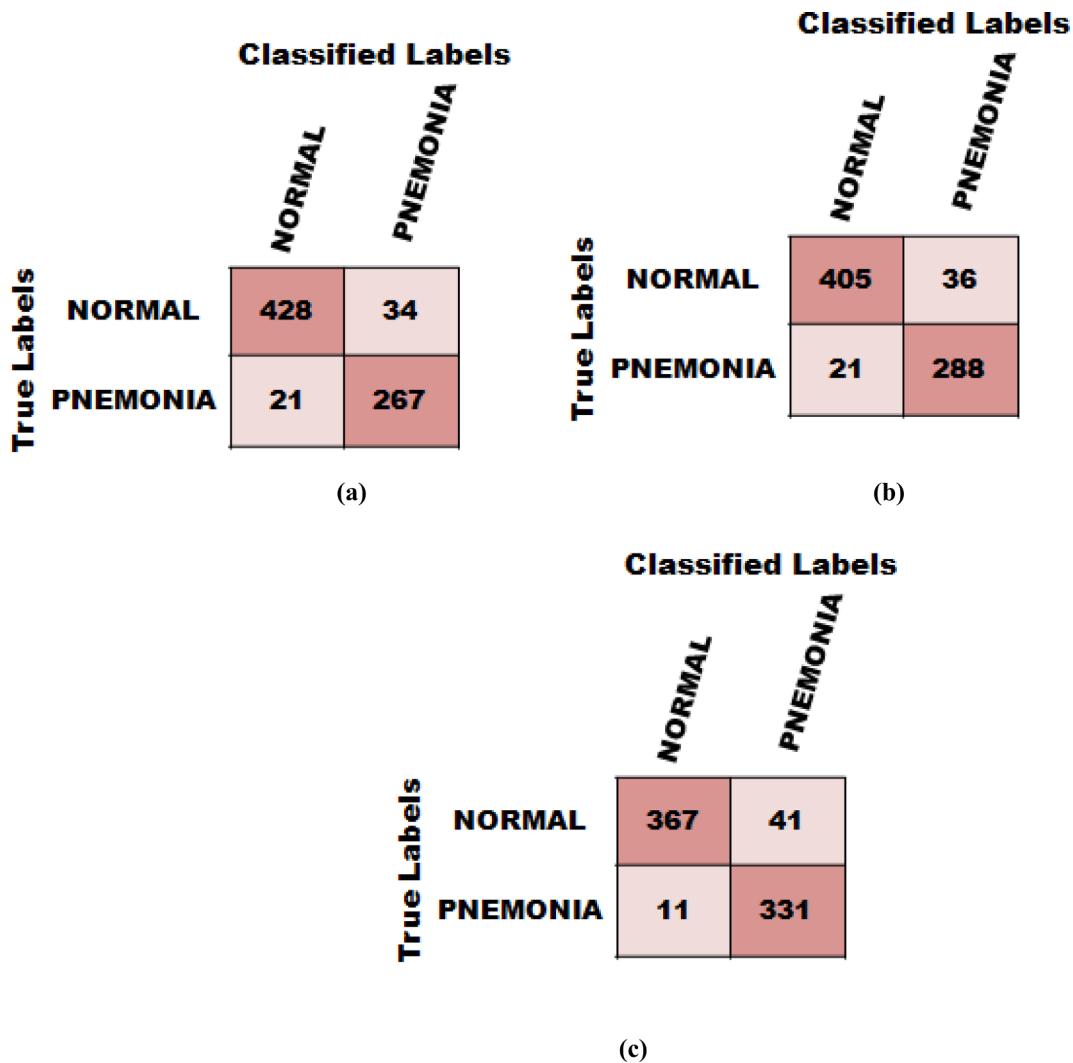


Fig. 15. Confusion matrices (a) InceptionV3 (b) InceptionResNetV2 (c) ResNet50.

were divided into three parts: 70% training set, validation set 15% and 15% testing set. In recent years, transfer learning method is widely adopted pre-trained deep learning model to address new problems in computer vision and image processing. The value of pre-trained model used in the experiment is given in Table 5. To achieve good results, we have prepared and trained all the three models of proposed methodology separately. For training, the Adam and SGD (Stochastic Gradient Descent) optimizers were used. The initial learning rate was set to 0.001 and reduced by 2 for every three epochs. Early stopping method was used to avoid over-fitting i.e. stop training when generalization gap is getting worse. Fig. 10 depicts accuracy and loss against epochs. From Fig. 10 inferred that the training models were completed with high accuracy.

In order to evaluate the proposed model, the performance measures such as precision, recall, specificity, accuracy and F1 score was considered. The formulas for the measures are defined in Eqs. (3) to (6). Precision is a measure of fidelity and a good measure to detect the false positive i.e. the chest x-ray images which are not having bacterial pneumonia and virus pneumonia infection. High precision value indicates the low number of false positives. Recall is widely used in many applications to detect the pneumonia which is a measure of completeness. High recall value corresponds to high detection rate. Accuracy is the best measure for validation or classification problem. F1-score is a measure to test the accuracy [37]. It uses both precision and recall to compute the score.

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \quad (4)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \times 100\% \quad (5)$$

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (6)$$

From Eqs. (3) to (6), TP, TN, FP and FN indicate true positives, true negatives, false positives and false Negatives respectively. The following sections discusses about histogram of pixel intensities, visual feature maps, experimental analysis of the three models and comparative study with the existing models presented in the paper.

5.2.1. Histogram of pixel intensities

In deep learning model, pixels are playing the major role to extract the feature maps. Histogram is a graphical representation of an image which depicts the number of pixels in an image at different intensity value found in that image. It processes the source image in a single pass and a running count of the number of pixels found at each intensity value is kept which will be used to construct a suitable histogram.

Table 10

Comparison of proposed method with the existing methods using accuracy metrics.

| Methodology | Topic | Accuracy |
|--|---|----------|
| Neuro-heuristic approach [12] | A neuro-heuristic approach for recognition of lung diseases from X-ray images | 79.06% |
| Different Pre-trained CNN [16] | Transfer Learning with Deep Convolutional Neural Network for Pneumonia Detection using Chest X-ray | 93.3% |
| VGG16 [21] | Pneumonia Detection in chest X-ray images using Convolutional Neural Networks and Transfer Learning | 87.18% |
| VGG19 [21] | Pneumonia Detection in chest X-ray images using Convolutional Neural Networks and Transfer Learning | 88.46% |
| CNN Models along with DenseNet-169 and SVM [1] | Pneumonia Detection Using CNN based Feature Extraction | 80.02% |
| VGG16 and CNN [31] | A transfer learning method with deep residual network for pediatric pneumonia diagnosis | 74.2% |
| RetinaNet + Mask RCNN [32] | Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database | 75.8% |
| VGG16 and Xception [33] | Diagnosis of Pneumonia from Chest X-Ray Images using Deep Learning | 87% |
| Fully connected RCNN [34] | Chest X-ray Image Classification Using Faster R-CNN | 62% |
| MobileNet + AEO [42] | A Novel Method for Detection of Tuberculosis in Chest Radiographs Using Artificial Ecosystem-Based Optimisation of Deep Neural Network Features | 90.20% |
| Proposed - ResNet50 | — | 93.06% |
| Proposed-InceptionV3 | — | 92.97% |
| Proposed-InceptionResNetV2 | — | 92.40% |

5.2.2. Visualization of feature maps

Basically in deep learning, feature maps are playing the vital role in detection of pneumonia on chest X-ray images. Hence, the visualization of feature maps explained in this section. Feature maps helps the researchers to identify how the model is learning at every layer. As the depth of model increases is able to learn more spatial information. Additionally, visualization of feature maps is helps the researchers to do following things to improve the architecture:

- Fine tune the hyper parameter
- Finding error in the neural network
- Find the mistake in the algorithm
- Allow improving the overall design of the models.

Further, feature detectors to the input image to generate the activation maps or the feature maps using the Relu activation function for the first four convolutional and max pooling layer is shown in Fig. 11. Fig. 11 contains the feature maps assist the researchers to analyse the different features present in an image like edges, vertical and horizontal line, blending etc.

5.2.3. Experimental analysis of lung classification

For the experimental result evaluations, three classes of X-ray images such as normal patients, bacterial pneumonia and viral pneumonia were converted into two classes, namely, normal patients and infected patients (bacterial and viral pneumonia merged into one class). The evaluation metrics results at two fixed learning rates of 0.001 and 0.0001 for ResNet50, InceptionV3 and InceptionResNetV2 were implemented separately. The effect of using different learning rates on the evaluation metrics are shown in Table 6.

Next, we present the evaluation metrics results at two fixed batch size of 16 and 32 for ResNet50, InceptionV3 and InceptionResNetV2 separately. The reason for choosing these values was that the model was

over-fitting for lesser as well as greater than the values aforementioned. One of the most important hyper parameters to tune in deep learning is batch size since using a larger batch size to train the models allows computational speedups from the parallelism of GPUs. However, it is well known that larger the batch size, will likely lead to poor generalization. In fact, using a batch equal to the entire dataset guarantees convergence to the global optima of the objective function which is what we aim to achieve in this study. However, this happens at the cost of slower convergence to those global optima. Literature studies have shown that using smaller batch sizes will have faster convergence which produces good results. This is intuitively explained by the fact that smaller batch sizes allow the model to start learning before having to observe all the data. The problem in using a smaller batch size is that the model is not guaranteed to converge to the global optima. Therefore, it is often advised that one starts at a small batch size while using the benefits of faster training dynamics and growing the batch size slowly through training. The effect of using different batch size values on the evaluation metrics are shown in Table 7. To easily interpret the data present in Table 7, it converted as graphical form in Fig. 12.

Further, the evaluation metrics results for two different optimizers - Adam and Stochastic Gradient Descent (SGD) with momentum for InceptionV3, InceptionResNetV2 and ResNet50 are individually presented. The effect of using both these optimizers on the evaluation metrics is shown in Table 8. Fig. 13 shows the graphical representation of the Table 8.

Finally, we present our final results in Table 9 by classifying the end results as True Positive, True Negative, False Positive, False Negative Accuracy, Precision, Recall and F1 Score which is the weighted average for InceptionV3, InceptionResNetV2 and ResNet50 independently.

For our proposed work, classification model performance at various thresholds was evaluated by using Area underneath the Receiver Operating Characteristics (AUROC) and it is presented in Fig. 14. It measures the entire two dimensional area underneath the entire AUC curve from (0, 0) to (1, 1). It is most desirable by the researchers due to its scale invariant and classification threshold invariant features with some limitations. By analogy, higher the AUROC curve, better the model is at distinguishing between patients with disease and no disease. Fig. 14 (b) represents the ROC curve of this study and detection rates of undertaken pre-trained architectures.

Furthermore, The confusion matrices for the pneumonia classification by InceptionV3, InceptionResNetV2 and ResNet50 is shown in Fig. 15. From the Fig. 15 notice that the proposed method proves efficiency by correctly classify the pneumonia from chest x-ray images with the less number of false positives and false negatives.

5.2.4. Comparative analysis of classification results

In addition, the proposed method was compared with the previous research work on pneumonic classification problem and the results are presented in Table 10. From the Table 10, it is observed that the proposed method outperform than the some of existing methods. However, proposed method trained and tested with more number of sample images than the existing methods listed in Table 9. In future work, authors aim to improve the accuracy of all the models by fine-tuning every hyper parameters and parameters. Further, the fine-tuning model can also applied to early detection of COVID-19 and help the physicians for disease diagnosis [35,36].

6. Conclusion

In this paper, we proposed a deep learning based approach to diagnose and classify pneumonia from Chest X-ray images using transfer learning. Transfer learning approach was adopted and three pre-trained architectures were used, ResNet50, InceptionV3 and InceptionResNetV2 were trained on ImageNet dataset to extract spatial and temporal features. The extracted features were classified and the collected outputs from individual architectures were used in the model. The experimental

results achieved the accuracy 93.06%, 92.97%, and 92.40% for ResNet50, InceptionV3 and InceptionResNetV2 proposed model. From the results observed that the proposed work has offered better performance than other pre-trained models. In addition, depicted that the proposed work results were much better than other architectures in terms of accuracy, and the whole performance is much better than existing works. The performance can be improved further by improving data size, employing more pre-trained architectures. Therefore, deep learning methods offered much better results than traditional methods in terms of quality of treatment and accuracy.

CRediT authorship contribution statement

Adhiyaman Manickam: Conceptualization, Software, Validation. **Jianmin Jiang:** Conceptualization, Methodology, Formal analysis, Investigation. **Yu Zhou:** Conceptualization, Methodology, Formal analysis, Investigation. **Abhinav Sagar:** Software, Data curation, Resources. **Rajkumar Soundrapandian:** Methodology, Investigation, Writing – original draft. **R. Dinesh Jackson Samuel:** 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.

References

- [1] D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan, A. Mittal, Pneumonia detection using CNN based feature extraction, in: 2019 IEEE international conference on electrical, computer and communication technologies (IIECCT), 2019, pp. 1–7.
- [2] Y. Li, Z. Zhang, C. Dai, Q. Dong, S. Badrigilan, Accuracy of deep learning for automated detection of pneumonia using chest X-ray images: a systematic review and meta-analysis, *Comput. Biol. Med.* 123 (Jul. 2020) 1–8, <https://doi.org/10.1016/j.combiomed.2020.103898>.
- [3] N.M. Elshennawy, D.M. Ibrahim, Deep-pneumonia framework using deep learning models based on chest x-ray images, *Diagnostics* 10 (9) (2020) 1–16, <https://doi.org/10.3390/diagnostics10090649>.
- [4] D. Demner, M.D. Kohli, M.B. Roseman, S.E. Shooshan, L. Rodriguez, S. Antani, G. R. Thoma, C.J. McDonald, Preparing a collection of radiology examinations for distribution and retrieval, *J. Am. Med. Inform. Assoc.* 23 (2) (2015) 304–310.
- [5] T. Cherian, E.K. Mulholland, J.B. Carlin, H. Ostensen, R. Amin, M.D. Campo, D. Greenberg, R. Lagos, M. Lucero, S.A. Madhi, K.L. O'Brien, Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies, *Bull. World Health Organ.* 83 (2005) 353–359.
- [6] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers, Chestx-ray8: Hospital scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, in: in:2017 Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2097–2106.
- [7] H.R. Roth, L. Lu, A. Seff, K.M. Cherry, J. Hoffman, S. Wang, J. Liu, E. Turkbey, R. M. Summers, A new 2.5 d representation for lymph node detection using random sets of deep convolutional neural network observations, in: in:2014 Proceedings of the 17th International Conference on Medical Image Computing and Computer-Assisted Intervention, 2014, pp. 520–527.
- [8] H. Jung, B. Kim, I. Lee, J. Lee, J. Kang, Classification of lung nodules in CT scans using three-dimensional deep convolutional neural networks with a checkpoint ensemble method, *BMC Med. Imaging* 18 (1) (2018) 1–10.
- [9] A.K. Jaiswal, P. Tiwari, S. Kumar, D. Gupta, A. Khanna, J.J.P.C. Rodrigues, Identifying pneumonia in chest X-rays: A deep learning approach, *Measurement* 145 (2019) 511–518.
- [10] N. Nasrullah, J. Sang, M.S. Alam, M. Mateen, B. Cai, H. Hu, Automated lung nodule detection and classification using deep learning combined with multiple strategies, *Sensors* 19 (3722) (2019) 1–19.
- [11] A. Pezeshk, S. Hamidian, N. Petrick, B. Sahiner, 3-D convolutional neural networks for automatic detection of pulmonary nodules in chest CT, *IEEE J. Biomed. Health Inform.* 23 (2019) 2080–2090.
- [12] Q. Ke, J. Zhang, W. Wei, D. Polap, M. Wo'zniak, L. Kosmider, R. Damaševicius, A neuro-heuristic approach for recognition of lung diseases from X-ray images, *Expert Syst. Appl.* 126 (2019) 218–232.
- [13] D.S. Kermany, M. Goldbaum, W. Cai, C.C. Valentim, H. Liang, S.L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, J. Dong, Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning, *Cell* 172 (2018) 1122–1131.
- [14] A.K. Acharya, R. Satapathy, A Deep Learning Based Approach towards the Automatic Diagnosis of Pneumonia from Chest Radio-Graphs, *Biomed Pharmacol J* 13 (1) (2020) 229–455.
- [15] P. Lakhani, B. Sundaram, Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks, *Radiology* 284 (2) (2017) 574–582.
- [16] T. Rahman, M.E. Chowdhury, A. Khandakar, K.R. Islam, K.F. Islam, Z.B. Mahbub, M.A. Kadir, S. Kashem, Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection using Chest X-ray, *Applied Sciences* 10 (9) (2020) 3233.
- [17] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, and M.P. Lungren, Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225 2017.
- [18] J. Rubin, D. Sanghavi, C. Zhao, K. Lee, A. Qadir, and M. Xu-Wilson, Large scale automated reading of frontal and lateral chest x-rays using dual convolutional neural networks. arXiv preprint arXiv:1804.07839 2018.
- [19] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, Lung pattern classification for interstitial lung diseases using a deep convolutional neural network, *IEEE Trans. Med. Imaging* 35 (5) (2016) 1207–1216.
- [20] M. Cicero, A. Bilbily, E. Colak, T. Dowdell, B. Gray, K. Perampaladas, J. Barfett, Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs, *Invest. Radiol.* 52 (5) (2017) 281–287.
- [21] R. Jain, P. Nagrath, G. Kataria, V.S. Kaushik, D.J. Hemanth, Pneumonia Detection in chest X-ray images using Convolutional Neural Networks and Transfer Learning, *Measurement* 165 (2020), 108046.
- [22] A. Bhandary, G.A. Prabhu, V. Rajinikanth, K.P. Thanaraj, S.C. Satapathy, D. E. Robbins, C. Shasky, Y.D. Zhang, J.M.R. Tavares, N.S.M. Raja, Deep-learning framework to detect lung abnormality—A study with chest X-Ray and lung CT scan images, *Pattern Recogn. Lett.* 129 (2020) 271–278.
- [23] K. Kallianos, J. Mongan, S. Antani, T. Henry, A. Taylor, J. Abuya, M. Kohli, How far have we come? Artificial intelligence for chest radiograph interpretation, *Clin. Radiol.* 74 (2019) 338–345.
- [24] S. Albawi, T.A. Mohammed, S. Al-Zawi, Understanding of a convolutional neural network, in: In:2017 Proceedings of the 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1–6.
- [25] C. Bailer, T. Habtegebrial, K. Varanasi, D. Stricker, Fast Feature Extraction with CNNs with Pooling Layers. arXiv 2018, arXiv:1805.03096.
- [26] M. Raghu, C. Zhang, J.M. Kleinberg, S. Bengio, Transfusion: Understanding Transfer Learning with Applications to Medical Imaging. arXiv 2019, arXiv: 1902.07208.
- [27] H. Ravishankar, P. Sudhakar, R. Venkataramani, S. Thiruvenkadam, P. Annangi, N. Babu, V. Vaidya, Understanding the Mechanisms of Deep Transfer Learning for Medical Images, in: in:2016 proceedings Deep Learning and Data Labeling for Medical Applications; DLMIA 2016, LABELS 2016; 2016; Volume 10008.
- [28] D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, in:2015 proceeding 3rd International Conference for Learning Representations, San Diego 2015.arXiv:1412.6980v9.
- [29] L. Luo, Y. Xiong, Y. Liu, and X. Sun, Adaptive Gradient Methods with Dynamic Bound of Learning Rate, arXiv preprint arXiv:1902.09843. 2019 Feb 26..
- [30] K. He, X. Zhang, S. Ren, J. Sun, Identity Mapping in Deep Residual Networks, in: 2016 proceeding European conference on computer vision, 2016 pp. 630–645.
- [31] G. Liang, L. Zheng, A transfer learning method with deep residual network for pediatric pneumonia diagnosis, *Comp. Methods Programs Biomed.* 187 (2019), 104964.
- [32] I. Siraztdinov, M. Kholiavchenko, T. Mustafaev, Y. Yixuan, R. Kuleev, B. Ibragimov, Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database, *Comp. Electr. Eng.* 78 (2019) 388–399.
- [33] E. Ayan, H.M. Ünver, Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning, in: in: 2019 proceeding Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), 2019, pp. 1–5.
- [34] A. Ismail, T. Rahmat, S. Aliman, Chest X-ray image classification using faster RCNN, *Malaysian J. Comput.* 4 (1) (2019) 225–236.
- [35] A.M. Ismael, A. Şengür, Deep learning approaches for COVID-19 detection based on chest X-ray images, *Expert Syst. Appl.* 164 (2020), 114054.
- [36] E. Deniz, A. Şengür, Z. Kadıroğlu, Y. Guo, V. Bajaj, U. Budak, Transfer learning based histopathologic image classification for breast cancer detection, *Health information Sci. Syst.* 6 (1) (2018) 1–7.
- [37] R. Soundrapandian, An approach for infrared image pedestrian classification based on local directional pixel structure elements' descriptor, *Int. J. Comput. Aided Eng. Technol.* 13 (3) (2020) 271–290.
- [38] S.S. Koushik, K.G. Srinivasa, Detection of respiratory diseases from chest X rays using Nesterov accelerated adaptive moment estimation, *Measurement* 176 (2021), 109153.
- [39] M. Loey, G. Manogaran, M.H.N. Taha, N.E.M. Khalifa, A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic, *Measurement* 167 (2021), 108288.
- [40] M. Ezhilan, I. Suresh, N. Nesakumar, SARS-CoV, MERS-CoV and SARS-CoV-2: A Diagnostic Challenge, *Measurement* 168 (2021), 108335.
- [41] T. Rahman, A. Khandakar, Y. Qiblawey, A. Tahir, S. Kiranyaz, S.B.A. Kashem, M. T. Islam, S. Al Maadeed, S.M. Zughaier, M.S. Khan, M.E. Chowdhury, Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images, *Comput. Biol. Med.* 132 (2021) 104319.
- [42] T.S. Ahmed, A.E. Mohamed, T.J. Amani, D. Robertas, F.H. Osama, A Novel Method for Detection of Tuberculosis in Chest Radiographs Using Artificial Ecosystem-

- Based Optimisation of Deep Neural Network Features, *Symmetry* 1146 (2020) 1–17.
- [43] X. Li, S. Luo, Q. Hu, J. Li, D. Wang, F. Chiong, Automatic lung field segmentation in X-ray radiographs using statistical shape and appearance models, *J. Med. Imag. Health Inform.* 6 (2) (2016) 338–348.
- [44] T.Y. Goh, S.N. Basah, H. Yazid, M.J.A. Safar, F.S.A. Saad, Performance analysis of image thresholding: Otsu technique, *Measurement* 114 (2018) 298–307.