LEC Journal 2022, 4(1): 43-48
ISSN: 2565-5205 (print)

Research Article

PNEUMONIA OF CHEST X-RAY IMAGES DETECTION USING VGG ARCHITECTURES

BAL KRISHNA NYAUPANE
Department of Electronics and Computer Engineering
Paschimanchal Campus
Institute of Engineering, Tribhuvan University
bkn@wrc.edu.np

ABSTRACT

Pneumonia is one of the world's most frequent diseases. Pneumonia, a serious respiratory infection, damages the lungs and makes it difficult to breathe. Chest X-ray scans are one of the most significant and common procedures for diagnosing pneumonia condition. An effective model for the identification of pneumonia based on digital chest X-ray pictures is proposed in this paper, which could help radiologists make better decisions. This is a supervised learning technique in which the proposed model predicts the outcome depending on the dataset's image quality. In this research, Kaggle dataset has been used to train the model. To improve training and validation accuracy, fine-tuning of Visual Geometry Group (VGG) pre-trained model are conducted with different hyperparameters. The deep Convolutional Neural Networks (CNNs) based VGG-16 and VGG-19 architectures are used to extract features from given chest X-ray images. These features are then used for classification. In addition to the accuracy and F1-score as an evaluation matrix, the results of pre-trained models are compared using the model loss, and model accuracy graphs. The model's performance in detecting pneumonia demonstrates that the proposed VGG pre-trained model can efficiently categorize normal and pneumoniachest X-rays images in practice. As a result, the proposed model can be utilized to make a speedy diagnosis of pneumonia and can assist radiologists with their work.

Keywords—*Pneumonia, Chest X-rays, fine-tuning, VGG, pre-trained model, extract features*

I. INTRODUCTION

In underdeveloped countries, pneumonia is a calamitous disease that kills people like the elderly, and babies. Pneumonia, which affects 450 million people globally, causes breathing problems by infecting the lungs and alveoli[1].Pneumonia can be caused by a variety of germs, but it is most commonly caused by bacteria or viruses in the air we breathe. Pneumonia is divided into two types: bacterial and viral[2]. Bacterial pneumonia, in general, is associated with more severe symptoms. Treatment is the most significant distinction between bacterial and viral pneumonia. Antibiotic therapy is used to treat bacterial pneumonia, although viral pneumonia normally improves on its own [3]. A high degree of pollution is the primary culprit. In the United States, pneumonia is listed eighth among the top ten causes of mortality. Especially in the case of elderly people, the condition is usually neglected and untreated until it has progressed to the point of death. It is the leading cause of death in children around the world. According to the World Health Organization (WHO), pneumonia kills over two million children under the age of five every year. According to the WHO, nearly all instances of childhood pneumonia (95 percent) occur in underdeveloped countries, notably in Southeast Asia and Africa [1].

Chest X-rays, CT scans of the lungs, ultrasound scans of the chest, needle biopsy of the lung, and MRI scans of the chest can all be used to diagnose pneumonia. Chest X-rays are currently one of the most effective ways for detecting pneumonia. X-ray imaging is favoured over CT imaging because CT imaging takes significantly longer than X-ray imaging and many underdeveloped locations may lack sufficient high-quality CT scanners[3]. X-rays, on the other hand, are the most popular and readily available diagnostic imaging technology, and they play an important role in clinical care and epidemiological research. Several regions around the world have a scarcity of experienced healthcare personnel and radiologists, whose capacity to forecast such diseases is critical[3]. Deep-learning based computer-aided diagnostics is getting increasingly prevalent these days. In this paper, a model based on deep learning and transfer learning is described that is capable of automatically classifying whether a patient has pneumonia or not. The suggested method employs a transfer learning algorithm to extract features from the X-ray image that automatically describe the presence of disease and indicate whether or not it is a case of pneumonia.

Methods for examining the identification of disease by chest X-ray have also been conducted previously using a different approach. Researchers have developed many deep-learning-based techniques for pneumonia detection in recent years. M. F. Hashmi et al. [3]used the Large Dataset of labelled Optical Coherence Tomography (OCT) and Chest X-Ray Images using weighted classifier of five pre-trained models such as deep learning models such as ResNet18, Xception, InceptionV3, DenseNet121, and MobileNetV3 and achieved test accuracy of 98.43 percent. D. Zhang et al. [6] used a dataset from the Kaggle competition which consists of 5786 X-ray images and performed the pre-processing using Dynamic Histogram Equalization (DHE) to improve the quality of images. The researcher has used Convolutional Neural Network with four two convolutional layers with max-pooling and cascading with three fully connected layers. The maximum obtained accuracy rate of 96.07% and the precision rate of 94.41%. A. Akgundogdu et al.[1] used a dataset containing 5856 chest X-ray images which include 4273 cases of pneumonia and 1583 normal case. In this work, feature extraction was done by using the 2D Discrete Wavelet method and then applied the classification by using Random Forest (RF). To improve the performance of RF 10-fold Cross-validation technique had used. The maximum achieved classification accuracy of the proposed was 97.11%.

II. Materials and Methodology

In this section, image acquisition, transfer learning, pre-trained VGG models, proposed methodology, fine-tuning and hyperparameter used for pre-trained models and performance metrics for evaluation are briefly explained to fulfill the objective of this research.

A. *Image Acquisition*

The Kaggle dataset "RSNA Pneumonia Detection Challenge" has been used in this research. Kaggle Pneumonia Detection Challenge contains 29,700 images for pneumonia and normal case. Among them, for this research 2585 randomly selected images are used for the classification. Each image is labelled by 0 for normal and 1 for Pneumonia. The main objective of this research is to detect the pneumonia of chest X-ray images. As mentioned in Table 1, the training dataset includes 902 images of normal; 1203 images of pneumonia; and the testing dataset includes 235 images of normal; and 1245 images of pneumonia.

TABLE II. TABLE 1. DATASET DESCRIPTIONS

Classes	Training Dataset	Test Dataset
Normal	902	235
Pneumonia	1203	245

A.

B. *Transfer Learning*

Transfer learning is the process of reusing previously trained models for a large-scale dataset on a new problem domain. Transfer learning is a method for adapting a pre-trained model to a new dataset by adjusting the weights and learned parameters. In transfer learning, the proposed new model uses the previously acquired knowledge from a prior model to increase generalization about a new domain. Transfer learning can be used in a variety of ways. Each method can be beneficial in terms of constructing and training a deep convolutional neural network model while also saving time. Because it's not always evident which pre-trained model will perform best on new computer vision tasks, some trial and error may be required. Transfer learning provides a number of advantages, the most important are reduced training time, improved neural network performance (in most circumstances), and requires less amount of data[4]. Transfer learning has a number of approaches for implementation. Among them, some of the most used approaches are as follows.

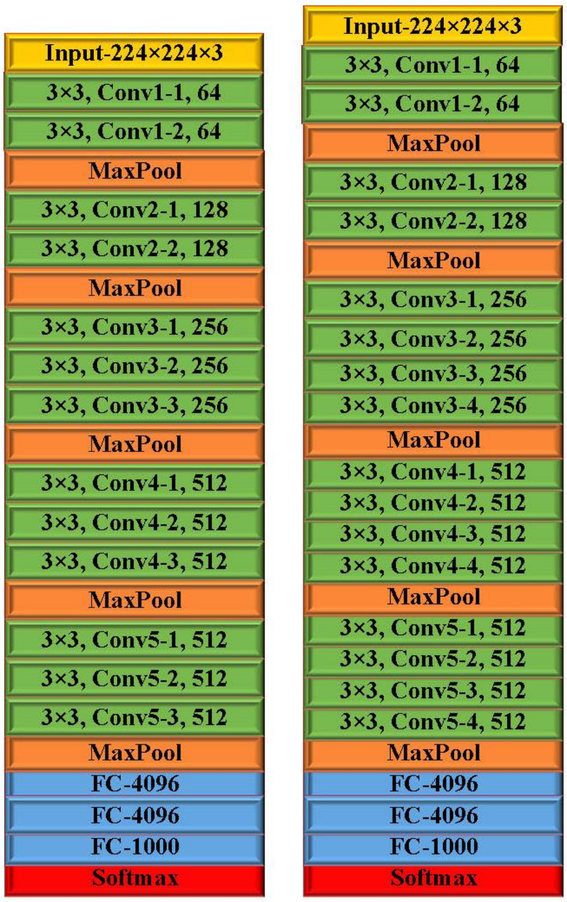
Freeze all the Convolution Layer: in this method, all the layers of the pre-trained model make untrainable and add a number of the fully connected layer as required to train the model for the new dataset.

Feature extraction: in this approach, most of the lower layers of the pre-trained model make trainable false and some higher layers of the pre-trained model make trainable. How many numbers of layers make trainable or frozen depends on the new domain of the data set and model desired accuracy. The fully connected layer can be added as per the requirement to train the new model with the new dataset.

C. *Pre-trained VGG Architectures*

VGG stands for Visual Geometry Group. The VGG architecture is a convolutional neural network model introduced by Simonyan and Zisserman in their 2014 paper "Very Deep Convolutional Networks for Large Scale Image Recognition" [5]. For Training of VGG architectures, the input image is three channels with fixed-size 224 × 224. The input image is passed through a stack of convolutional layers, where all the convolutional layers use the filter with a size of 3 × 3. The convolution stride is set to 1 pixel, and the spatial padding of the convolution layer input ensures that the spatial resolution is maintained after convolution.

Five Max-pooling layers perform spatial pooling and it follows part of the convolutional layers (not all the convolutional layers). With stride 2, max-pooling is done over a 2× 2pixel window. Three fully connected layers follow a stack of convolutional layers: the first two each have 4096 channels, while the third performs 1000-way ILSVRC classification and so has 1000 channels. The soft-max layer is the final layer. All the configuration of the fully connected layer is the same for all networks [5]. The VGG architecture has two popular pre-trained architectures- VGG16 and VGG19. The VGG16 architecture has a stack of thirteen convolutional layers followed by three fully connected layers. The VGG19 has a stack of sixteen convolutional layers followed by three fully connected layers. The Figure 1 depicts the overall operation of VGG16 and VGG19 architectures.



(a) (b)

Fig. 1. VGG Architectures (a) VGG16 (b) VGG19

D. *Fine tuning of pre-trained models*

Fine-tuning is the process of unfreezing the entire model or a portion of a model or all of the model's layers and it depends on model and new dataset. Remove the full connected layer of architectures and replace it with some full connected layers to match the number of classes, then retrain the model with a very low learning rate on the new dataset. By gradually modifying the pre-trained model weight to the new dataset, the low learning rate could produce considerable improvements. If randomly initialized trainable layers are mixed with pre-trained model trainable layers during training, the randomly initialized layers will produce very large gradient updates, destroying pre-trained model weights. Because training is often a tiny dataset, it is crucial to employ a relatively low learning rate at this stage. As a result, if the model updates a big weight, there is a risk of overfitting. To avoid overfitting, the model needs to incrementally readjust the pre-trained weights.

Figure 2 shows the how fine-tuning process is conducted in the pre-trained model. In case 1- make all the trainable layers of pre-trained models false and add some fully connected layers with SoftMax function. In case 2- make some of the trainable layers false and the rest leave true. Table 2 shows the number of frozen and unfrozen layers during the fine-tuning process and trainable and untrainable parameters respectively.

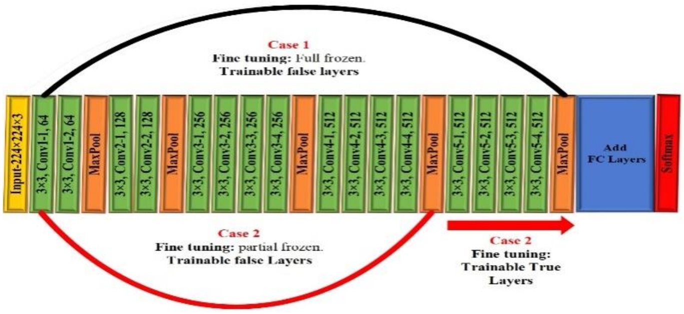


Fig. 2. Fine tuning of VGG19 pre-trained model.

TABLE III. TABLE 2. DETAILS OF EACH CASE FOR FINE-TUNING OF TABLE IV. PRE-TRAINED MODELS

Pre-trained Model	Case	Fully Connected layers	Frozen and Unfrozen Layers	Trainable Parameters	Non-Trainable Parameters
VGG16	1.Full Frozen	[128, 64, 32]	All frozen	16,788,161	14,715,136
	2.Partialy Frozen	[128, 64, 32]	**Frozen:** Conv1,2,and 3 **Unfrozen:** 4 and 5	29,767,361	1,735,936
	3. Complete Unfrozen	[4]	All unfrozen	15,238,993	8
VGG19	1.Full Frozen	[128, 64, 32]	All frozen	16,788,161	20,024,832
	2.Partialy Frozen	[128, 64, 32]	**Frozen:** Conv1,2,and 3 **Unfrozen:** 4 and 5	34,486,977	2,326,016
	3. Complete Unfrozen	[4]	All unfrozen	20,548,689	8

E. *Block diagram for proposed methodology*

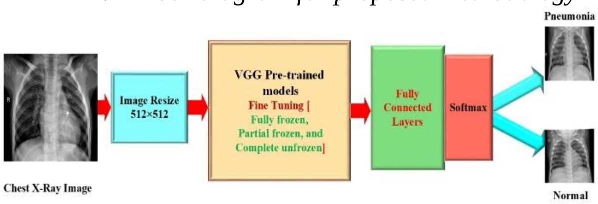


Fig. 3. Overall block diagram of purposed method.

Figure 3 explains the overall structure of this research. The chest X-ray image is the input image for this model and the input image is resized to 512×512. Two pre-trained models are trained on a new dataset of Chest X-ray images. In this research, the main focus is to fine-tune the pre-trained models VGG16 and VGG19 to achieve the highest accuracy on the new dataset. In this research, fine-tuning is carried in three cases as mentioned in Table 2. Finally, the model is used to predict the test chest X-ray images as Normal or Pneumonia.

III. Experimental Result and Analysis

To conduct this research, Google Collaboratory has used to run code directly through a browser utilizing cloud computing. The three pre-trained models of VGG architecture are trained one after another to complete this research. Google Collaboratory gives a decent GPU for free to run continuously for 12 hours. This research has used the GPU runtime equipped up to Tesla K80 with 12 GB of GDDR5 VRAM, intel Xeon Processor with two cores@ 2.20 GHz and 13 GB RAM, and available disk space of 69GB for 12 hours to perform research work.

A. *Result in Terms of Testing Accuracy and Testing Loss*

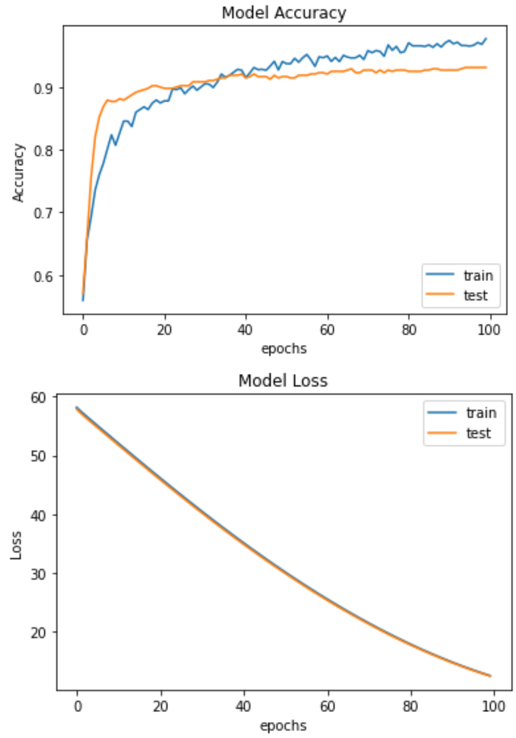


Fig. 4. VGG16 model accuracy and loss plot for all frozen layer

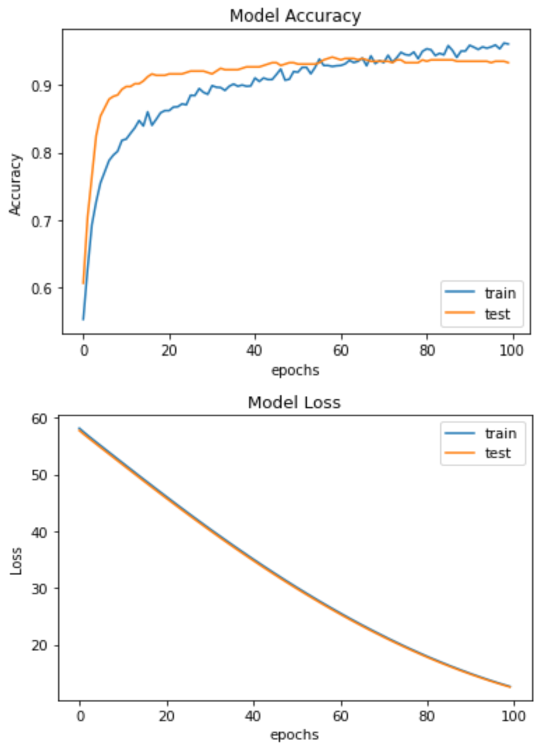


Fig. 5. VGG19 model accuracy and loss plot for all frozen layer

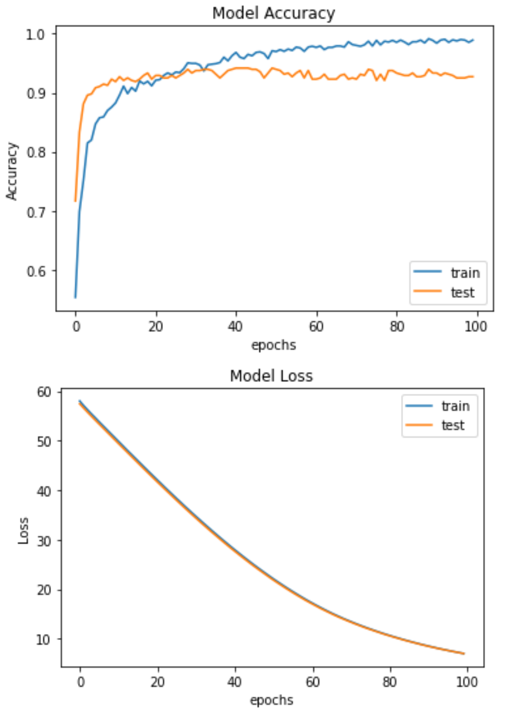


Fig. 6. VGG16 model accuracy and loss plot for Partial Frozen layer

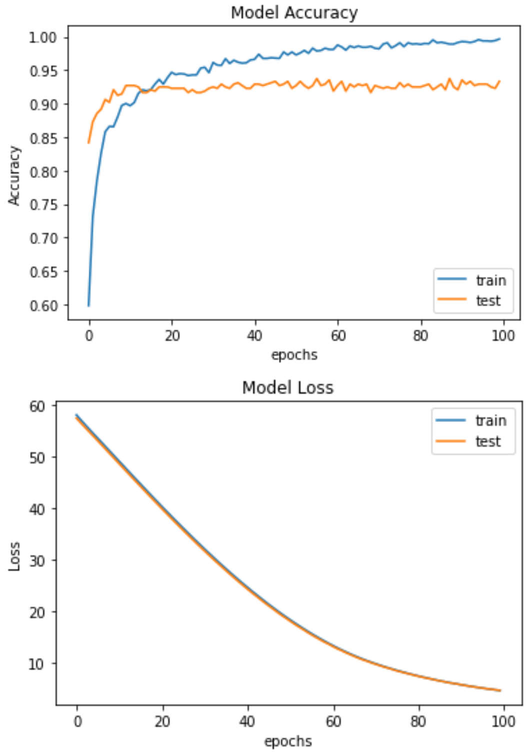
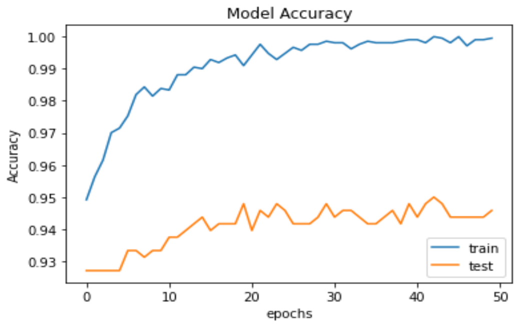


Fig. 7. VGG19 model accuracy and loss plot for Partial Frozen layer



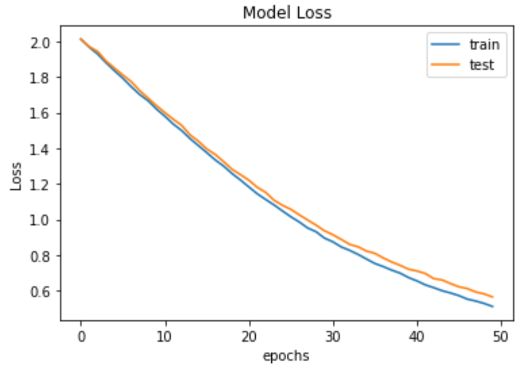


Fig. 8. VGG16 model accuracy and loss plot for completely unfrozen layer

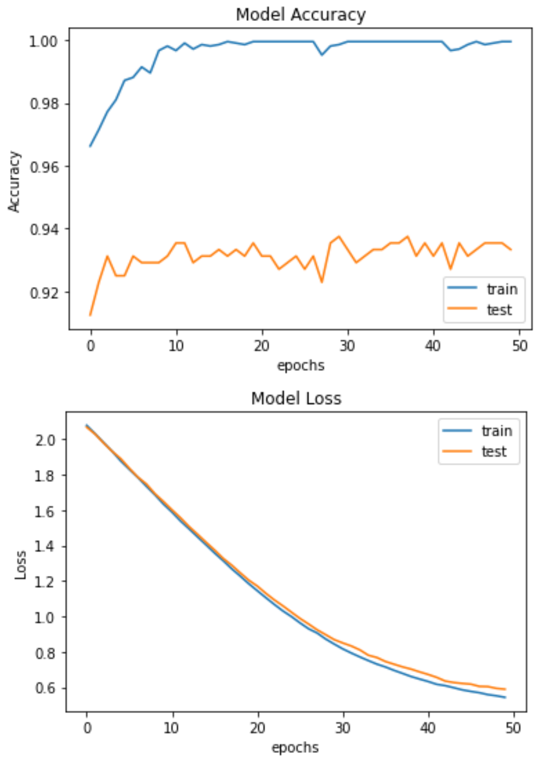


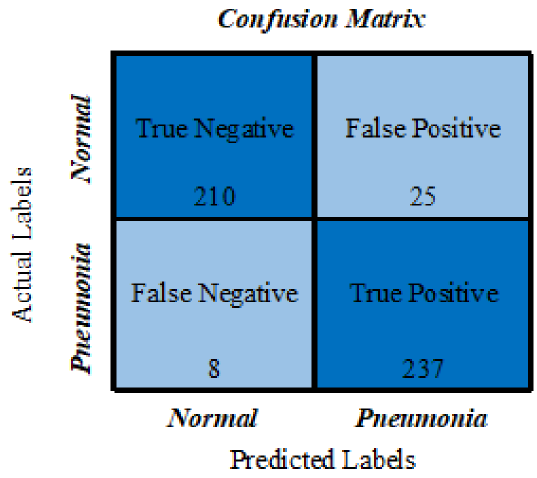
Fig. 9. VGG19 model accuracy and loss plot for completely unfrozen layer

In Figures 4 and 5, even though the model loss plot seems to same for both VGG16 and VGG19 pre-trained models, the training accuracy is better in the VGG19 pre-trained model. The gap between training and test accuracy started to increase from 75 epochs on the VGG19 pre-trained model, whereas the gap started to increase from 25 epochs in VGG16.From figures 6 and 7, it is shown that both model loss and accuracy plots almost resemble the same. In figures 8 and 9, however, the training accuracy is 99% in both pre-trained models, but the VGG16 model has the better test accuracy with comparison to VGG19 i.e., 95%.

B. *Performance Analysis*

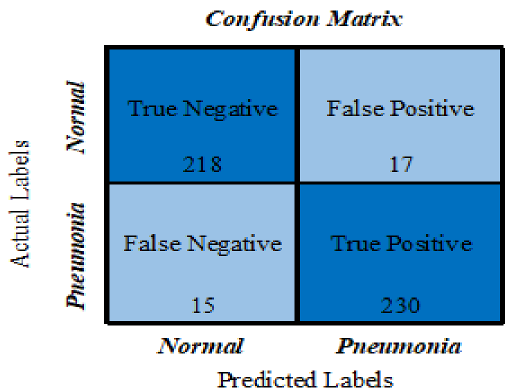
To further test the robustness of the proposed methodology confusion matrix, accuracy, recall, precision, and F1 score is calculated to compare the different cases of fine-tuning of VGG architectures pre-trained models.

In disease diagnosis, false negative means you suffered from disease, but your model predicts you are normal. That means you will not go further investigation which could be a disaster. False-positive means you are not suffering from disease, but your model predicts you suffered from the disease. That means you will go further treatment plan and investigation and ultimately you will find you have no disease. Therefore, a false-negative is more sensitive than the false positive for disease diagnosis.



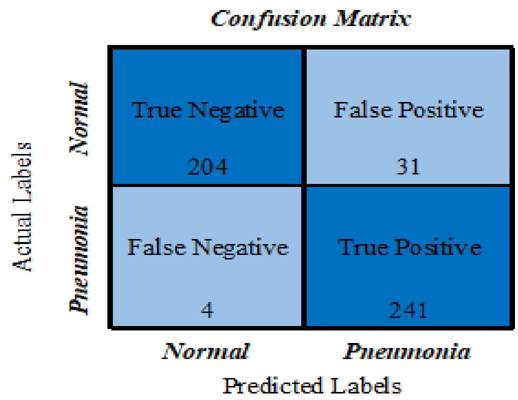
	Precision	*Recall*	*F1-score*	*Accuracy*
Normal	*0.96*	*0.89*	*0.93*	*0.93*
Pneumonia	*0.90*	*0.97*	*0.93*	

Fig10.Confusion matrix and Classification report for VGG16 Frozen all layer



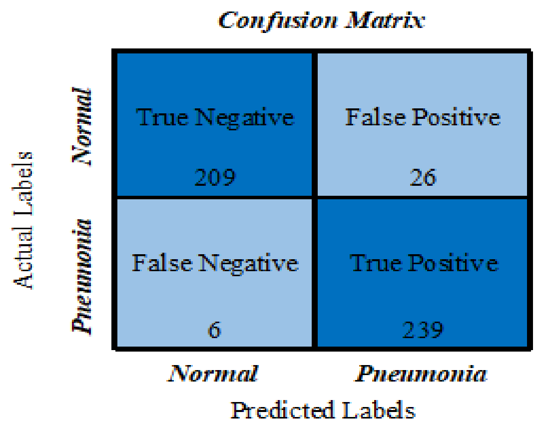
	Precision	*Recall*	*F1-score*	*Accuracy*
Normal	*0.94*	*0.93*	*0.93*	*0.93*
Pneumonia	*0.93*	*0.94*	*0.93*	

Fig. 11. Confusion matrix and Classification report for VGG19 Frozen all layer



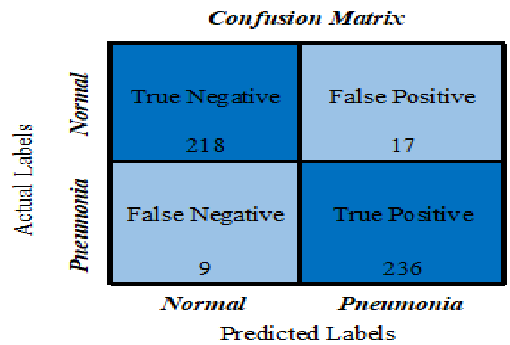
	Precision	*Recall*	*F1-score*	*Accuracy*
Normal	*0.98*	*0.87*	*0.92*	*0.93*
Pneumonia	*0.89*	*0.98*	*0.93*	

Fig. 12. Confusion matrix and Classification report for VGG16 Partial Frozen layer



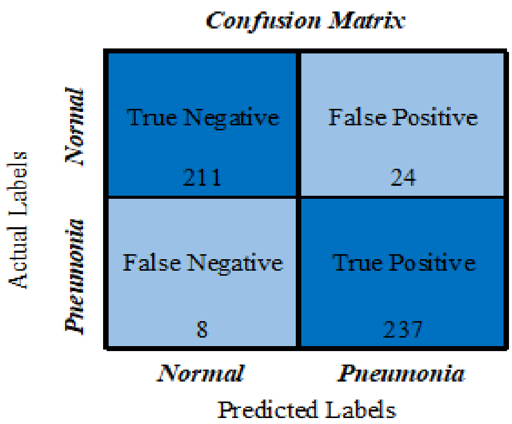
	Precision	*Recall*	*F1-score*	*Accuracy*
Normal	*0.97*	*0.89*	*0.93*	*0.93*
Pneumonia	*0.90*	*0.98*	*0.94*	

Fig. 13. Confusion matrix and Classification report for VGG19 Partial Frozen layer



	Precision	*Recall*	*F1-score*	*Accuracy*
Normal	*0.96*	*0.93*	*0.94*	*0.95*
Pneumonia	*0.93*	*0.96*	*0.95*	

Fig. 14. Confusion matrix and Classification report for VGG16 complete unfrozen layer



	Precision	*Recall*	*F1-score*	*Accuracy*
Normal	*0.96*	*0.90*	*0.93*	*0.93*
Pneumonia	*0.91*	*0.97*	*0.94*	

Fig. 15. Confusion matrix and Classification report for VGG19 complete unfrozen layer

C. *Advantages and Disadvantages*

The major advantages of this research are as follows:

1. For a medical diagnosis system, a false-negative rate plays a major role to determine the performance of the system. In our result, we achieved that the false-negative rate is low in comparison to the false-positive rate.
2. Even though this research was conducted with a limited dataset in a limited GPU environment, we achieved competitive accuracy (95%) with comparison to state-of-art models.
3. In this research, multiple cases of fine-tuning for pre-trained models were conducted to determine the better configuration of VGG architectures.

Even if there are some advantages, this research has following disadvantages:

1. A good medical diagnosis system should have a low false-positive rate, but this research has still a high false-positive rate.
2. Because of the limited dataset, the training phases are suffered from overfitting, even though regularization and dropout methods were used during training.

IV. Conclusion and Future Enhancement

In this research, VGG based architectures approach to chest X-ray classification has been proposed and the performance of the different models was calculated and compared. Two thousand one hundred and five chest X-ray images are utilized for training the model. Four hundred and eighty chest X-ray images are utilized for test the model. These images are chest X-ray images of normal and pneumonia patients. In this research, features are extracted from the images by using VGG pre-trained models, and the obtained features are used as input for the fully connected layers with SoftMax for classification.

This research used different fine-tuning cases for VGG pre-trained models. The VGG16 pre-trained model with a partially frozen layer was found to be very successful according to the false negative rate of pneumonia cases i.e.,4. The proposed method achieves 95% accuracy on chest X-ray images when VGG16 pre-trained model hascomplete unfrozenlayer with minimum learning rate.

In the future, more robust image preprocessing steps can be applied to enhance the features of pneumonia and normal X-ray images. Also, the performance of the proposed method can be improved by adding a number of training images for the training phases. In addition to this, the ensemble method can be applied after the feature extraction process.

References:

[1] Akgundogdu, “Detection of pneumonia in chest X-ray images by using 2D discrete wavelet feature extraction with random forest,” *International Journal of Imaging Systems and Technology*, vol. 31, no. 1, pp. 82-93, 2020.

[2] T. Rajasenbagam, S. Jeyanthi and J. A. Pandian, “Detection of pneumonia infection in lungs from chest X-ray images using deep convolutional neural network and content-based image retrieval techniques,” *Journal of Ambient Intelligence and Humanized Computing*, 2021.

[3] M. F. Hashmi, S. Katiyar, A. G. Keskar, N. D. Bokde and Z. W. Geem, “Efficient Pneumonia Detection in Chest Xray Images Using Deep Transfer Learning,” *Diagnostics (Basel)*, vol. 10, no. 6, 2020.

[4] B. K. Nyaupane and R. K. Sah, “Fine-tuning of ResNet Pre-trained Models for Diabetic Retinopathy Stages Classification,” *International Journal of Advanced Engineering*, vol. 4, no. 2, pp. 101-111, 2021.

[5] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” in *ICLR*, 2015.

D. Zhang, F. Ren, Y. Li, L. Na and Y. Ma, “Pneumonia Detection from Chest X-ray Images Based on Convolutional Neural Network,” *electronics*, vol. 10, no. 13:1512, 2021.