ORIGINAL RESEARCH

Detection of pneumonia infection in lungs from chest X-ray images using deep convolutional neural network and content-based image retrieval techniques

T. Rajasenbagam[1] · S. Jeyanthi[2] · J. Arun Pandian[3]

Abstract

In this research, A Deep Convolutional Neural Network was proposed to detect Pneumonia infection in the lung using Chest X-ray images. The proposed Deep CNN models were trained with a Pneumonia Chest X-ray Dataset containing 12,000 images of infected and not infected chest X-ray images. The dataset was preprocessed and developed from the Chest X-ray8 dataset. The Content-based image retrieval technique was used to annotate the images in the dataset using Metadata and further contents. The data augmentation techniques were used to increase the number of images in each of class. The basic manipulation techniques and Deep Convolutional Generative Adversarial Network (DCGAN) were used to create the augmented images. The VGG19 network was used to develop the proposed Deep CNN model. The classification accuracy of the proposed Deep CNN model was 99.34 percent in the unseen chest X-ray images. The performance of the proposed deep CNN was compared with state-of-the-art transfer learning techniques such as AlexNet, VGG16Net and InceptionNet. The comparison results show that the classification performance of the proposed Deep CNN model was greater than the other techniques.

Keywords Content-based image retrieval · Deep convolutional neural network · Deep convolutional generative adversarial network · VGG19Net · Data augmentation

1 Introduction

Pneumonia infection is a serious disease of the lungs with a range of possible causes. Bacteria, viruses or fungi can cause Pneumonia infection Wale and Sonawani (2018), Varadharajan et al. (2018), Muthazhagan et al. (2020). Pneumonia is a serious complication of the new coronavirus (COVID-19). Pneumonia infection is ranked as the eighth leading cause of death in the US at 2017 Jalil and Fraig (2019). Also, it causes death in children younger than five years of age worldwide. There are three different types in pneumonia infection such as bacterial pneumonia, viral pneumonia, and fungal pneumonia. The bacterial pneumonia is the most common Pneumonia infection mainly caused by bacterium Streptococcus pneumoniae Krishna et al. (2017). The Viral pneumonia is the most dangerous Pneumonia infection caused by the respiratory syncytial virus (RSV) and influenza types A and B. Fungal pneumonia is caused by the coccidioides fungus and creates the valley fever Beuy and Viroj (2019).

Chest X-rays are the commonly used method to detect the Pneumonia infection and locate the infected area in the lungs Jaeger et al. (2014). Also, the chest X-ray is the widely used radiological examination technique toward diagnosis of several lung diseases Demner-Fushman et al. (2016). Finding radiological examiners in remote places for analysis for more number of Chest X-rays is an extremely challenging task Dou et al. (2016). In recent times, artificial intelligence approaches are used to solve the challenges in several of medical diagnosis processes (Hwang and Kim 2016; Leaman et al. 2015). Mainly, the deep learning and computer vision techniques are supports to the diagnosis of various cancer and genome diseases Moeskops et al. (2016).

✉ T. Rajasenbagam
trajasenbagam@gct.ac.in; rajasenbagamcse@gmail.com

1 Department of CSE, Government College of Technology, Coimbatore, India

2 Department of CSE, PSNA College of Engineering and Technology, Dindigul, India

3 Department of CSE, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Avadi, India

Deep learning is a subset of machine learning in artificial intelligence (AI) that has learnings from large volume of unstructured data Jamaludin et al. (2017). Also, Designing and developing the deep learning model to solve any problem consumes more time and computational resources Roth Holger et al. (2015). Transfer learning techniques are introduced to avoid the deep learning model development challenges Ronneberger et al. (2015). Transfer learning make use of the knowledge gained by the deep learning network while solving one problem and applying this knowledge to solve a similar problem Azimi et al. (2018). The new dataset for the new problem was used to fine-tune the existing deep learning model weights Setio et al. (2016). A Convolutional Neural Network is a multi-layer neural network and recognizes visual patterns directly from pixel images with minimal preprocessing (Roth et al. 2014; Yao et al. 2016). The most common of transfer learning techniques using the convolutional Neural Network for solving images classification problems are the AlexNet, VGGNet, ResNet, and InceptionNet Shin et al. (2016).

VGGNets are classified as a two types based on the number of convolutional layers such as VGG16Net and VGG19Net. In general, the VGG19Net gives a better performance than the VGG16Net. The VGG19Net consists of nineteen convolutional layers and is very appealing because of its very uniform architecture Obinikpo and Kantarci (2017). It is currently the most preferred choice in the community for extracting features from images. The weight configuration of the VGG19Net is publicly available and has been used in many other applications and challenges as a baseline feature extractor. However, the VGG19Net consists of 46 Number of layers and 143 million parameters, which can be a bit challenging to handle Roth.Holger et al. (2015).

In this research, the VGG19Net based Pneumonia infection diagnosis model was proposed using chest X-ray images and trained with the GPU system. The subsequent sections of the article were organized as follows: Sect. 2 provides the literature survey of the Pneumonia disease classification related works. Section 3 describes the materials and methodologies of the VGG19Net for Pneumonia infection diagnosis model. Section 4 contains experimental results and related discussions. Section 5 gives the conclusions and future directions of the research.

2 Related works

Scan and X-ray images of human organs comprises of more valuable medical information that can help diagnose and treat diseases in hospitals. The Manual examination approach was not suitable to handle the large amounts of medical imaging data. The purpose of this review is to show the capability of image processing techniques that can efficiently handle these different but closely related human disease diagnosis. It also conducts a comparative study of various image classification and image retrieval techniques that are applied to the human lung disease identification systems. The authors Ma et al. (2017) proposed a content based medical image technique to retrieve the CT Imaging Signs of Lung Diseases using considering fused, and context-sensitive similarity measures. The fused pairwise similarity was used to minimize the semantic gap for obtaining a more accurate pairwise similarity measure. The authors Ning et al. (2018) proposed the Hausdorff distance combining Tamura texture features and wavelet transform algorithm based brain MRI database and the lung CT image retrieval technique. The performance of the technique is better than the single feature texture technique.

The Three dimensional Convolutional Neural Network framework was proposed by Jamaludin et al. (2017) to produce radiological grading of spinal lumbar MRIs and also localize the predicted pathologies using intervertebral disc volumes. The authors Azimi et al. (2018) proposed an ECG data classifications model using Convolutional Neural Network, deployed the model in Internet of Things devices. This approach uses the existing ECG dataset to classify the patient's health status. The edge computing devices reduced the cost of implementation and increased the portability of the device. The authors Obinikpo and Kantarci (2017), reviewed applications and challenges of Internet of Things and deep learning techniques in smart connected health systems through wearable or invasive devices. Also, the authors discussed the open challenges in sensed and captured medical data. The authors Kumar and Gandhi (2018) proposed a scalable Internet of Things device toward diagnosis of heart diseases. The logistic regression technique was used to process the sensed data form Internet of Things device. The cloud services were used to store and retrieve the large volume of collected data from patients. Performance of the logistic regression was examined using the Receiver Operating Characteristic (ROC) analysis to predict heart disease.

Likewise, the authors Abdelaziz et al. (2018), proposed a diagnosis for healthcare services for chronic kidney diseases using the linear regression (LR) and neural network (NN) approaches. The model was deployed in the cloud platform to enhance the execution performance. The author compared the performance of various machine learning techniques. The model predicts the chronic kidney diseases with an accuracy of 97.8 in percentage. The authors Wale and Sonawani (2018) reviewed various machine learning and deep learning techniques to diagnose the various human health challenges using imaging and sensing data. The authors discussed the advantages and challenges of numerous machine learning based healthcare applications. The succeeding section presents the fundamentals of the implemented

Pneumonia infection diagnosis model and the training and testing dataset.

3 Materials and methods

The complete processes of the proposed model for Pneumonia infection identification are described further in detail. The entire method is separated into a number of stages in the following subsections, starting with collecting the images for the classification process.The proposed Pneumonia detection model consists of five phases. The complete flow of the proposed system is shown in Fig. 1.

The complete development process of the Pneumonia infection detection system was given on subsequent sub sections. It begins with dataset preparation and the preprocessing phase.

3.1 Dataset preparation and preprocess

The Chest X-ray 8 dataset was downloaded from the open data repository for diagnosing pneumonia from chest images. The Content-based visual information retrieval technique was used to retrieve the input images from the large database with minimum time conception. Data augmentation techniques were then used to enhance the size of the dataset. Also, the content-based image retrieval technique helped avoid image duplication problems. The basic data manipulation and Deep Convolutional Generative Adversarial Network (DCGAN) were used to create the augmented images. The data augmentation techniques increased the number of images in each class to 6000 images. There are 11,900 images which were used to train the proposed neural network. Table 1 illustrates the details of the training data for the Pneumonia infection identification model.

The data augmentation process was implemented by using python programming language with tensorflow and augmentor library. The dataset was splitted as a training, validation and testing set. There were 11,900 images used to train and

Table 1 Number of images in proposed dataset

Class ID	Class name	No. of training images	No. of testing images
0	No finding	5950	150
1	Pneumonia	5950	150
Total training images		11,900	300

validate the proposed VGG19Net for the detection of Pneumonia infection. There were 300 unseen images which were used to test the performance of the proposed and existing state-of-art models. Figure 2 shows the random samples from the dataset of proposed Pneumonia infection identification model.

3.2 VGG19Net design and training

The proposed pneumonia infection identification model uses the standard VGG19Net with a few modifications. The Keras application library was used to import the pre trained VGG-19Net. There are sixteen convolutional layers which are used in VGG19Net. The convolutional layers are grouped as a set and after the each set of convolutional layers the max pooling layer was introduced to reduce the dimension of the features. The dimension of the input images are represented using the below Eq. 1,

$$\dim(image) = (nH, nW, nC) \tag{1}$$

where nH and nW represents the size of the Height and Width and nC represents the number of Channels.

The each convolutional layer uses unique filters to extract the features from the input data. The dimension of the filter in VGG19Net is shown in the below Eq. 2.

$$dim(filter) = (f, f, nC) \tag{2}$$

where f represents the size of the Height and Width and nC represents the number of Channels.

Fig. 1 Flow of proposed pneumonia detection model
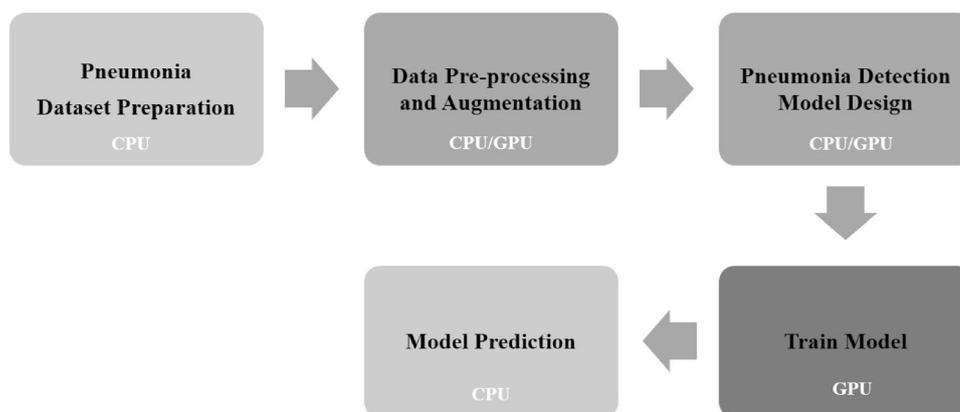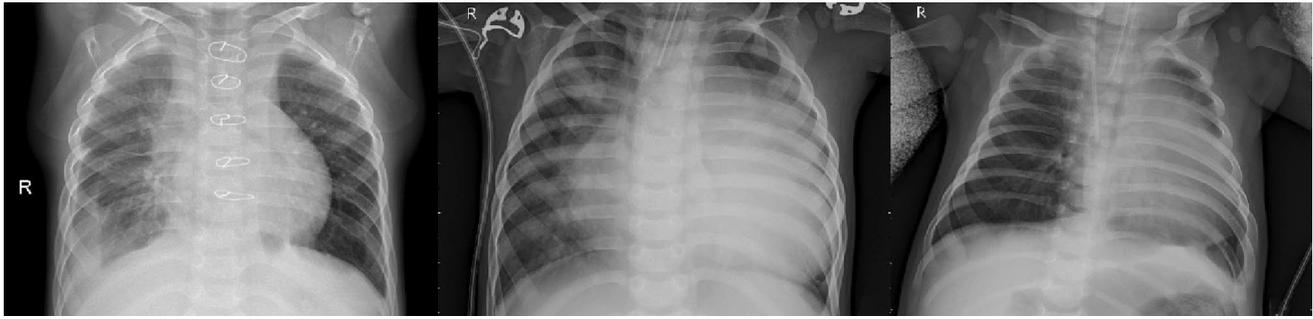
Fig. 2 Random sample images of Chest X-ray dataset

The convolutional layers extracts the feature information from the input data using the filter function and produces the feature data as an output. The convolutional function from the given input and filter is as shown in the below Eq. 3,

$$conv(I, K)_{xy} = \sum_{i=1}^{nH} \sum_{j=1}^{nW} \sum_{k=1}^{nC} K_{i,j,k} I_{x+i-1, y+j-1, k} \qquad (3)$$

where I represents the input image of the convolutional function and K represents the kernel filter. To improve the performance of the first five layers of the standard VGG-19Net, it was freezed for the training process. The training time of the model without freezing the first five layers taking very long. But, There is no changes on the training performance of the model including the first five layers. So, the layers are freezed for the training process. Also, two dense layers were added in the network for changing the number of classes from 1000 to 2 classes. The two output classes are for not finding and pneumonia. The softmax activation function was used to classify the model output to class.

Additionally, the dropout layer with dropout of 0.2 was added between the first two dense layers for avoiding the overfitting problem. Different dropout values were compared in this network and varied from 0.2 to 0.8 and the 0.5 gives the better training performance. The mini-batch gradient descent was used in this model with batch size of 16. Also, the various batch sizes were applied and the performance was compared with the batch size 16. Implementation of the proposed VGG19Net was developed and done by using the python programming language and Tensorflow library. Number of layers and input and output dimensions of the each layers are represented in the model summary. The primary designing and customizing processes of the proposed VGG19Net were developed in HP Z800 workstation. After the customization of VGG19Net for Pneumonia infection detection, the model was trained using the NVIDIA DGX1 deep learning server. There are 54,631,490 parameters which are trained in the proposed VGG19Net. The NVIDIA DGX1 deep learning server boosts the training process of the proposed VGG19Net. The summary of the proposed VGG19Net for Pneumonia infection diagnosis is shown in Fig. 3.

The Fig. 4 illustrates the training result of the proposed model with 500 epochs and of 16 batch size.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 256, 256, 3)	0
block1_conv1 (Conv2D)	(None, 256, 256, 64)	1792
block1_conv2 (Conv2D)	(None, 256, 256, 64)	36928
block1_pool (MaxPooling2D)	(None, 128, 128, 64)	0
block2_conv1 (Conv2D)	(None, 128, 128, 128)	73856
block2_conv2 (Conv2D)	(None, 128, 128, 128)	147584
block2_pool (MaxPooling2D)	(None, 64, 64, 128)	0
block3_conv1 (Conv2D)	(None, 64, 64, 256)	295168
block3_conv2 (Conv2D)	(None, 64, 64, 256)	590080
block3_conv3 (Conv2D)	(None, 64, 64, 256)	590080
block3_conv4 (Conv2D)	(None, 64, 64, 256)	590080
block3_pool (MaxPooling2D)	(None, 32, 32, 256)	0
block4_conv1 (Conv2D)	(None, 32, 32, 512)	1180160
block4_conv2 (Conv2D)	(None, 32, 32, 512)	2359808
block4_conv3 (Conv2D)	(None, 32, 32, 512)	2359808
block4_conv4 (Conv2D)	(None, 32, 32, 512)	2359808
block4_pool (MaxPooling2D)	(None, 16, 16, 512)	0
block5_conv1 (Conv2D)	(None, 16, 16, 512)	2359808
block5_conv2 (Conv2D)	(None, 16, 16, 512)	2359808
block5_conv3 (Conv2D)	(None, 16, 16, 512)	2359808
block5_conv4 (Conv2D)	(None, 16, 16, 512)	2359808
block5_pool (MaxPooling2D)	(None, 8, 8, 512)	0
flatten_1 (Flatten)	(None, 32768)	0
dense_1 (Dense)	(None, 1024)	33555456
dropout_1 (Dropout)	(None, 1024)	0
dense_2 (Dense)	(None, 1024)	1049600
dense_3 (Dense)	(None, 2)	2050

Total params: 54,631,490
Trainable params: 54,518,914
Non-trainable params: 112,576

Fig. 3 Summary of proposed VGG19Net
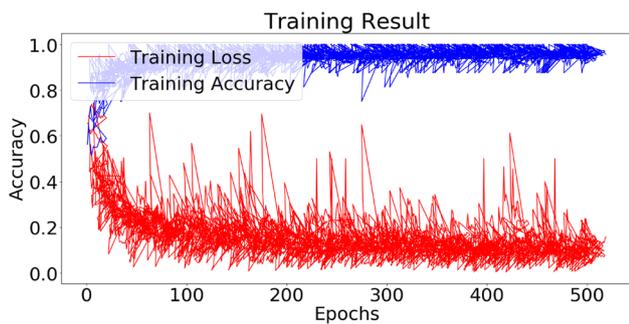
Fig. 4 Training result of proposed VGG19Net
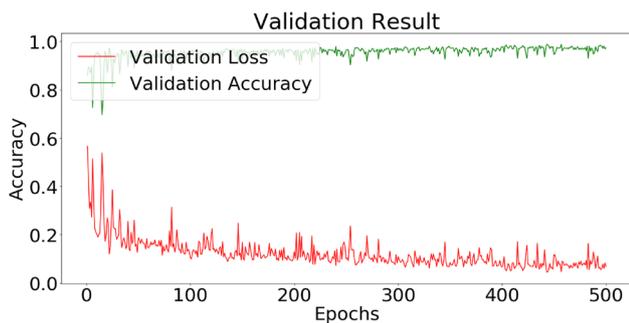
Fig. 5 Training result of proposed VGG19Net

Figure 5 represents the validation accuracy and loss of the proposed VGG19Net in proposed Pneumonia Chest X-ray dataset.

After completion of the training process the layer architecture and weight values of the proposed VGG19Net were stored as H5 for future purpose. The validation performance of the model using various datasets are shown in the Fig. 6. The datasets are original dataset, Basic manipulation techniques based augmented dataset, DCGAN based augmented dataset and combined dataset using original, basic manipulation and DCGAN techniques.
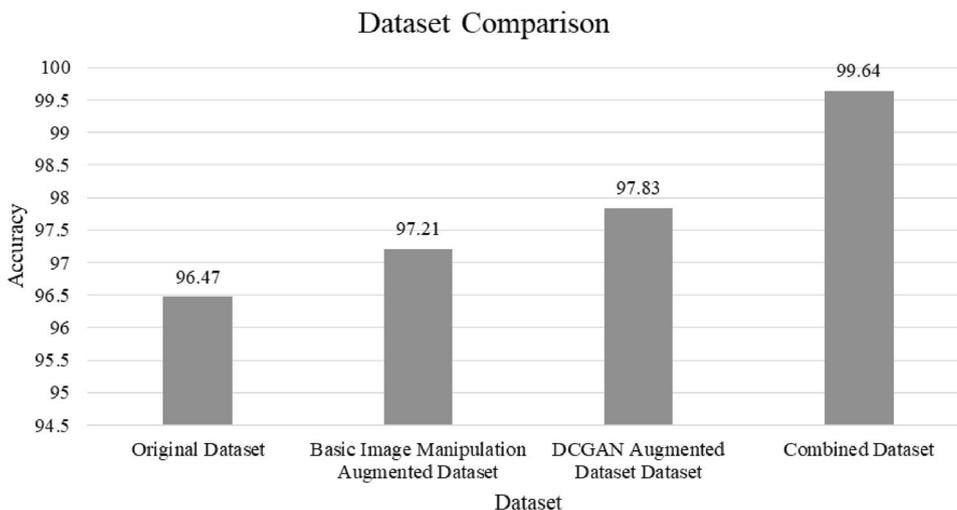
The dataset performance comparison results shows the significance of the combined data augmentation techniques in image classification datasets.

4 Results and discussions

The proposed VGG19Net and state-of-the-art networks are tested with 300 testing images for classifying the Pneumonia disease from chest X-ray images. The most common state-of-the-art classification algorithms are AlexNet, VGG-16Net and InceptionV3Net Geetharamani and Arun Pandian (2019). Confusion matrix represents the performance of the classification algorithms. The confusion matrix shows the True Positive (TP), True negative (TN), False positive (FP) and False Negative (FN) of the classification algorithm. Fig. 7 illustrates the confusion matrix of the proposed Pneumonia infection detection model.

Label 0 represents "Pneumonia" and label 1 represents "No Finding" of the confusion matrix of the proposed VGG-19Net model. The confusion matrix represents the classification performance of the proposed model on testing data.

The Receiver Operating Characteristic (ROC) curve is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. Figure 8 illustrates the roc curves of the proposed binary classifier.

The ROC curves show the classification superiority of the proposed VGG19Net on Pneumonia diagonsis.

Classification accuracy is most important performance metrics of classification algorithm. The following Eq. 4 was used to calculate the classification accuracy from True Positive (TP), True negative (TN), False positive (FP) and the False Negative (FN) of confusion matrix.

Fig. 6 Dataset performance comparison
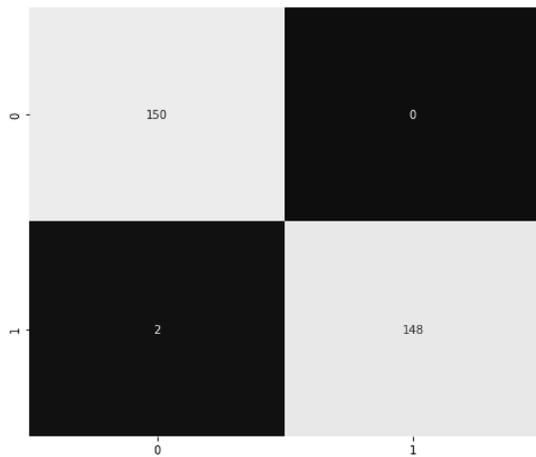
Fig. 7 Confusion matrix of proposed VGG19Net

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (4)$$

The classification accuracy of the proposed VGG19Net is much higher than the other approaches. Figure 9 compares the classification accuracy of proposed VGG19Net and state-of-the-art classification techniques.

Another most important of performance metrics is precision. Precision is the fraction of relevant instances among the retrieved instances of classification techniques. The precision value of the proposed VGG19Net is superior to the existing techniques. The following Eq. 5 represents the calculation process of precision,

$$\text{Precision} = \frac{TP}{TP + FP} \qquad (5)$$
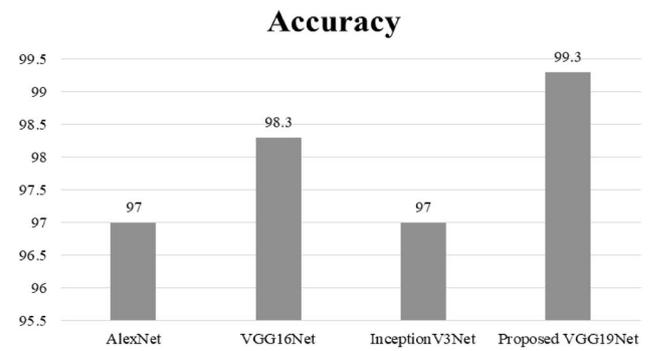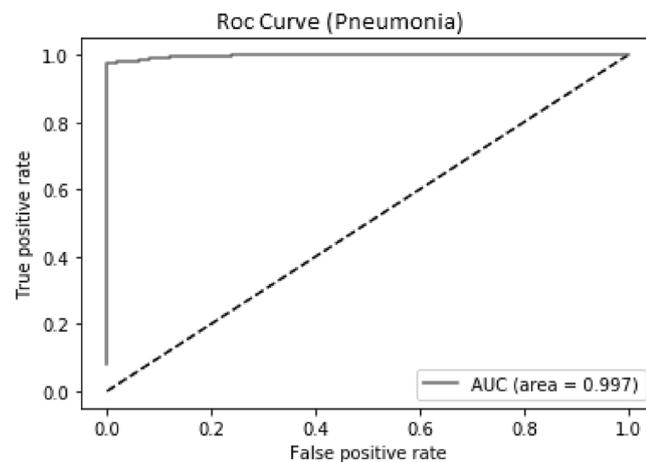
Fig. 9 Comparison of accuracy

Figure 10 illustrates the precision value of the proposed VGG19Net and other state-of-the-art approaches.

Recall is also one of the most common performance metrics to estimate the performance of the classification model. The Recall is the fraction of the total amount of relevant instances that were actually the retrieved instance of the classification algorithm. The following Eq. 6 represents the calculation process of recall,

$$\text{Recall} = \frac{TP}{TP + TN} \qquad (6)$$

Figure 11 shows the recall value of proposed VGG19Net and the existing classification techniques such as AlexNet, VGG16Net and InceptionV3Net.

The recall value of the proposed VGG19Net is found better than in the state-of-the-art classification techniques.

In statistical analysis of binary classification, the F1 score is a measure for testing accuracy. It considers both the precision and the recall of the classification algorithm to compute the F1 score. The following Eq. 7 represents the calculation of the F1 score,
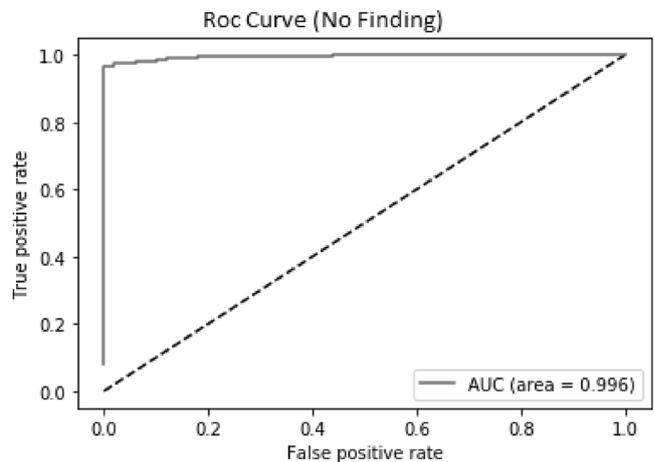
Fig. 8 AuC–RoC curves of proposed VGG19Net
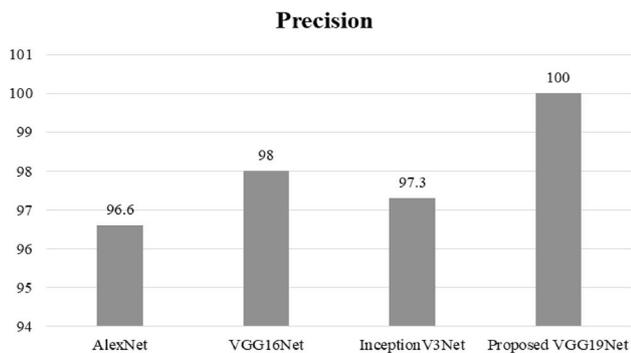
Precision

Fig. 10 Comparison of precision
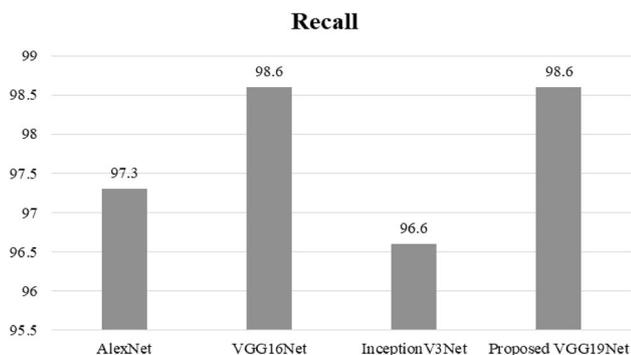
Recall

Fig. 11 Comparison of recall
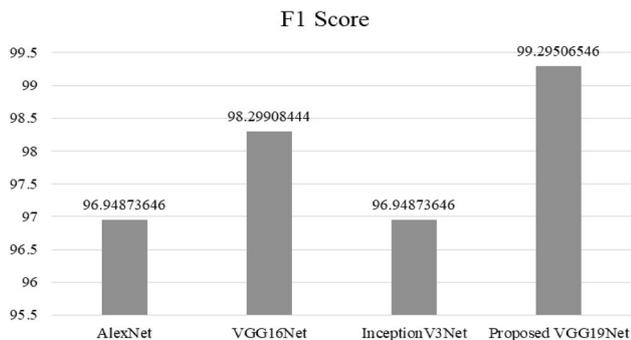
F1 Score

Fig. 12 Comparison of F1 score

$$\text{F1 Score} = \frac{\text{Precision*Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

The F1 score of the proposed VGG19Net is superior to the existing approaches. Figure 12 shows the F1 score of the proposed VGG19Net and in existing AlexNet, VGG16Net and InceptionV3Net.

The outcomes of the most common performance metrics such as classification accuracy, precision, recall and F1 score, shows that the proposed VGG19Net gives a better performance than the AlexNet, VGG16Net and InceptionV3Net.

Also the proposed VGG19Net model was deployed on the embedded device and mobile application for the real-time implementation using the tensorflow lite package.

5 Conclusion and future works

Reliable recognition of infections in the lung is a key step in the diagnosis of Pneumonia disease. X-ray imaging examination of Chest is usually performed by trained human examiners or doctors, making the process time-consuming and hard to standardize. This research proposed and developed a Pneumonia detection model using the Deep Convolutional Neural Network and Pneumonia Chest X-ray dataset. This data was collected from the various patients and clinically examined and categorized by human examiners. The proposed Deep Convolutional Neural Network was trained on by using 1000 training epochs with NVidia tesla v100 GPU and TensorFlow framework. The training process of the model uses 7000 chest X-ray Images and the testing process uses 200 images. The performance of the proposed model used, evaluated thus using different metrics such as Classification accuracy, Sensitivity, Specificity and the F1 score. The Classification accuracy of the proposed model achieved the average accuracy of 99.53 percentage in unseen chest X-ray images. Also, this accuracy was greater than the existing transfer learning approaches such as Alexnet, Resnet, and InceptionNet. The proposed Deep Convolutional Neural Network was found most suitable to detect Pneumonia infection from Chest X-ray images.

In the future, different lung disease classes will include this model to detect various lung diseases using the chest X-ray images. Also, the performance of the proposed Deep CNN model can be improved with more number of layers and parameters. This will allow clinicians to recognize lung diseases from chest X-ray images with lower prevalence at an earlier stage of the disease.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Abdelaziz A et al (2018) A machine learning model for improving healthcare services on cloud computing environment. Measurement 119:9

Azimi I et al (2018) Empowering healthcare IoT systems with hierarchical edge-based deep learning. IEEE Int Conf Connect Health Appl Syst Eng Technol 2018:8

Beuy J, Viroj W (2019) Pulmonary pathology of early phase 2019 novel coronavirus pneumonia. J Thoracic Oncol 15:5

Demner-Fushman D, Kohli MD, Rosenman MB et al (2016) Preparing a collection of radiology examinations for distribution and retrieval. J Am Med Inform Assoc 2016:5

Dou Q, Chen H, Yu L, Zhao L, Qin J, Wang D, Mok V, Shi L, Heng P (2016) Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. IEEE Trans Med Imaging 2016:35

Geetharamani G, Arun-Pandian J (2019) Identification of plant leaf diseases using a nine-layer deep convolutional neural network. Comput Electr Eng 76:9

Hwang S, Kim HE (2016) Self-transfer learning for weakly supervised lesion localization, MICCAI

Jaeger S, Candemir S, Antani S, Wng YXJ, Lu PX, Thoma G (2014) Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. Quant Imaging Med Surg 2014:4

Jalil B, Fraig M (2019) Transmission electron microscopy in severe pneumonia: back to the future? Am J Med Sci 357:5

Jamaludin A, Kadir T, Zisserman A (2017) SpineNet: Automated Classification and Evidence Visualization in Spinal MRIs. Med Image Anal 2017:5

Krishna R, Zhu Y, Groth O, Johnson J, Hata K, Kravitz J, Chen S, Kalantidis Y, Li LJ, Shamma DA, Bernstein M, Fei-Fei L (2017) Visual genome: connecting language and vision using crowd sourced dense image annotations. Int J Comput Vis 123:5

Kumar PM, Gandhi UD (2018) A novel three-tier Internet of Things architecture with machine learning algorithm for early detection Of Heart diseases. Comput Electr Eng 65:8

Leaman R, Khare R, Lu Z (2015) Challenges in clinical natural language processing for automated disorder normalization. J Biomed Inf 57:8

L Ma, X Liu, Y Gao, Y Zhao, X Zhao, C Zhou.(2017) A new method of content based medical image retrieval and its applications to CT imaging sign retrieval. Journal of Biomedical Informatics, 2017:66.

Moeskops P, Wolterink J, Van der Velden B, Gilhuijs K, Leiner T, Viergever M, Isgum I (2016) Deep learning for multitask medical image segmentation in multiple modalities, MICCAI

Muthazhagan B, Ravi T, Rajinigirinath D (2020) An enhanced computer-assisted lung cancer detection method using content based image retrieval and data mining techniques. J Ambient Intell Hum Comput. https://doi.org/10.1007/s12652-020-02123-7

Ning Z, Haibin W, Xiaoyang Y, Xue W, Shuang Y (2018) Medical image retrieval approach by texture features fusion based on Hausdorff distance. Math Probl Eng 2018:66

Obinikpo A, Kantarci B (2017) Big sensed data meets deep learning for smarter health care in smart cities. J Sens Actuator Netw 2017:5

Ronneberger O, Fischer P, Brox T(2015) U-net: convolutional networks for biomedical image segmentation, MICCAI 2015, lecture notes in computer science, vol 9349

Roth HR, Lu L, Seff A, Cherry K M, Hoffman J, Wang S, Liu J, Turkbey E, Summers RM (2014) A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations. In: MICCAI

Roth Holger R, Lu L, Farag A, Shin HC, Liu J, Turkbey EB, Summers RM (2015) DeepOrgan: multi-level deep convolutional networks for automated pancreas segmentation, MICCAI 2015, Lecture Notes in Computer Science, vol 9349

Setio AAA et al (2016) Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imaging 2016:35

Shin H, Lu L, Kim L, Seff A, Yao J, Summers R (2016) Interleaved text/image deep mining on a large-scale radiology database for automated image interpretation. J Mach Learn Res 17:8

Varadharajan R, Priyan MK, Panchatcharam P et al (2018) A new approach for prediction of lung carcinoma using back propagation neural network with decision tree classifiers. J Ambient Intell Hum Comput. https://doi.org/10.1007/s12652-018-1066-y

Wale AS, Sonawani SS (2018) A survey on deep learning approaches in healthcare. Int J Innov Res Comput Commun Eng 6(5):5

Yao J et al (2016) A multi-center milestone study of clinical vertebral CT segmentation. Comput Med Imaging Graph 49:8