A Novel Method for Convolution Neural Network in Deep Learning for Detecting Tuberculosis Disease

Dr. Rajalakshmi R[1], Roobini.M.S.[2], Aman Deep Singh[3], Babu.K[4], Usha N.S.[5]
[1]Department of CSE
Sathyabama Institute of Science & Technology
Chennai, Tamilnadu, India
rajalakshmi.cse@sathyabama.ac.in
[2]Department of CSE Sathyabama Institute of Science & Technology
Chennai, Tamilnadu, India
roobinims.cse@sathyabama.ac.in
[3]Department of CSE Sathyabama Institute of Science & Technology
Chennai, Tamilnadu, India
amandeepsingh.cse@sathyabama.ac.in
[4]Department of CSE Sathyabama Institute of Science & Technology
Chennai, Tamilnadu, India
babu.cse@sathyabama.ac.in
[5]Department of CSE Sathyabama Institute of Science & Technology
Chennai, Tamilnadu, India
ushans.cse@sathyabama.ac.in

Article Info
Page Number: 1545-1556
Publication Issue:
Vol 71 No. 3s2 (2022)

Abstract

The common ancient disease widely prevailed in the world which is one of the top 10 main causing of crucial death is widely known as Tuberculosis. The chest x-ray which is the modest way for finding out tuberculosis is the model preparation of the main objective in this paper. Classification is done in this model to achieve high accuracy with attaining normal and abnormal classes infected by Tuberculosis. The techniques of deep neural networks is used in our study of approach in order to obtain and improve high accuracy of the model to meet the goal objective in classifying the new chest x- ray which is given as the input. Pre-processing of images are needed for achieving good accuracy and is obtained through the datasets achieved from both datasets of Shenzhen and Montgomery and together totally there are 800 chest x rays. Data augmentation is done over the 680 training set images and normalized them. It is followed by giving these pre- processed images as inputs to our models for supervised training. Then we performed model testing over the set of 120 images out of 800. In this paper we used two models. baseline CNN model and pre-trained VGG16 model and gave pre- processed images as inputs to these both models and evaluated the models to see which performed accurately better results. On comparing the results using different performance metrics like accuracy rate, specificity, sensitivity, high precision and f1- score. Finally it is depicted through using graphs and tables, where baseline CNN model gave an accuracy of 82% and the model VGG16 gave an accuracy of 90%.

Article History
Article Received: *28 April 2022*
Revised: *15 May 2022*
Accepted: *20 June 2022*
Publication: *21 July 2022*

Keywords: - Tuberculosis (TB), Chest X Ray(CXR), Computer aided detection, Shenzhen and Montgomery dataset, baseline CNN, pretrained VGG16, ImageNet dataset, Normalization, Data augmentation.

I. INTRODUCTION

The most ancient dreadful disease Tuberculosis is at the top ten position of causing crucial death across the world. Most people who get infected with tuberculosis can be saved with proper treatment but due to lack of medical support to detect tuberculosis almost various parts in the World which causes the mortality rate higher due to tuberculosis. According to WHO[1] ,TB (tuberculosis) is one of the top 10 causes of death across the world in 2018 where the people of 10.4 million fell ill due to Tuberculosis i.e. around 28,500 people per day out of them 1.8 million people have died i.e. around 4500 per day. The diagnosis of tuberculosis using accurate methods is one of the crucial steps involved to control the occurrence and prevalence of TB. However, the manual diagnosis of tuberculosis is quite complex these days, so there is no standard method at present. Hence, we aim to build a model which can recognize the Tuberculosis from the given chest x-ray with high accuracy.

II. RELATED WORK

The Research of various studies have been performed for the accurate detection of tuberculosis through a number of deep learning models in the past 10 years. Some are mentioned as follows. In the method followed by Bhuvaneswari, C Aruna in [2], different techniques are used for classification and feature extraction of images which are shown using mode like decision trees ,SVM.

In the method followed by Yaniv Bar, Idi Diamant, Lior Wolf, Hayit Greenspan in [3], depicted of deep learning method with non-medical training is given by the best performance using CNN and GIST features. Finally got the result of an AUC od 0.86-0.95 for different chest pathological chest. The method adopted by Paras Lakhani, Baskaran Sundaram in [4], finalised that Deep learning method with DCNNs can accurately classify Tuberculosis at radiography of chest with an AUC of 0.98.

The method adopted by Rahul Hooda, Sanjeev sofat, SimranPreet kaur, Ajay Mittal in [5], for the accurate detection of tuberculosis with radiography of chest using a CNN architecture method of 7 convolutional layers and fully connected layers of three gives an high accuracy rate of 82%.

The method adopted by Chunli Qin, Denim Yao, Yonghong shi, Zhijian shong in [6], for radiography of chest using Computer Aided Detection method which is basically connected Artificial Intelligence. It is concluded that in segmentation deep learning models are better compared to rule-based methods and in classification SVM and random forest are traditional Deep Learning methods have become mainstream.

Yin, Chuanlong in [7] performed binary and multiclass classification methods. The total number of neurons and different learning impact rates on the best performance of the proposed model. It is then compared with J48, Artificial Neural Networks, random forest, Support Vector Machines. The machine learning methods proposed by previous researchers proves onto the benchmark data set. The various experimental results show that RNN-IDS is actually suitable for classification method of modeling with the performance rate of higher accuracy.

III. PROPOSED METHODOLOGY

The Proposed model has the high performance of accuracy by using the method of Classification for Pre- processing of images. The perfect diagnosis of Tuberculosis disease is examined by the performance metrics by using ImageNet Dataset.

The various Performance metrics are used to do the comparison of results by achieving accuracy, specificity, sensitivity, precision and F1-score. By Pre-training VGG16 Model using Convolution layer, Activation Layer and Pooling Layer to obtain validation accurately.

A. Dataset

In this project we used two datasets Shenzhen and Montgomery[8]. Montgomery County dataset is source from "Department of Health and Human Services, Montgomery County, Maryland, USA". Shenzhen dataset has been collected from "Shenzhen No.3 People's Hospital".
The MC dataset includes 138 frontal CXRs out of which 58 cases are infected and remaining cases are with normal i.e. no TB infection.
The Shenzhen dataset includes 662 frontal CXRs out of them 336 cases are infected with tuberculosis, and the remaining cases are normal i.e. without infection of TB.

Each X-ray consists of a text file and a .png file text file contains details about X-ray and .png image is X-ray itself.

B. Preprocessing

Initially all the images are pre-processed before training the model all images are reduced to the same sizes for giving them as basic input to the given model.
For increasing the calculation of accuracy in the model Image data augmentation is performed (Image data augmentation is the process of increasing the number of images available from existing training data images because deep learning models perform better with large amount of data) initially a total of 800 images are available from both datasets. They both are segmented into validational training and testing sets. There are total of 680 images used for intial training over these 680 images. Data augmentation[9] is performed by using Image Data Generator class from Keras library[10] to do this with this total image count for training data used became 2040 images.
Generally, augmentation is performed for getting better accuracy as deep learning models provide better accuracy with large data and model will not be overfit i.e. it does not recognise only one type of images as given thus with augmentation model becomes robust and overfitting is removed. Images are also normalized[11] for getting better contrast i.e. we rescale the ranges of RGB values in images which is between 0-255 to 0-1 we do this by using Image Data Generator class and resize method of that class.

C. CNN Description

In this project we used two models first one is baseline CNN[12] and second one is pretrained Convolution Neural Network. This model contains ImageNet dataset is well trained. Both the models are convolution neural networks works by updating the weights and bias upon training. CNNs are mainly used for image recognition and classification. They are made up of perceptions of learnable weights and bias. Generally, the neural networks take input in the form of vectors. The input taken from the neurons is passed through the activation function and output if obtained. and layers in convolutional neural networks are:

- Convolution Layer –Convolutional Kernels whose depth needs to be equal to that of the image depth, and whose height and width are hyper-parameters. Convolutional layers take the input[14] and produce the output and this output is passed to the next layer. This is like the response given by neuron to a stimulus in visual cortex. Input is used to give an image shape it is tensor with form (image width) x (number of channels) based upon channel first or last.
- Activation Layer – Activation function[15] is the function which decides that a neuron should be activated or not. This is done by the calculation of adding weighted sum to the bias. The main purpose of activation function to obtain the output of neuron is by the introduction of the non-linearity. Rectified Linear Unit carries out non-linear operations and the function is

$$f(x) = \max(0,x)$$

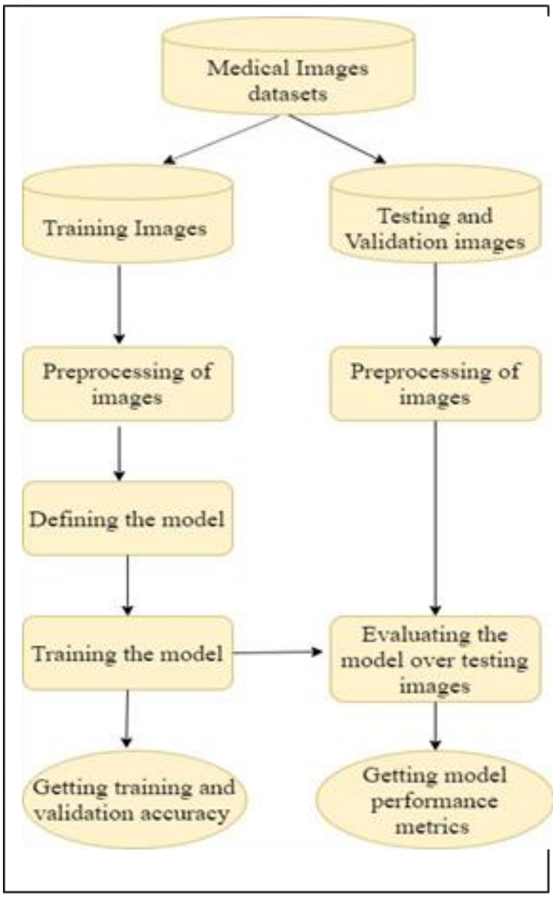

Fig. 1. Proposed Baseline of CNN model

- Pooling Layers - It works on the output of convolutional Layer and reduce the scale of feature map obtained from convolution layer while not distorting the knowledge. It divides the feature map into some parallelogram regions to perform pooling operation on that. It converges quicker, better, generalisation and sturdy. It's used at the top of convolution layer. There are 3 kinds of pooling are average, Max and Min Pooling. The ReLU activation function is similar to step function, which outputs 0 when the value is less than the threshold and 1 otherwise.
- Fully Connected Layers – In these dense layers, every neuron of one layer will be linked to every other neuron in the following layer i.e. if there are n, m neurons in two layers respectively then a total of n*m connections will be available on whole ion fully connected layers. This works on the same principle of multilayer perceptron neural networks. After flattened (decrement of conversion of higher dimensions to lower dimensions) the matrix it is given as input to a fully connected layer these layers then classify the images.
- Flatten Layer - It transforms 2-dimensional matrix into a linear vector so that it could be fed to fully connected layer. The concept of transfer learning proves to be helpful for enhancing the performance accuracy with a much less effort than the previous model. The VGG-19 is a very popular pre-trained CCN model due to a more in-depth architecture.

D. Model Description

Baseline CNN – In this model we used 9 convolutional and 2 dense layers and 3 max pooling layers initially to first convolutional layer input is given as an image of size 96*96*3 and output from max pooling layer after 3 convolutional layers is 45*45*32 which is taken as input to next convolutional layer and max pooling layer after that converts image and its output is 19*19*64 after this a flatten layer is present to reduce the dimensions of image to give it as input to dense layers and after two dense layers image is classifies into any one of two classes i.e. either infected or not. Kernel size used for this model is 3*3 and pool size used for this model is 2*2 and a dropout of 0.3 is used the total number of trainable parameters comes out to be 1,661,186.

Pretrained VGG16 – It is a 16 layered deep neural network consisting of convolution, fully connected layers in original VGG16 model is trained on ImageNet dataset which contains totally 1000 number of classes of images. VGG16 contains a set of convolution layers, dense layers, max pooling layers which gives input to VGG16 gives an image in the shape of 224*224. The above model contains 4 max pooling, 13 convolution and 3 fully connected layers or dense and activation function used is ReLU activation. The basic input of cov1 layer is through fixed size of 224 x 224 RGB image. This image is then passed through a full stack of convolution layers. Here the filters which were used is of with a compact and tiny receptive field.

The fully connected layers of three follows a full stack of convolution layers. The first of two layers which has 4096 channels of each. The third layer contains and performs 1000-way of classification methods. Therefore it contains 1000 channels of one layer for each class. The final layer is otherwise called as the soft-max layer. In this model we use we remove this final fully connected layers and replace them by two fully connected layers of 512 neurons and 2 neurons one for each class it needs to classify.

From original VGG16 model which is trained over ImageNet dataset all those weights are imported and used and only last two fully connected layers are changed remaining layers are made non trainable and remains same.

E. Model Evaluation Parameters

Confusion Matrix – It is a matrix representation of summarized predicted results [17] which are obtained from the evaluation of classification problems. There are four major classes for result prediction. They are

- True -Positive
- False -Positive
- True -Negative
- False -Negative

The model evaluation of various performance parameters [18] that could be defined from Confusion matrix valuesare:

Accuracy - Accuracy is defined as the actual proportion of correct estimates to the total number of sample estimates.

- Precision - Precision is defined as what fraction of predicted positives which are actually correct.
- Recall – Recall is considered as Sensitivity of a classifier which is the ratio between truepositive and truenegative. Here, both Recall and sensitivity are equal to same.
- F1-Score - F1-Score is given by the harmonic mean of recall and precision.

B. Training and Evaluation Models

1) Baseline CNN Model - In this model we used 9 convolutional and 2 dense layers and 3 max pooling layers initially to first convolutional layer input is given as an image of size 96*96*3 and output from max pooling layer after 3 convolutional layers is 45*45*32 which is taken as input to next convolutional layer and max pooling layer after that converts image and its output is 19*19*64 after this a flatten layer is present to reduce the dimensions of image to give it as input to dense layers and after two dense layers image is classifies into any one of two classes i.e. either infected or not. Kernel size used for this model is 3*3 and pool size used for this model is 2*2 and a dropout of 0.3 is used the total number of trainable parameters comes out to be 1,661,186. After defining model layers training phase starts and model is trained for 100 epochs over a total of 2040 augmented images and for testing and validation of model 120 images are used various performance metrics like specificity, accuracy, sensitivity and f1-score are calculated over the predictions made by model on these 120 imagesand results are compared between models.

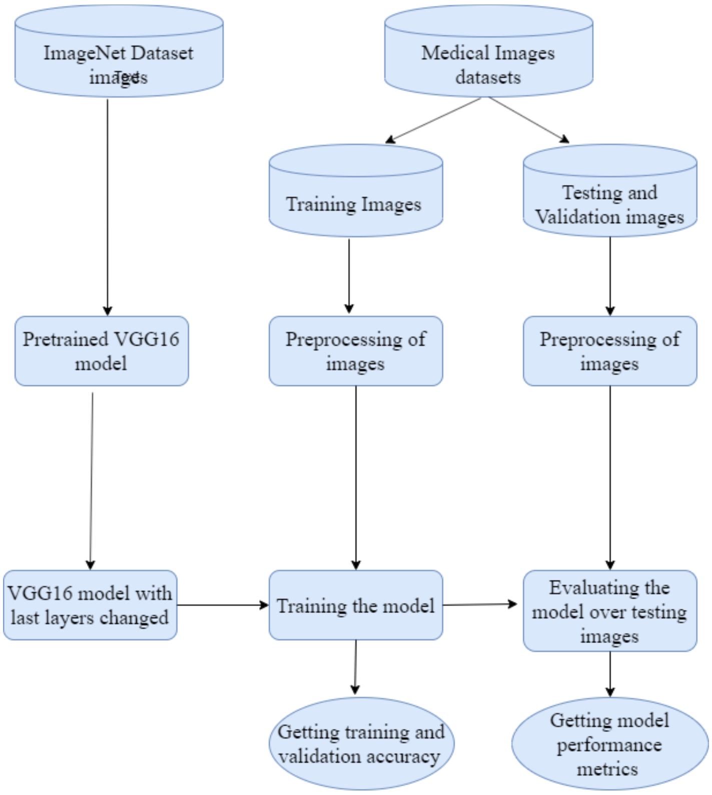

Fig. 2. Proposed VGG16 model

IV. EXPERIMENTATION RESULTS AND DISCUSSION

A. Preproccessing

Initially all the images from both Shenzhen and Montgomery are collected both together make 800 images and for these images dataframes are generated in python and a directory structure is created for base directory containing training, validation and testing images to divide the images amongst training and testing we used a test_train_split ratio of 0.15

i.e. 680 images for training and 120 images for testing the models then over 680 training images data augmentation is performed for increasing the number of training images count this is done by using ImageDataGenerator class of Keras library and properties used for generating object of ImageDataGenerator are:

- Rotation range: -10
- Width shift limit for images: - 0.1
- Height shift limit for images: -0.1
- Zoom: -0.1
- Horizontal flip

We then perform normalization over these images i.e. rescale the ranges of RGB values in images which is between 0-255 to 0-1 we do this by using ImageDataGenerator class and resize method of that class. Then generators are created from these images and given to model for best training and actual testing of the model.

2) *Pretrained VGG16 Model* - PretrainedVGG16 is a 16 layered deep neural network which consists of convolution, and the fully connected layers. The original VGG16 model is completely trained on ImageNet Dataset which contains 1000 number of classes of images. VGG16 model contains a set of convolutional, dense and max pooling layers which gives input to VGG16 model is taken as an image of shape 224*224. This model contains 4 maximum pooling and 13 convolution and fully 3 connected layers or dense layers and activation function used is ReLU activation. The final input to cov1 layer is about fixed size of 224 x 224 RGB image. The pretrained image which is passed through the full stack of convolution layers has tiny receptive fields in which the filters were used. The fully connected layers of three follows the full stack of convolution layers. The first of two layers should have 4096 channels of each. The third layer performs 1000-way of classification methods. This contains 1000 number of channels one in each of the class.
The final layer which is also named as the soft-max layer. In this model we use we remove this final fully connected layers and replace them by two fully connected layers of 512 neurons and 2 neurons one for each class it needs to classify. From original VGG16 model which is trained over ImageNet dataset all those weights are imported and used and only last two fully connected layers are changed remaining layers are made non trainable and remains same.In training VGG16 model we used images of both Shenzhen, and Montgomery datasets and images are augmented so total of 2040 images are used for training the model, model is run for 20 epochs we used Adam optimizer for training the model and also used binary cross-entropy loss. Baseline CNN shows performance analysis after model training which is shown below, wo plots are drawn: 1) Number of Epochs versus Accuracy value and 2) Number ofEpochs versus Loss value.

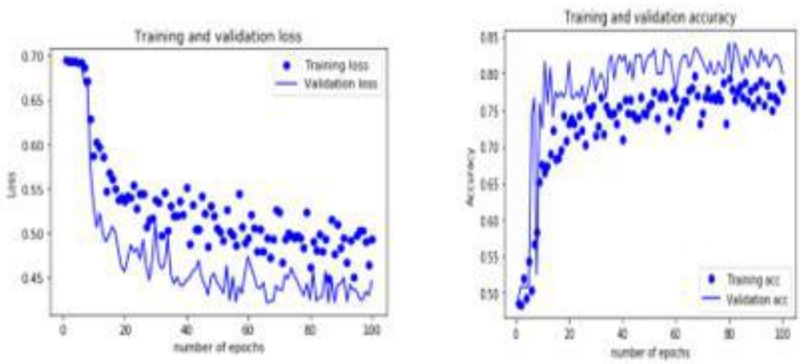

Fig. 3. Performance of CNN model

As the rate of accuracy value for both training and validation set is increasing for epochs and loss is decreasing with number of epochs this shows baseline Convolution Neural Network model is not overfitted which gives good accuracy rate. After evaluating the confusion matrix model which depicts an accuracy rate of 82% and the final precision, sensitivity or recall, the f1-score and specificity states 0.81, 0.82, 0.871 and 0.81 respectively.

ISSN: 2094-0343
2326-9865

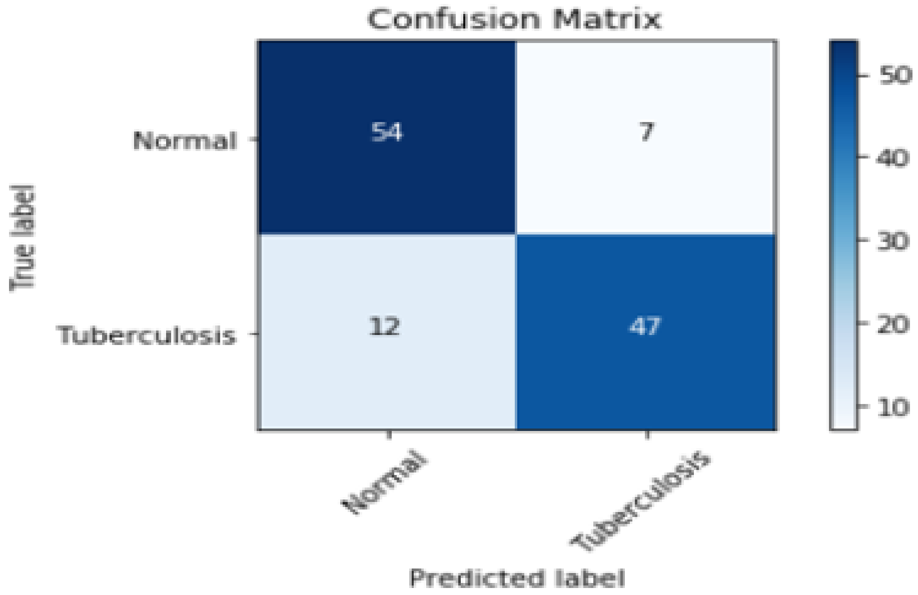

Fig. 4. Confusion Matrix from Baseline CNN model

Similarly, performance of VGG16 model is also evaluated and plots are mentioned below

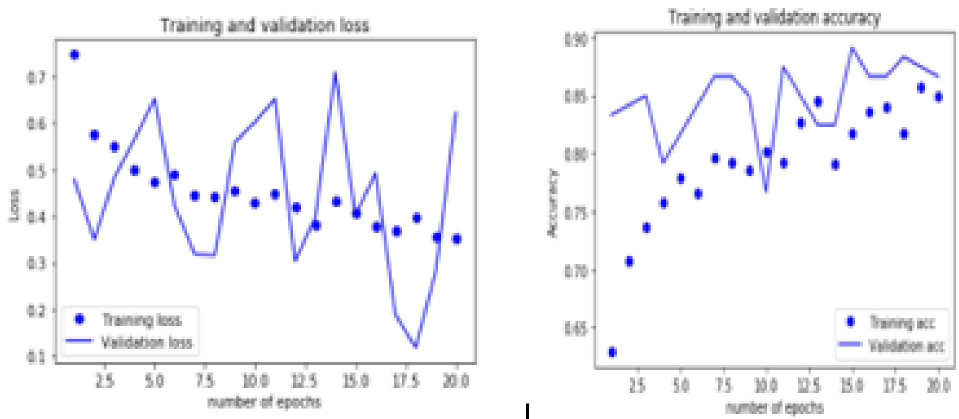

Fig. 5. Performance of VGG16 model

After evaluating the confusion matrix model which shows an accuracy rate of 90% and final precision, sensitivity or recall, the f1-score and specificity states 0.87, 0.877, 0.87 and 0.85 respectively.

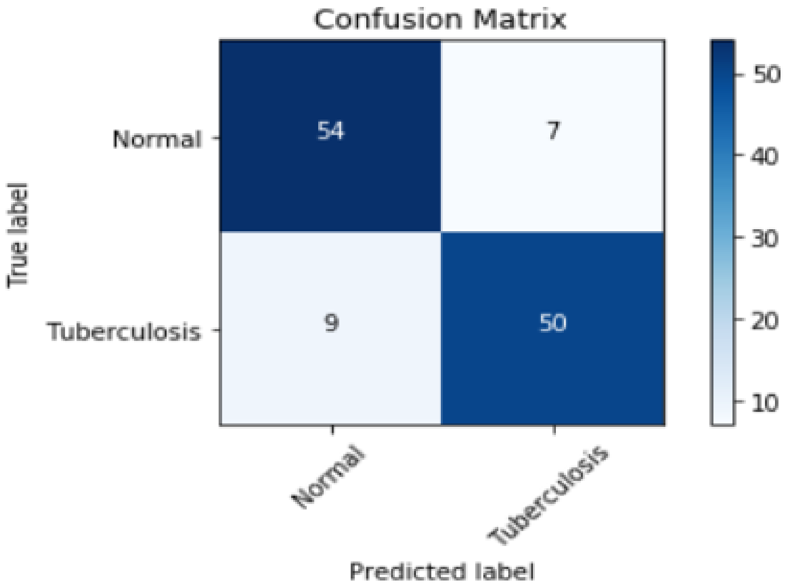

Fig. 6. Confusion Matrix of VGG16 model

The actual Performance of both VGG16 model'scomparison is shown below.

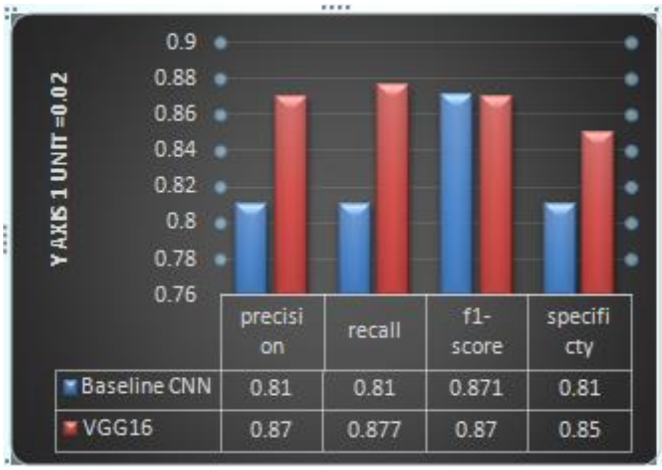

Fig. 7. Performance Metrics Comparison Graph

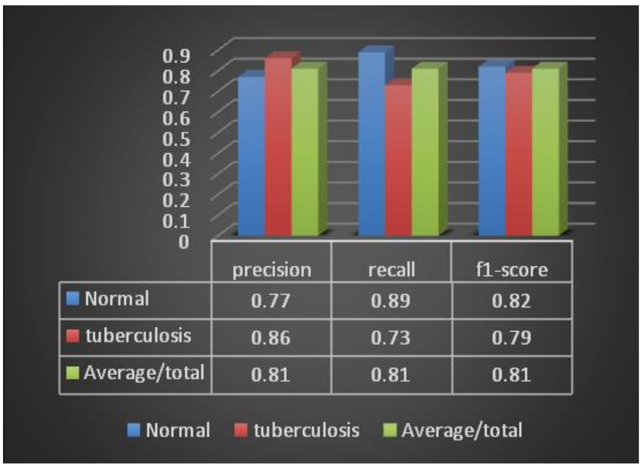

Fig. 8. Classification Report for Baseline CNN model

The above figures shows different performance parameters for baseline CNN model for both normal and tuberculosis.

V. CONCLUSION AND FUTURE WORK

Having large amounts of data is necessary for good accuracy of deep neural networks as we can see from results of both model's baseline CNN and VGG16, VGG16 performed better than baseline CNN as VGG16 is already trained over ImageNet dataset which is a large dataset of images. Thus, pretrained models perform better than completely customized deep neural networks, in conclusion with more data and better preprocessing the performance of models can still be increased. Thus the dreadful disease Tuberculosis can be found at the intial stage itself by the diagnosis of accuracy through considering various performance metrics with the help of confusion matrix from Baseline Convolution Neural Network Model.

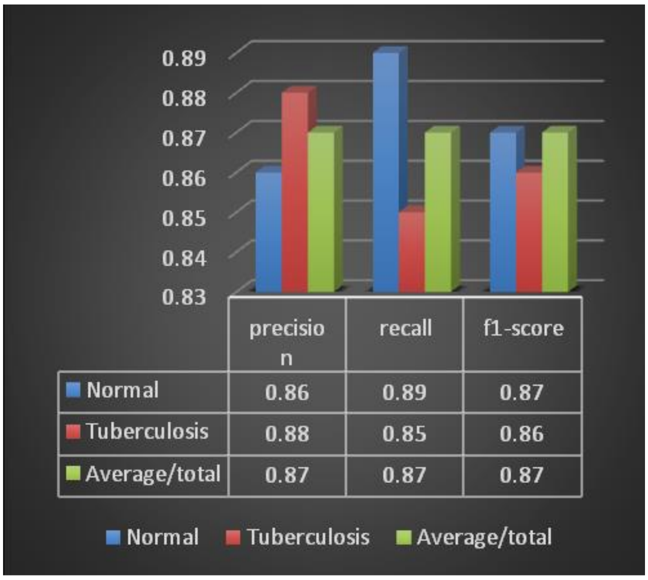

Fig. 9. Classification Report for VGG16 model

The above figure shows different performance parameters for VGG16 model for both normal and tuberculosis.

REFERENCES

[1] World Health Organization, Global tuberculosis report,World Health Organization, http://www.who.int/iris/handle/10665/274453, ,2018

[2] Bhuvaneswari, C., Aruna, P., and Loganathan, D., "Classification of Lung Diseases by Image Processing Techniques Using Computed Tomography Images", Int. J. Adv. Comput. Res. 4, 2277–7970, 2014.

[3] Bar, Y., Diamant, I., Wolf, L., and Greenspan, H., "Deep learning with non-medical training used for chest pathology identification",Medical Imaging Computer-Aided Diagnosis, 2015.

[4] Lakhani, P., and Sundaram, B. , "Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks", Radiology 284, 574–582,2017.

[5] Hooda, R., Sofat, S., Kaur, S., Mittal, A., and Meriaudeau, F. "Deep- learning: A potential method for tuberculosis detection using chest radiography", Proc. 2017 IEEE Int. Conf.

Signal Image Process. Appl. ICSIPA , 497–502, 2017.
[6] Qin, C., Yao, D., Shi, Y., and Song, Z. , "Computer-aided detection in chest radiography based on artificial intelligence: A survey", Biomed. Eng. Online 17, 1–23. Available at: https://doi.org/10.1186/s12938- 018-0544-y.
[7] Yin, C., Zhu, Y., Fei, J., and He, X., "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks", IEEE Access, 2017.
[8] Shenzhen and Montgomery datasets Available at: https://ceb.nlm.nih.gov/repositories/tuberculosis-chest-x-ray-image- data-sets/.
[9] Karnkawinpong, T., and Limpiyakorn, Y., "Classification of pulmonary tuberculosis lesion with convolutional neural networks", In Journal of Physics: Conference Series, 2019.
[10] Kears library documentation , Available at: https://keras.io/, 2019.
[11] Avants, B., Tustison, N., and Song, G., "Advanced Normalization Tools (ANTS)", Insight J, 2009.
[12] Razavian, A.S., Azizpour, H., Sullivan, J., and Carlsson, S., "CNN features off-the-shelf: An astounding baseline for recognition", IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 512–519.
[13] Jmour, N., Zayen, S., and Abdelkrim, A., "Convolutional neural networks for image classification", International Conference on Advanced Systems and Electric Technologies, IC_ASET, 2018.
[14] Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D.D., and Chen, M., "Medical image classification with convolutional neural network, 13th International Conference on Control Automation Robotics and Vision, ICARCV, 2014.
[15] Liu, J., Liu, Y., Wang, C., Li, A., Meng, B., Chai, X., and Zuo, P., "An original neural network for pulmonary tuberculosis diagnosis in radiographs", Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).
[16] Krizhevsky, A., Sutskever, I., and Hinton, G.E., "ImageNet classification with deep convolutional neural networks. Commun. ACM 60, 84–90, 2017.
[17] Sivaramakrishnan, R., Antani, S., Candemir, S., Xue, Z., Thoma, G., Alderson, P., Abuya, J., and Kohli, M., "Comparing deep learning models for population screening using chest radiography, p. 49, 2018.
[18] Oquab, M., Bottou, L., Laptev, I., and Sivic, J., "Learning and transferring mid-level image representations using convolutional neural networks", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1717– 1724, 2014.