Lung Disease Prediction And Segmentation Using Deep Learning

Vidhi Pitroda
Master in Computer Science
Lakehead University
Thunder Bay, ON

***Abstract*—Machine learning plays a significant role in medical system as early and accurate detection of disease can save many people lives. Lung disease such as Tuberculosis (TB) is a chronic infectious disease is a major cause of death worldwide. For the diagnosis, chest x-ray is performed to detect TB. To predict the disease from the patient's chest x-ray, many machine learning approaches are developed giving decent results. In this project, two small datasets i.e. Montgomery County and Shenzhen dataset are considered. The task performed is detection of tuberculosis disease from the chest x-ray images. I have used feature combination approach for classification. A pipeline is created, the images are first segmented using U-net algorithm and then are classified with feature combination method. In this method features are extracted using VGG-16 and ResNet 50 methods with predefined weights. These extracted features are combined and then are trained on fully connected dense layers. There are 3 different feature combination models are prepared of VGG16, ResNet 50, and InceptionV3. Furthermore, feature combination of CNN + CNN is also performed.**

***Index Terms*—Deep neural networks, Transfer learning, Segmentation, Feature combination**

I. Introduction

Tuberculosis is fifth leading disease causing death worldwide with about 10 million new cases and 1.5 million deaths per year. It is commonly caused by bacteria which is known as Mycobacterium tuberculosis and mostly affected to the lungs of the human beings. TB is spread through the air from everyone or everywhere [1]. Early detection and taking treatment from the physicians are the best prevention for TB disease. Chest X-Ray is one of the most frequently used diagnostic modality in detecting lung diseases such as pneumonia or tuberculosis. With the advancement in technology medical imaging can be used in medical treatment and analysis. It can be defined as a process and technique of making optical representations of the interior of a body. Medical imaging along with the detection can also be used to diagnose different disease. With the development of deep learning the disease can be detected early stage and can be very helpful.

It is very significant to make a deep learning model which can give good performance metrics based on less amount of data. The main motivation behind this project is to make a neural network which can give good results for classification of x-ray images considering the small size of data. The pipeline contains two parts segmentation and classification. Initially the images are segmented using U-net neural network architecture which works with small dataset and give precise segmentation. The unit results in capturing important information about lungs. In the next classification step, feature combination method is used with deep neural network models. The VGG16 and resent 50 models are used. Features are extracted from both the models and then these features are combined and trained on fully connected layers. To compare the results the feature combination of different models are also performed and transfer learning model. Feature combination gives good result as compared to transfer learning model.

The contribution of the project can be summarized as follows:

- Performed segmentation of chest x-ray using U-net neural network
- Proposed a new feature combination method for classification

The paper is organized as follows Section II presents related work on lung segmentation and disease classification. Section III elaborates about the methodology of present work with the details of model architecture, segmentation, configuration of model and proposed model. Section IV discuss the results and Section V concludes the paper and provides future directions.

II. Literature Review

Medical imaging can give significant information for diseases, and it is an indispensable component for disease diagnosis and treatment[1]. With medical imaging playing a considerable role, doctors monitor and analyze the occurrence, development, and feedback of treatment. Chest X-ray has been used to find diseases such as tuberculosis, pneumonia, and lung cancer[2].

The paper “Deep Learning Based Multi-Label Chest X-Ray Classification with Entropy Weighting Loss [2]” explains about diagnosis of chest x-ray with multilevel classification. They gave a new approach of inter-label dependencies. The classes having instances less than other were taken into consideration. For the improvement of model performance entropy weighting loss function is applied. Furthermore deep CNN model i.e. denseness101 is developed for classification of tasks depending on labels where the labels cannot wholly determine images.

In the paper “ Advances in automatic tuberculosis detection in chest x-ray images [3]” the authors presented a survey of different approaches for disease detection and classification with chest radiography in tuberculosis screening. Medical imaging is essential for detection and diagnosis of lung diseases. Image preprocessing is applied to reduce noise, meet the requirements of physician for the image quality, correction of wrong pixels and contrast enchantment. After preprocessing segmentation performs clustering of pixels having same intensity, separating regions or desired part. Different segmentation techniques were discussed in the paper such as graph cut based lung segmentation, watershed segmentation, otsu segmentation, active shape model and multi segment active shape model. In the feature extraction geometry and texture features are extracted. Furthermore in the classification different algorithms were explained such as SVM, ART neural networks and decision tree.Thus, the authors elaborated the steps of TB classification from chest X-ray by preprocessing, feature extraction and classification.

Diving into deep into deep learning algorithms and segmentation, the paper ‘Reliable Tuberculosis Detection using Chest X-ray with Deep Learning, Segmentation and Visualization [4]’ uses different transfer learning approaches to classify the lung disease. The authors created a database of 700 TB infected and 3500 normal chest X-ray by combining different publicly available datasets. The methods used were image pre-processing, data augmentation, image segmentation, and deep-learning classification techniques. U-net algorithm was used for segmentation and achieved a decent accuracy. Classification was performed on three shallow networks and 6 deep networks. The shallow networks includes MobileNetv2, SqueezeNet and ResNet18 and deep networks includes Inceptionv3, ResNet 50, ResNet101, CheXNet, VGG19 and DenseNet201. In total nine CNN models having pre defined weights were trained and 15 back propagation epochs and cross validation were used for classification problem. The performance metrics were ROC curve, confusion matrix and evaluation matrices.

In the paper “Classification of Lung Diseases Using Deep Learning Models [5]” the authors addressed the problem of medical data scarcity by considering small volume dataset of chest x-ray. Three different deep convolutional neural networks were implemented for lung disease classification i.e. VGG16, ResNet-50 and InceptionV3 with pre trained weights of imagenet. A new framework was created where the images are first segmented before passing to classification model. For segmentation they used U-net neural network architecture by changing few layers at the end. It was observed that InceptionV3 gives good results as compare to any other transfer learning models used. Furthermore, the accuracy was found to be improved by passing segmented images to classifier rather than passing raw images.

The paper ‘Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks[6]’ evaluated deep learning techniques for detecting tuberculosis on chest radiograph. For classification four HIPAA compliant datasets were used consisting 1007 posteroanterior chest radiographs. AlexNet and GoogleNet transfer learning approaches were used with weights and without pre trained weights. Augmentation and preprocessing of the images improved accuracy which gave good results of almost 90 percent AUC score. Along with AUC score Delong method was also used for statistical comparison of receiver operating characteristic curves.

III. METHODOLOGY

A. Dataset Description

In this project 2 small datasets are taken I.e. Montgomery County and the shenzhen dataset [7]. The Montgomery county dataset is the standard digital image database for Tuberculosis is created by the National Library of Medicine in collaboration with the Department of Health and Human Services, Montgomery County, Maryland, USA. The set contains data from X-rays collected under Montgomery County‘s Tuberculosis screening program. In that total are 80 normals cases and 58 cases tuberculosis. Each image is of size 4020 x 4892 pixels. The notation in file 0 represents normal lung x-ray while 1 represents abnormal lung [7].

Shenzhen dataset is The database of chest x-ray is created by the National Library of Medicine, Maryland, USA in collaboration with Shenzhen No.3 People’s Hospital, Guangdong Medical College, Shenzhen, China. There are total 336 normal chest x-ray and 336 abnormal chest x-ray. Each image is in png format and having pixel size 3k x 3k[7]. The dataset is not very balanced external segmentation of lungs can improve the prediction accuracy. To extract the lungs information by eliminating outside lung information manually prepared mask are used.

B. Data Preprocessing

Every image file in dataset is between 6-7 MB of size. There are in total 704 mask and 800 images so one to one correspondence is done from images to mask or vice-versa. Initially every file and mask is read and stored in two different array after resizing the images with size 256 x 256 pixels. One array stores images while the other array stores its corresponding masks. While reading the dataset is divided into training and testing part. From the training and testing files labels are seperated and are stored in different list as it will be needed later for transfer learning. This preprocessing operation of resizing and storing the image data takes approximately 30 minutes.

C. Software and Hardware:

For transfer learning the code of algorithms are available on Keras documentation. The models were trained on google colab which is a free online cloud based Jupyter notebook environment used for training machine learning or deep learning models on GPU or TPU.

D. Segmentation of Lung images

For medical image processing it is expected to have a class assigned to each pixel. Thus, every pixel should be associated with every class. For this purpose U-net neural network architecture is used for segmentation. In U-net neural network CNN is used which is fast and precise segmentation of images. The convolutional layers in U-net are formed in top down and bottom up forming a U-shape network.

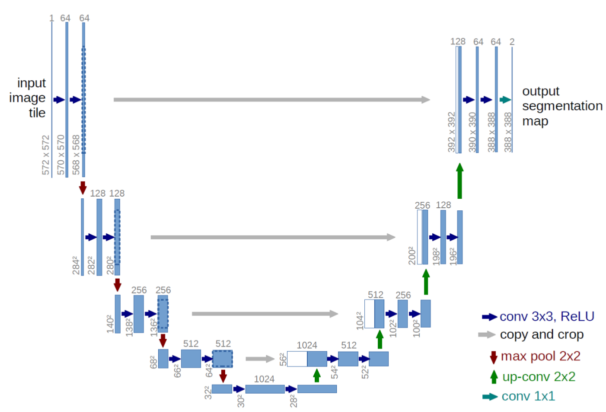

Fig. 1: Unet-architecture

This architecture in fig.[1] has two path followed from upward to downward and vice versa. The top down is called contracting section and bottom up is called expansive section. In contraction significant information of images are captured. While in expansion is used to localize the region of interest. The contracting section has in total 4 subsection each section consists of two convolutional layers having activation function as relu and max pooling layer. The expansive section consist of 4 sub sections. The first two sub sections comprises of convolutional layer, concatenation layer and upsampling layer. The last block has three convolution, upsampling, concatenation layers. The last section has dropout and output layer[8].

Layer	Filter	Dim	Concat
Input	-	256x256	-
conv1	32	256x256	input
conv1	32	256x256	conv1
pool1	32	128x128	conv1
conv2	64	128x128	pool1
conv2	64	128x128	conv2
pool2	64	64x64	conv2
conv3	128	64x64	pool2
conv3	128	64x64	conv3
pool3	128	32x32	conv3
conv4	256	32x32	pool3
conv4	256	16x16	conv4
pool4	256	16x16	conv4
conv5	512	16x16	pool4
conv5	512	16x16	conv5
up6	256	32x32	conv5 +conv4
conv6	256	32x32	up6
up7	128	64x64	conv6 +conv3
conv7	128	64x64	up7
up8	64	128 x 128	conv7 +conv2
conv8	64	128 x 128	up8
up9	32	256 x 256	conv8 +conv1
conv9	32	256 x 256	up9
conv10	1	256x256	conv9

Table1: configuration of U-net

The following table [1] represents the configuration of U-net architecture. There are 5 columns layer name, filter applied, pooling size, dimensions and concatenated together. In total there are 13 convolutional layers. The first layer is the input layer where the image size is 32 x 32 which takes 8 filters of size 3x3. con1 is joined with another layer with relu activation function followed by this in entire network. Then next pooling layers are used which reduces the dimensions of the image. From pool4 expansive path begins, conv5 has 32 filters and is concatenate with conv4. Upsampling up1 follows concat yielding increased image of size 8x8. The next upsampling layers 4 and 5 increases the size of image.

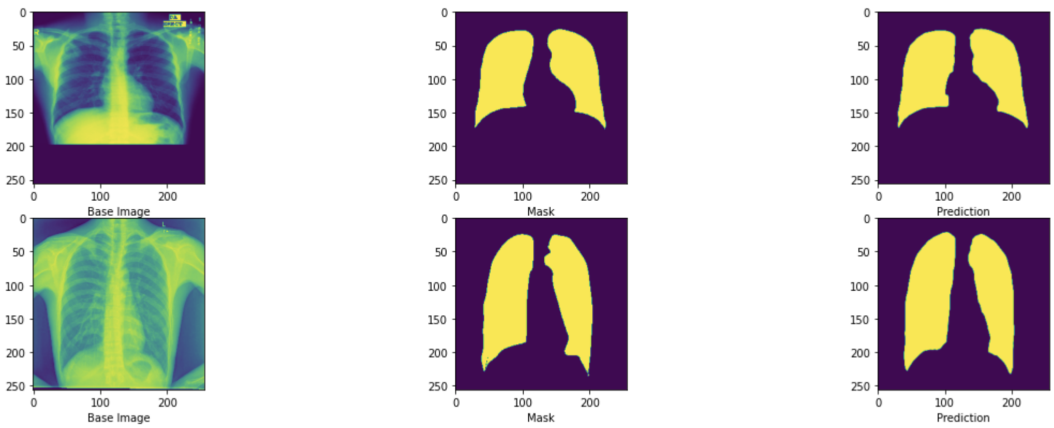

Fig. 2: U-net neural network prediction on test data

The above fig.2 shows the lung image prediction of segmented image and its corresponding lung mask. The performance metrics set was binary accuracy and loss. When the image is passed in the input layer the convolutional, non-linearity and downsampling layers are initiated. Each operation is applied which results in reduces size of image.Then after concatenation of corresponding layers in contracting and expanding path is performed which increases the size of image. Thus output layer generates the image with segmented lung region.

E. Transfer learning

Transfer learning focuses on storing the knowledge gained while solving one problem and applying it to a different problem but the goal is similar. It uses similar type of dataset type. For instance considering a transfer learning model where the model knowledge is used to recognize cars the same model can be applied to recognize cats. Transfer learning is popular approach when the data are in image form by using deep neural networks. This approach is difficult to use in medical image due to the scarcity of real samples. The feature extraction in deep neural network is done by passing raw data through models specialized in other tasks. In transfer learning models pretrained weights are used normally [5].
A pre-trained model can be explained as a model which was trained on a large dataset to solve a problem. So the weights of that model can be used to solve a problem of similar type. It will reduce computational costs of training from scratch. While using the pre trained models some changes needs to be done such as removing the original classifier and changing it with new one which fits the model. Depending upon the end result some layers can be trained while some can be kept frozen and fine tuning the model. There are different transfer learning models used in this project are VGG16, VGG19, ResNet 50, ResNet101 and InceptionV3.

VGG16 model comprises of 16 convolutional layers which are trained on a fixed size image while VGG19 has 19 layers. This network was first introduced by Simonyan and Zisserman in their paper[9] where the network is made using 3x3 convolutional layers stacked on top of each other in increasing depth. The dimensions are reduced by using Max pooling layers. At the end there is two fully connected layers having 4096 nodes followed by a classifier. VGG has smaller architecture which is desirable in classification problems[8].

Resnet relies on microarchitecture modules such as convolutional, maxpooling and many more. With the use of residual networks ResNet demonstrates deep networks trained using standard SGD. Resnet models are much deeper and smaller than VGG models because of using global average pooling compared to fully connected layers. Thus, even after the 50 and 101 layers in resnet models the size is 102 MB which is less than VGG[10].

InceptionV3 model is designed as a multilevel feature extractor by calculating 1x1, 3x3 and 5x5 convolutions within the same module of network. The output obtained from described filters are stacked with the channel dimensions and then passed to the corresponding layer in the network. The weights of this model is small as 96MB compare to ResNet and VGG.

IV. RESULTS

As explained in the methodology section the images are first resized to 256x256 and then passed to model. The shape if input image is (256,256,1) the dataset is splitted into 80:20 ratio and passed to the U-net model. After training the U-net model the data are predicted to get the segmented images. The predicted images from U-net are in gray scale so they are converted to RGB as pretrained models takes 3 channel input. These segmented images are further passed to feature combination model. In feature combination 3 separated models are prepared InceptionV3 and VGG16 model, ResNet50 and ResNet101 and VGG16 and ResNet50. These models are then trained on the following parameters.

Parameters	Values
Epochs	70
Epoch(U-net)	10
Batchsize	2
loss	Binary-crossentropy
optimizer(feature c)	RMSProp
optimizer(U-net)	adam(1e-5)

Table 2: Parameters

The above table 2 shows the parameters set for training model. The labels from training and testing data are separated and are converted to categorical. The combined features are trained on 4 fully connected layers.

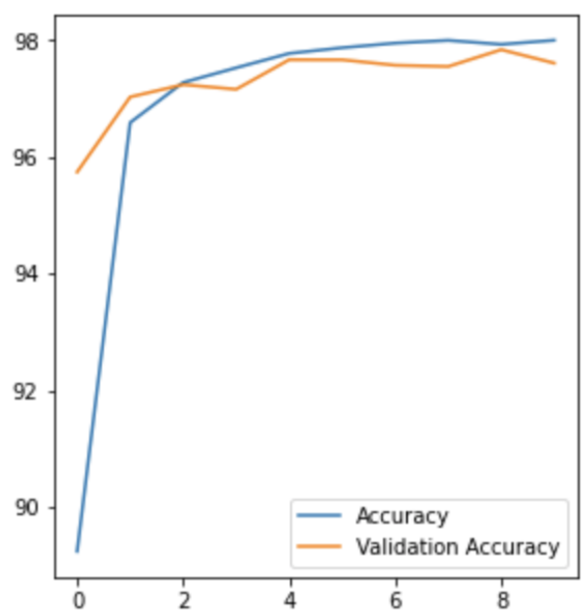

Fig. 3: Unet-Training

It can be observe from fig:3 that the model is trained for 10 epochs with the parameters defined in table 2. The validation accuracy got 97 percent.

Model	Accuracy	Loss
ResNet50 + ResNet101	77	63.2
VGG 16 + ResNet50	80.85	48.7
ResNet50+ InceptionV3	77.7	63.4
CNN+CNN(without weights)	54.61	79.2

Table2: Comparison of Feature combination models

From Table 2 it can be noted that feature combination of VGG16 and ResNet50 is the best model giving 80.85 percent accuracy. Feature combination of CNN model is also performed without using weights gives 54.61 percent accuracy. Thus, Feature combination of transfer learning models gives good results.

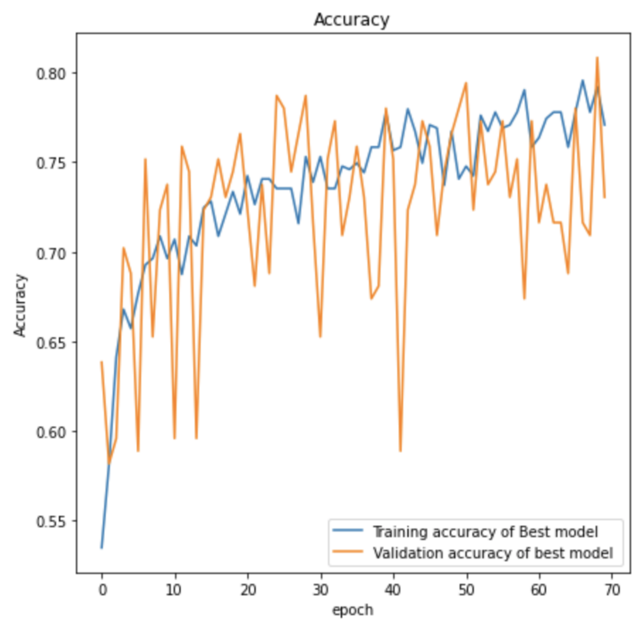

Fig. 4: Accuracy graph of VGG16 + ResNet50

From Fig 4. the model was trained for 70 epoch and initial accuracy obtained was from 50 and after increased gradually along with the validation accuracy.

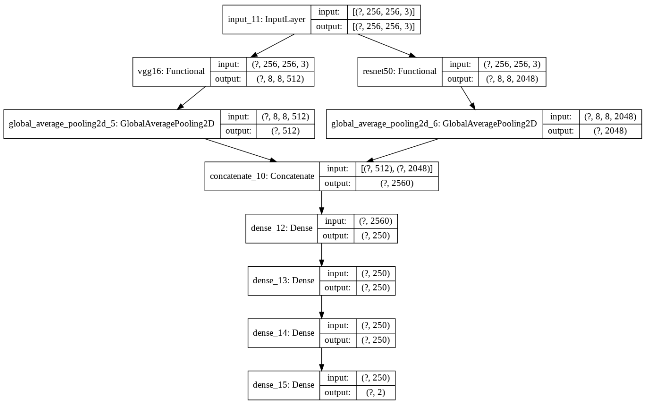

Fig. 5: VGG+ResNet model architecture

The above fig 5. shows the model architecture of best model the features from both the models are extracted and passed to global average pooling function which computes the mean value and passes it to fully connected layers. The features are combined and then are passed through fully connected dense layers. The last layer contains softmax as a classifier.

V. Conclusion and Future work

In this project a feature combination method for classification is proposed. The images are first segmented and then segmented images are passed to transfer learning models. The features of both transfer learning models are combined and then are trained on fully connected layers.The feature combination of ResNet50 +ResNet101, VGG16 + ResNet 50, ResNet50 +InceptionV3 and CNN + CNN(without weights). The model VGG16+ResNet50 gave a best accuracy of 80.85 percent in 70 epochs.The model can be improved by changing the architecture of U-net and passing original size of image. Furthermore, using real time data the can be improved as the datasize is large.

References

[1] Golkov V. Pfeiffer F. et al. Pasa, F. Efficient deep network architectures for fast chest x-ray tuberculosis screening and visualization. *Sci Rep 9*, (6268), 2019.

[2] Shaocong Mo ; Ming Cai. Deep learning based multi-label chest x-ray classification with entropy weighting loss. *12th International Symposium on Computational Intelligence and Design (ISCID*, 2019.

[3] Zaw Z. Htike wai Yan Nyein Naing. Advances in automatic tuberculosis detection in chest x-ray images. *Signal Image Processing : An International Journal (SIPIJ)*, 5(6), 2014.

[4] Tawsifur Rahman, Amith Khandakar, Muhammad Abdul Kadir, Khandaker R. Islam, Khandaker F. Islam, Rashid Mazhar, Tahir Hamid, Mohammad T. Islam, Zaid B. Mahbub, Mohamed Arselene Ayari, and Muhammad E. H. Chowdhury. Reliable tuberculosis detection using chest x-ray with deep learning, segmentation and visualization. 2020.

[5] Krzyżak A. Zak M. Classification of lung diseases using deep learning models. *International Conference on Computational Science*, (621-634), 2020.

[6] Baskaran Sundaram Paras Lakhani. Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. *Original Research Thoracic Imaging*, 2017.

[7] Antani SK. Tuberculosis chest x-ray image data sets.

[8] H. Shaziya, K.Shyamala, and R.Zaheer. Automatic lung segmentation on thoracic ct scans using u-net convolutional network. *International Conference on Communication and Signal Processing (ICCSP)*, pages 0643–0647, 2018.

[9] Andrew Zisserman Karen Simonyan. Very deep convolutional networks for large-scale image recognition. *Computer Vision and Pattern Recognition*, 2014.

[10] Adrian Rosebrock. Imagenet: Vggnet, resnet, inception, and xception with keras.