**X-Ray Image Classification using CNN**

Kartik Aggarwal Soumyo Dey Faiz Khwaja

*Northeastern University Northeastern University Northeastern University*

*(aggarwal.kart,dey.soum,khwaja.f)@northeastern.edu*

**ABSTRACT**

X ray and medical reports are an essential part of any healthcare system. The quick and accurate diagnosis of these X-ray images is critical for smooth and timely functioning of the healthcare system. If these medical images are not handled properly, problems such as inaccurate interpretation of X-ray images can occur. In order to achieve the best results in the healthcare system new technologies should be tested and brought into use. A convolutional neural network (CNN) is a type of artificial neural network used in image recognition and processing that is specifically designed to process pixel data. In this report we have developed a CNN model to classify pneumonia and brain tumor in X-ray images. The classification process consists of two main steps. First, the images were pre-processed by using different image processing techniques and then these images were classified using the CNN model. CNN can be very useful in order to improve the efficiency of the traditional health care system by giving accurate diagnosis of the medical images and in very less time. In this report we will help expand our knowledge in the application of CNN model to classify X-ray images. For a given image of the brain, the model will detect if the image has a tumor. Similarly, for a given image of the chest, it will check for the presence of pneumonia.

**1. INTRODUCTION**

Since the past few decades image processing and computer vision have helped many medical personnels for the identification of various diseases through automated diagnostic processes. In this project we tried to solve X-Ray image classification using a CNN model. An X-ray is a common imaging test that's been used for decades. It can help your doctor view the inside of your body without having to make an incision. This can help them diagnose, monitor, and treat many medical conditions. Different types of X-rays are used for different purposes. The purpose of this project is to develop automated methods to aid doctors in diagnosis in order to prevent misdiagnosis and decrease patient wait time.

First dataset we are using is the Chest X-ray image dataset. Chest diseases are one of the major health problems, in particular pneumonia is extremely dangerous for people already suffering from other diseases, infants, and older adults. Pneumonia has been diagnosed by radiologists by performing sophisticated radiological investigation on them. The computerized technique has been adopted because they give more precise results and are easily accessible for diagnosis. The image processing technique has a powerful ability to detect various objects together, extract deep features and classify them and therefore it is popularly the initial step achieved by convolutional layers of CNN.

The brain is the management center of the central nervous system and is responsible for the execution of activities all throughout the human body. Brain tumors can threaten human life directly. If the tumor is detected at an early stage, the patient’s survival chance increases. The other dataset we are using consists of brain x-rays comprising both tumor affected brains and tumorless brains. By including images of brains without tumors, neural networks can better learn the structure of a brain and take steps towards differentiating brains with and without tumors. More generally, this differentiates physiological structures through deep learning.

There are many different kinds of neural networks that have been used to research X-Ray images. There are mainly three types of neural networks that have been used for classifying X-Ray images - Fully Connected Neural Networks (FCNNs), Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). In this project we propose a X-ray image classification based on a Convolutional Neural Networks. CNNs consists of a convolutional network to perform automatic segmentation and feature extraction, followed by a conventional neural network to perform classification tasks.

**2. PROPOSED WORK**

**2.1 Dataset and labels**

We have worked on two datasets.

**2.1.1 Chest X-Ray Images (Pneumonia)**

We have taken an x-ray image dataset consisting of 5870 images. We have further divided it into three parts, Train, Test and Validation respectively, each of them consists of two parts Normal and Pneumonia. Normal images contain X-ray images of lungs free of pneumonia and Pneumonia contains images of lungs which are infected. The train dataset contains a total of 5230 images, the test dataset contains 624 images and the validation dataset contains 16 images , all of which belong to the two classes.

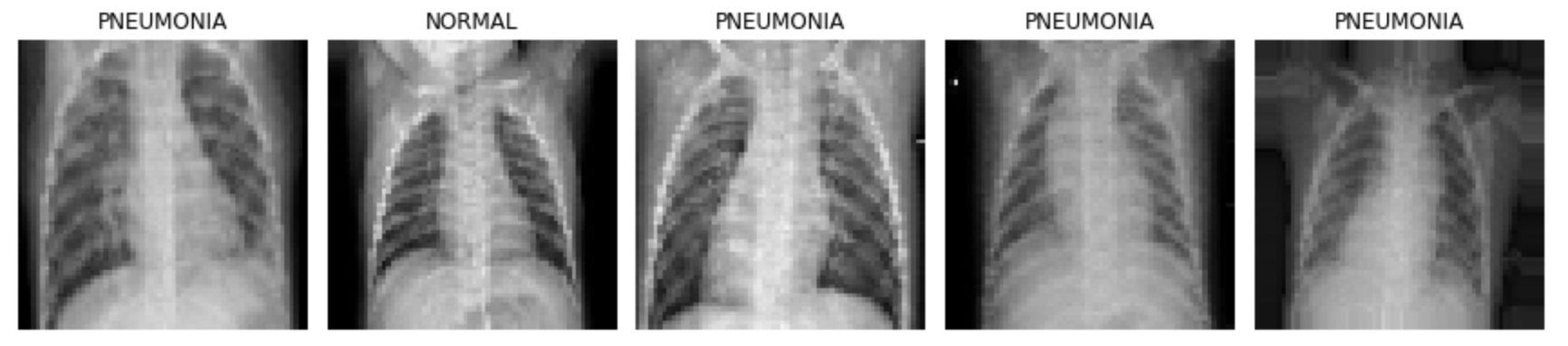


Fig 1. Chest X-Ray dataset Sample

**2.1.2 Brain Tumor Dataset**

We have taken a brain tumor image dataset consisting of 4600 images. We have further divided it into three parts, Train, Test and Validation respectively, each of them consists of two parts Normal and Unhealthy. Normal images contain Brain X-ray images free of tumor and Unhealthy contains images of brain which are damaged with tumor. The train dataset contains a total of 3249 images, the test dataset contains 853 images and validation dataset contains 498 images , all of which belong to the two classes.

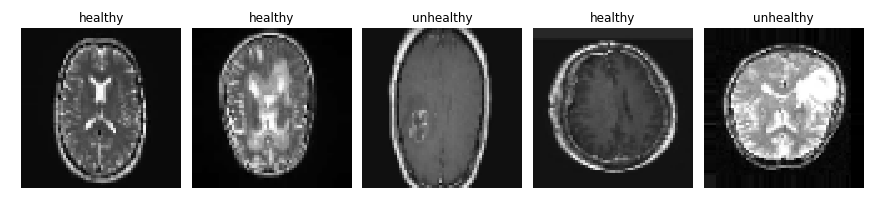


Fig 2. Brain Tumor Dataset Sample

**2.3 Data Preprocessing**

First we identified that all the images in both the datasets were of different size, thus we resized all the images in the dataset to 64x64 to have an uniformity in our dataset. Since the images have pixel values between 0 and 255, the images were normalized between 0 and 1 by dividing each pixel value by 255 where 0 is completely black and 1 is completely white. Doing this increases the speed by which the model gets trained and it makes the training computationally efficient. After which data augmentation techniques were applied on both the datasets to modify the images. First we performed horizontal Flip, this horizontally flips the image along the vertical axis. Then we performed ‘shear\_range’ which shears the image by 20%. After that a random zooming was performed by 20% using ‘zoom\_range’. These data augmentation techniques were performed to provide a robustness to our classification.

**2.3 Image Classification using CNN**

Classification model is built using a deep CNN model consisting of different layers. The structure of the model consists of the following layers.

**Convolutional layer** convolves learned features with input data and uses 2D convolutional layers. This means that this type of network is ideal for processing 2D images. Compared to other image classification algorithms, CNNs actually use very little preprocessing. This means that they can learn the filters that have to be hand-made in other algorithms. So repeated application of the filter to an input allows to detect the features of the input. Therefore, this layer can automatically learn a large number of filters and can help in predicting modelling problems. Parameters used in the convolutional layer are filters , kernel size, activation, input shape. The activation function used is ‘Relu’.

**Maxpool layers** are used after the convolution layers. So it helps in selecting the maximum element from the region of the feature map. Thus it shows the most prominent features of the previous map. The need of maxpool layers is that it reduces the dimensions of the feature map. So the number of parameters to learn and the amount of computations to perform. Parameters used in the maxpool layer are pool size and stride.

**Flatten Layer** is used to convert the data into a 1-dimensional array for inputting it to the next layer. We flatten the output of the convolutional layers to create a single long feature vector. And it is connected to the final classification model, which is called a fully-connected layer.

**Dense Layer** is used after the maxpool layers. Dense layer is the regular deeply connected neural network layer. It is the most common and frequently used layer. Dense layer does the below operation on the input and returns the output. output = activation(dot(input, kernel) + bias). Parameters used in the Dense Layer are units and activation. Activation function used in the output is sigmoid as we want binary classification.

Once we had applied all the layers, we compiled the model using ‘ADAM’ optimizer. We chose ‘ADAM’ optimizer because of its good learning rate and its parameter specific adaptive nature of the learning rates. We used ‘Binary cross entropy loss’ as our loss function because it compares each of the predicted probabilities to actual class output which can be either 0 or 1. Our batch size was taken to be ‘32’ as a very large batch size can adversely affect the quality of the model. Number of epochs was chosen to be ‘25’ to prevent overfitting.

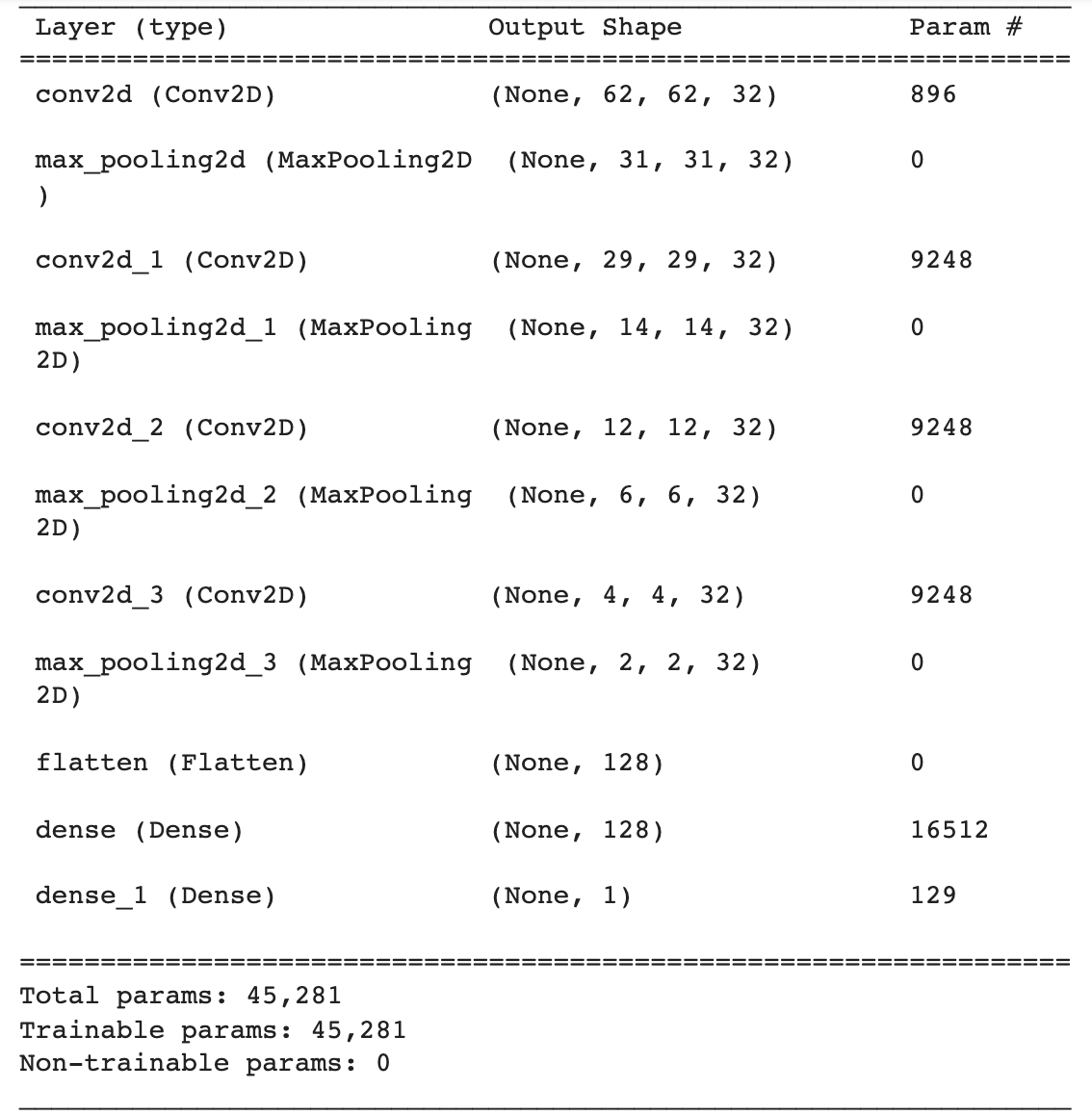


Fig 3. Model Summary

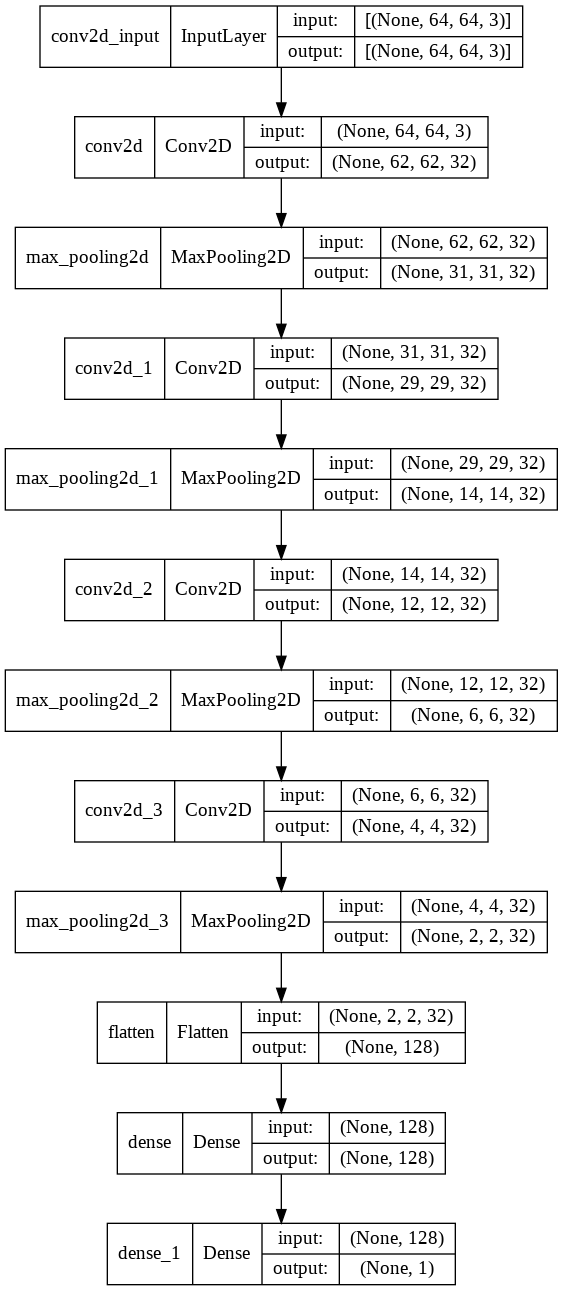


Fig 4. CNN model

**3. Experimental Result and Analysis**

**3.1 Confusion Matrix**

The confusion matrix is a table that consists of 4 different combinations of true and predicted values. It is an important table to evaluate the performance of the machine learning model.

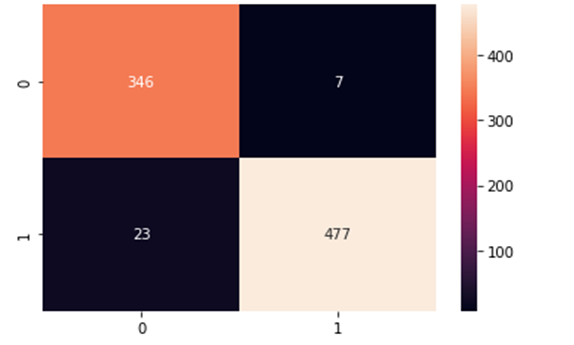
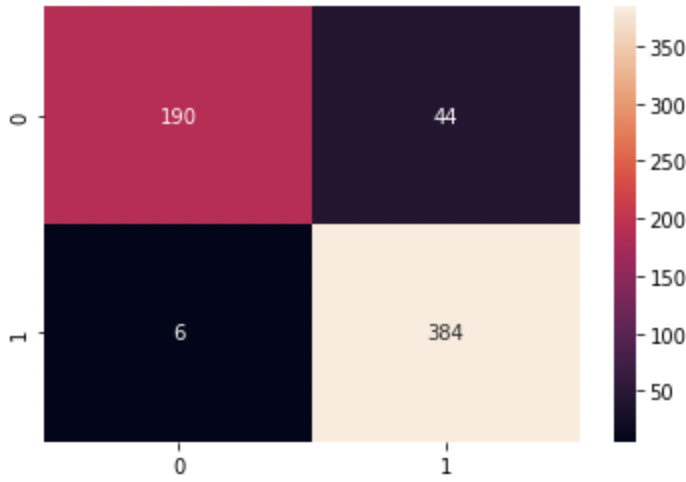
 

Fig 5.(a) Fig 5.(b)

The above is the confusion matrix obtained on the test set for our CNN model. The left is that of the brain tumor dataset while the right one is of chest x-ray dataset.

**3.1.1 True Positive**

This is the measure of outcomes in which the model is able to correctly classify the positive class. The top left grid of the confusion matrix shows the value of True Positives. In the above case the value is 346 and 190, implying that on the test set the model was able to correctly classify 346 images with tumor and 190 images with pneumonia respectively.

**3.1.2 True Negative**

This is the measure of outcomes in which the model is able to correctly classify the negative class. The bottom right grid of the confusion matrix shows the value of True Negatives. In the above case the value is 477 and 384, implying that on the test set the model was able to correctly classify 477 images without tumor and 384 images with pneumonia respectively.

**3.1.3 False Positive**

This is the measure of outcomes in which the model is able to incorrectly classify the positive class. The top right grid of the confusion matrix shows the value of False Positive. In the above case the value is 7 and 44, implying that the model incorrectly predicted7 healthy images as unhealthy and 44 normal cases as pneumonia.

**3.1.4 False Negative**

This is the measure of outcomes in which the model is able to incorrectly classify the negative class. The bottom left grid of the confusion matrix shows the value of False Negative. In the above case the value is 23 and 6, implying that the model incorrectly predicted 23 unhealthy images as healthy and 6 pneumonia cases as normal.

**3.2 Classification Report**

The classification report is an evaluation report of how well the Machine Learning Model performs on the test set. The classification report consists of precision, recall, F1 score and accuracy.

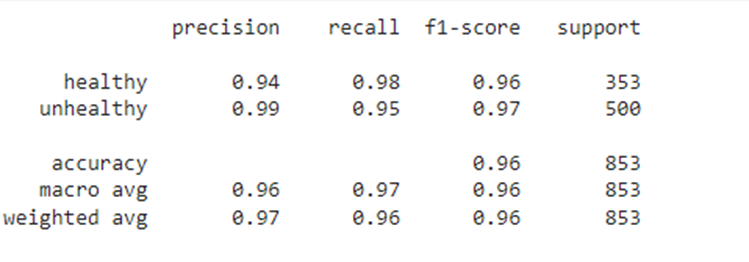
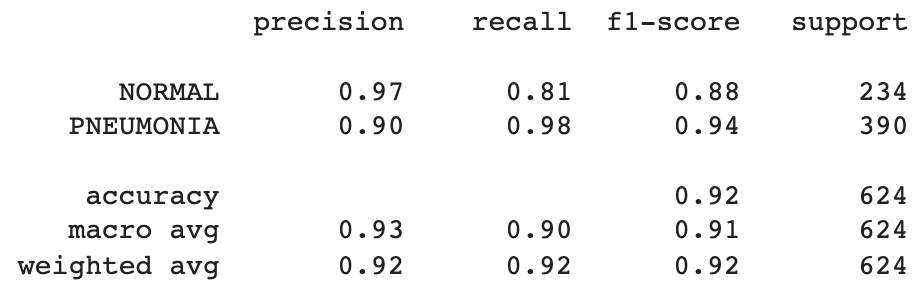
 

Fig 6.(a) Fig 6.(b)

The above classification report on the left is obtained for the CNN model trained on the brain tumor dataset and the right classification report is obtained for the chest x-ray dataset.

**3.2.1 Precision**

Precision quantifies the number of positive class predictions that actually belong to the positive class.

Precision = True Positives / (True Positives + False Positives)

The precision values of our CNN model for the brain tumor dataset for the healthy and unhealthy images is 0.94 and 0.99 respectively and for the chest x-ray dataset for the normal and pneumonia images is 0.97 and 0.90 respectively.

**3.2.2 Recall**

Recall is a metric that quantifies the number of correct positive predictions made out of all positive predictions that could have been made.

Recall = True Positives / (True Positives + False Negatives)

The recall values of our CNN model for brain tumor dataset for the healthy and unhealthy images is 0.98 and 0.95 respectively and for the chest x-ray dataset for the normal and pneumonia images is 0.81 and 0.98 respectively.

**3.2.3 F1 score**

F1 score takes into account both precision and recall values into a single number.

F1 score = (2 \* Precision \* Recall) / (Precision + Recall)

The F1 score values of our CNN model for brain tumor dataset for the healthy and unhealthy images is 0.96 and 0.97 respectively and for the chest x-ray dataset for the normal and pneumonia images is 0.88 and 0.94 respectively.

**3.2.4 Accuracy**

Accuracy is the measure of the number of correctly predicted data points out of all the data points in the test set.

The accuracy of our CNN model for brain tumor dataset, obtained on the test set was 0.96 and for chest x-ray dataset the accuracy is 0.92.

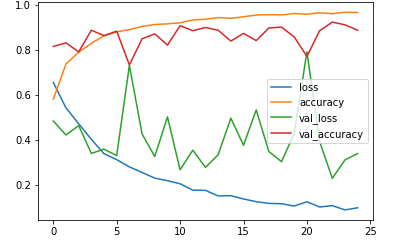
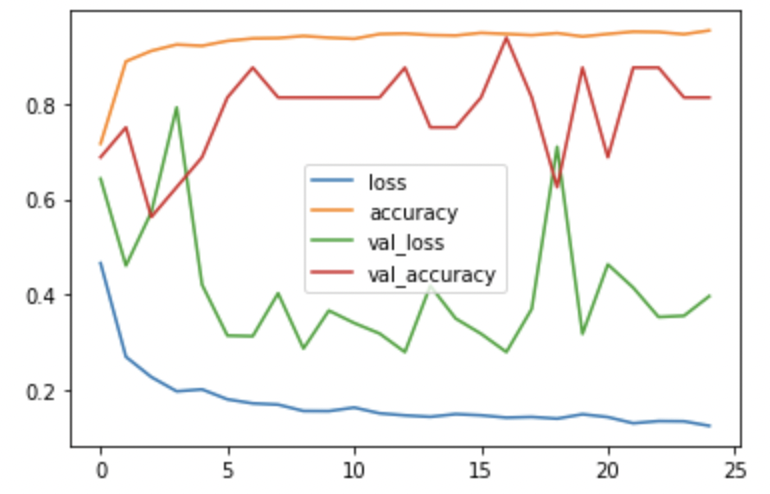
 

Fig 7.(a) Fig 7.(b)

The above diagrams are plots of number of epochs(on the X-axis) vs loss, accuracy, validation loss, validation accuracy(all on the Y-axis). Left image is of the brain tumor dataset while the right image is of the chest x-ray dataset.

It can be inferred from the above plot tha as the number of epochs increase the validation accuracy (red) and the test accuracy increase whereas the validation loss and the test loss decrease .

**3.3 ROC Curve**

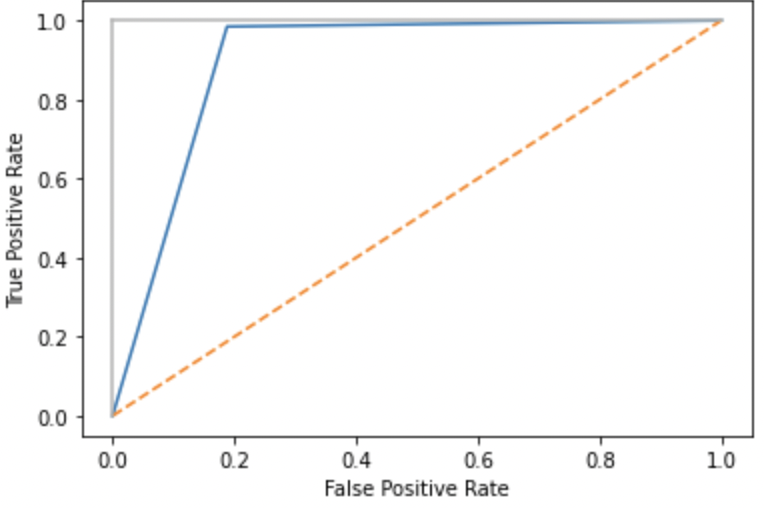
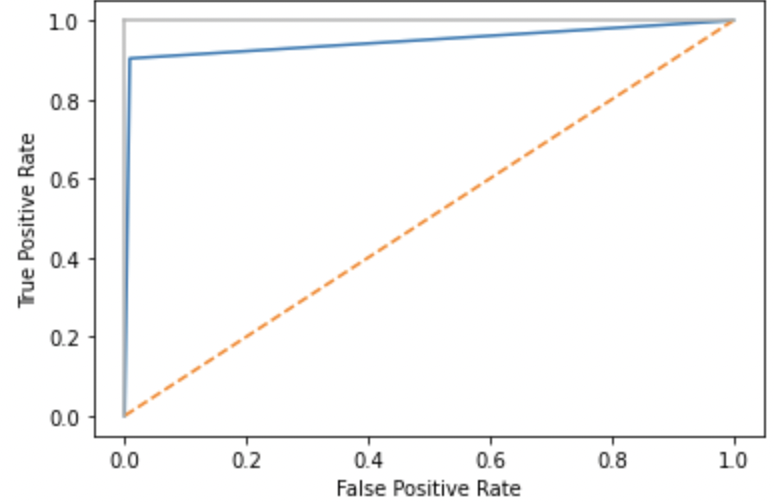
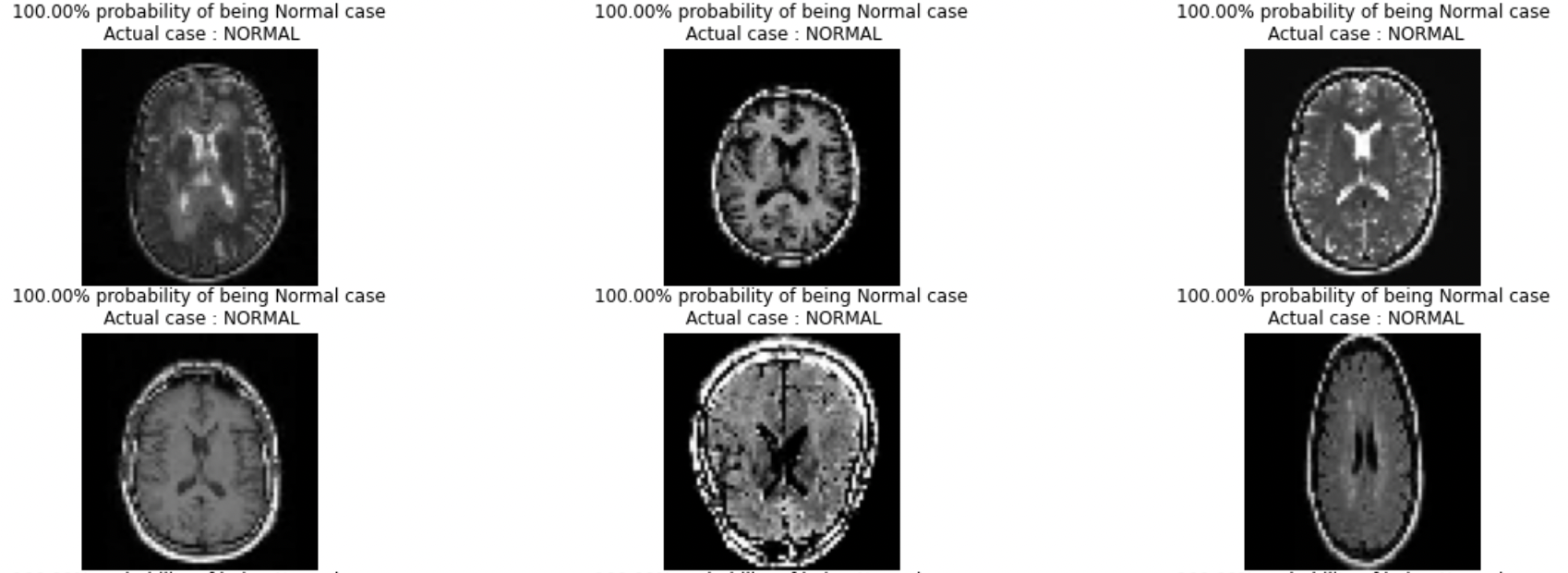
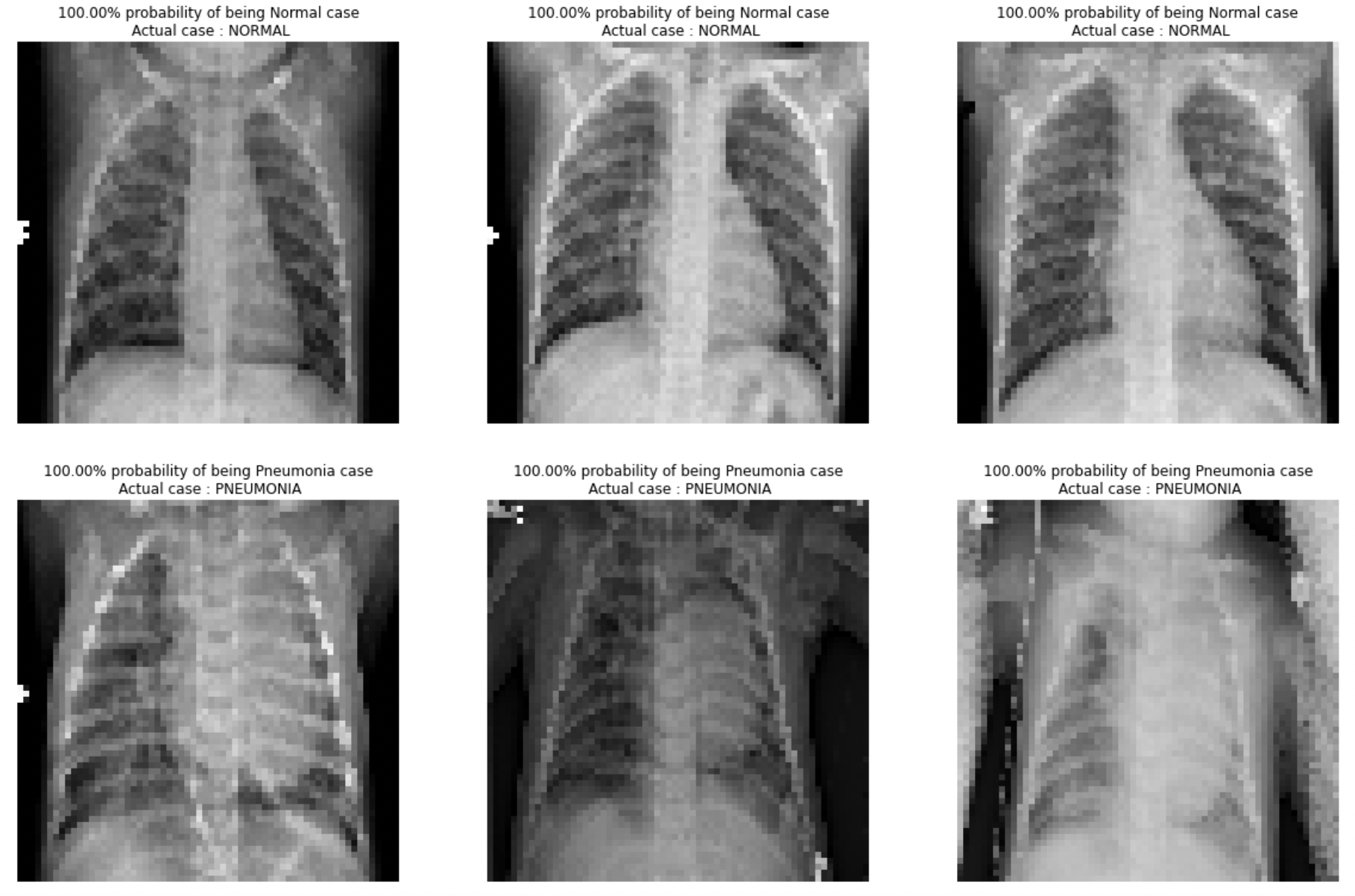


Fig 8.(a) Fig 9.(b)

A ROC curve is constructed by plotting the true positive rate (TPR) against the false positive rate (FPR).The above figures are the ROC curves of the brain tumor dataset and the chest x-ray dataset respectively. The vertical gray line represents the true positive rate and the orange line represents no predictive value. Closer the blue line (actual test) to the gray line, the better the performance of the CNN model.

**3.4 Final Predictions**

****

****

The above images show a part of the actual predictions of our model. We got a desirable result as most of the images were predicted correctly.

**4.Conclusion**

In this project we have obtained results for image classification of brain and chest images. Both datasets were fit using the exact same CNN model. The accuracy obtained by fitting this model on the Brain tumor dataset was 96% and the accuracy obtained for the Chest was 92%. This implies that this model correctly diagnoses tumor 96% and pneumonia 92% of the time.

The precision,recall and F-1 values for the Brain image dataset for the healthy and unhealthy images is 0.94 and 0.99 , 0.98 and 0.95, 0.96 and 0.97 where as precision,recall and F-1 values obtained for the Chest image dataset the normal and pneumonia images is 0.97 and 0.90, 0.81 and 0.98, 0.88 and 0.94 respectively.

**5. References**

1. Brain Tumor Classification Using Convolutional Neural Network by Sunanda Das, O.F.M. Riaz Rahman Aranya, Nishat Nayla Labiba Department of Computer Science and Engineering Khulna University of Engineering & Technology Khulna-9203, Bangladesh
2. Brain tumor classification using deep CNN features via transfer learning by S. Deepak, P.M. Ameer Department of Electronics & Communication Engineering, National Institute of Technology, Calicut, India