PNEUMONIA DETECTION CHALLENGE

Team Members

1. **Maneesh Kumar**
2. **Kannan**
3. **Sreejith**
4. **Ananthu**
5. **Anish**
6. **Geeta**
7. **Izas Ahmed**

AIML CAPSTONE PROJECT ON PNEUMONIA DETECTION

Table of Contents

Abstract

Pneumonia is one of the common respiratory problems, considered as inflammation in lungs due to viral and other bacterial infection. Pneumonia patients suffer from serious breathing problems and the infection in lungs can also lead to other health problems such as fever and cough. The treatments of this disease depend on the severity. Doctors diagnose this disease based on the symptoms and observing the breathing pattern. The bacterial pneumonia can cause severe unrest compared to viral one. Hence a chest X-Ray followed by optional blood tests are recommended for proper diagnosis and subsequent medication. In this aspect Machine learning and deep learning models play an important role as it would be difficult for the doctors to read and infer the outcome from the X-ray images. In addition, the factors of external influence such as Lung cancer, fluid overload (pulmonary edema), bleeding, volume loss (atelectasis or collapse), or post-radiation or surgical changes appear as increased opacity on CXR make it more difficult for optimal diagnosis accurately. The aspect of sensitivity of CXR with respect to the positioning of the patient and depth of inspiration cannot be ignored. Most importantly, the doctors and medical community are faced with reading high volumes of images every shift on a regular basis and hence a smart system such as machine learning based diagnostics help to improve the efficiency and reach of diagnostic services. Sparse material tissues like lungs appear black in X-ray images as they don't absorb X-rays. Dense tissues like bones absorb X-rays and appear white. Pneumonia manifests as an area of opacity in X-ray.

In the present work the dataset contains X-ray images of patients' lungs. Here we develop models for both segmentation techniques to detect lung opacities on chest X-ray film images, i.e., i.e., to locate the position of inflammation in the DICOM image. The present work aims at developing CNN architectures to detect lung opacities on chest X-ray film images and subsequently building classifiers pertaining to this disease. The subsequent part of the work aims at applying transfer learning, fine tuning the parameters and improving the model accuracy architectures for building classifiers pertaining to this disease.

1. Introduction:

Respiratory infections are found to be prominent cases for hospitalization, for example, in Iraq they represent 60% mean consultations corresponding to 45% of the average population hospitalized [1]. On the other hand, 22-42% of adult pneumonia patients require hospitalization and 5-10% of them require ICU [2].

Considering Among adults suffering from pneumonia, it is estimated that between 22 and 42% require hospitalization and between 5 and 10% need an intensive care unit, and the lethality varies between 5 and 50% depending on the severity of the condition, which is higher in the elderly and immunosuppressed patients [2]. According to UNICEF data [3], pneumonia is increasingly becoming a life threatening disease among children compared to other diseases among them claiming the lives of over 700,000 children under five every year, or around 2,000 every day. South Asia is prominent as 2,500 cases per 100,000 children are noticed by West and Central Africa (1,620 cases per 100,000 children).

Why do we need to solve it?

Pneumonia accounts for over 15% of all deaths of children under 5 years old internationally. In 2015, 920,000 children under the age of 5 died from the disease. Pneumonia accounts for over 500,000 visits to emergency departments in the United States. There were 50,000 deaths in 2015 keeping the ailment on the list of top 10 causes of death in the United States.

Why can't doctors themselves handle the issue?

Well, reading the X-ray/CXR is a complicated thing because of following reasons:

1. Lung cancer, fluid overload (pulmonary edema), bleeding, volume loss (atelectasis or collapse), or post-radiation or surgical changes appears as increased opacity on CXR.
2. Outside of the lungs, fluid in the pleural space (pleural effusion) also appears as increased opacity on CXR.
3. Positioning of the patient and depth of inspiration can alter the appearance of the CXR
4. Clinicians are faced with reading high volumes of images every shift.

To improve the efficiency and reach of diagnostic services, an automated Pneumonia detection system is very much necessary.

In the present work, our objective is to build a pneumonia detection system i.e., to locate the position of inflammation in the DICOM image. In other words, to build an algorithm that needs to automatically locate lung opacities on chest radiographs. We have developed CNN model to detect lung opacities on chest X-ray film images. The subsequent part of the work aims at applying transfer model learning, fine tuning the parameters and improving the model accuracy architectures for building classifiers pertaining to this disease. The dataset contains X-ray images of patients' lungs.

The data description for the present problem is provided below:

- In the dataset, some of the features are labeled "Not Normal No Lung Opacity". This extra third class indicates that while pneumonia was determined not to be present, there was nonetheless some type of abnormality on the image and oftentimes this finding may mimic the appearance of true pneumonia. Dicom original images: - Medical images are stored in a special format called DICOM files (*.dcm). They contain a combination of header metadata as well as underlying raw image arrays for pixel data.
- Dataset has been attached along with this project. Please use the same for this capstone project.
- Original link to the dataset: https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data

The methodology used and the associated tasks performed are described in the figure 1 below:

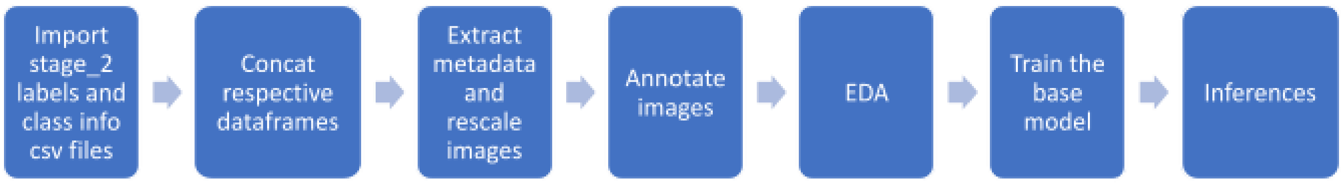


Figure 1: Pneumonia segmentation and prediction using deep learning methods

We took csv files and imported them first. We performed EDA and verified for null values. However, there are records with multiple pneumonia presence, so we concatenated files to create a dataframe. Metadata from the images is taken and the information so extracted is used as columns in the dataframe.

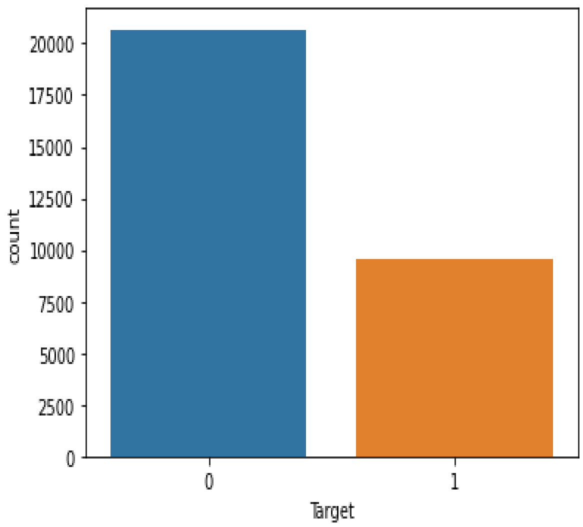
The csv data comprises patient IDs followed by associated bounding box attributes such as x, y, height, width and associated target label (pneumonia :1 its absence:0). The shape of the data frame is `(30227, 6)`

2. Exploratory Data Analysis:

The csv file 'stage_2_train_labels.csv' contains the patient id , the bounding box coordinates (x, y, width and height) along with the Target column consisting of values 0 and 1. There are 30,227 records in this file. The other file is 'stage_2_detailed_class_info.csv' which contains the patient id and the classes: **No Lung Opacity/Not Normal, Normal, Lung Opacity.** The following are the pre- processing / EDA analysis that has been performed on the data :

Out of (30227, 6), the total number of null values in bounding boxes columns is equal to the total number of 0s in the Target column. The class distribution of presence of pneumonia (class label 1);

its absence (class label 0) are: {0: 20672, 1: 9555}. These studies indicate that all the records with class label 0 do not contain bounding boxes as expected because the bounding boxes are indicative of opacity in the images.



Secondly, one can notice that the instances with class 0 are more than double that of class 1 instances which can be considered as class imbalance problem and to be handled as part of classification task during subsequent phases of this work. One can notice the relative numbers from the adjacent count plot

Figure 2: Distribution of the patients across classes 0, 1

From the above data, we can notice that there are around `30227` rows in `stage_2_train_labels.csv` but we have a total of `26684` unique patientIDs. Hence `3543` patients have more than one bounding box.

For example, the patientId: `c1f7889a-9ea9-4acb-b64c-b737c929599a` is the same from 4th and 5th rows of the dataframe:

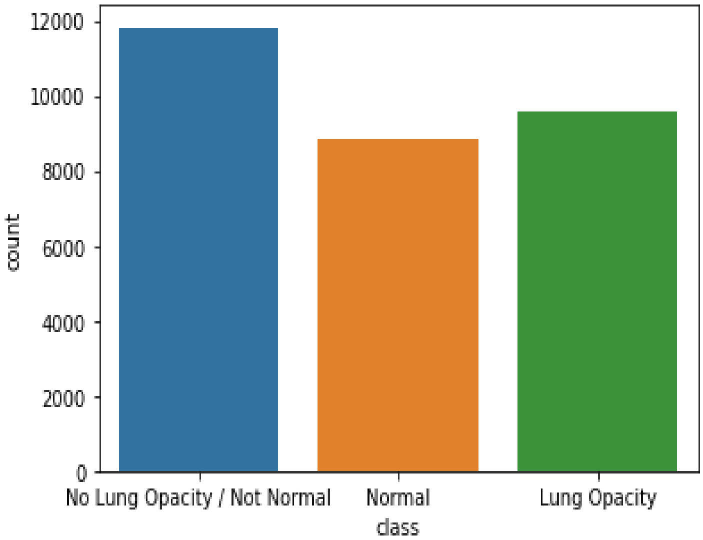
```
selected_data = trainlabels[trainlabels.patientId == 'c1f7889a-9ea9-4acb-b64c-b737c929599a']
```

	patientId	x	y	width	height	Target
30225	c1f7889a-9ea9-4acb-b64c-b737c929599a	570.0	393.0	261.0	345.0	1
30226	c1f7889a-9ea9-4acb-b64c-b737c929599a	233.0	424.0	201.0	356.0	1

On the other hand, it is observed that the total number of patients match with the value from training labels is `26684` and the associated classes are: `Normal`, `Lung Opacity` and `No Lung Opacity/Not Normal`

The relative population of these classes can be noticed from the count plot below:

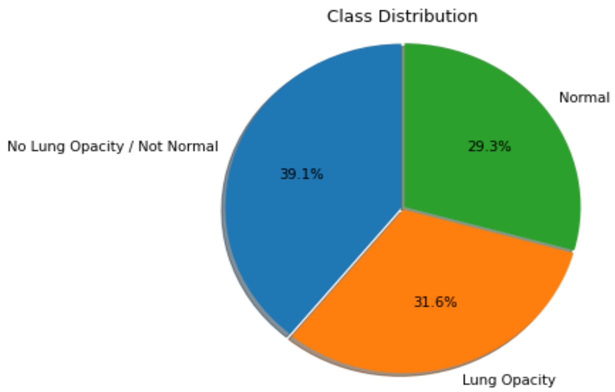
From the countplot, it is very much noticeable that all the classes are almost equally distributed. There are no missing values in the detailed_class_info data



It is observed that total number of null values for "Normal" and "No Lung Opacity / Not Normal" is equal to total number of 0s in Target column i.e., who do not have pnemonia have their bounding box columns as NaN.

Figure 3: Distribution of the patients based on the opacity information from the images.

<u>% of population across opacity groups:</u>



Hence the first two categories are labelled as 0 while the class with Lung Opacity as 1. The resulting distribution can be seen as given in the figure below:

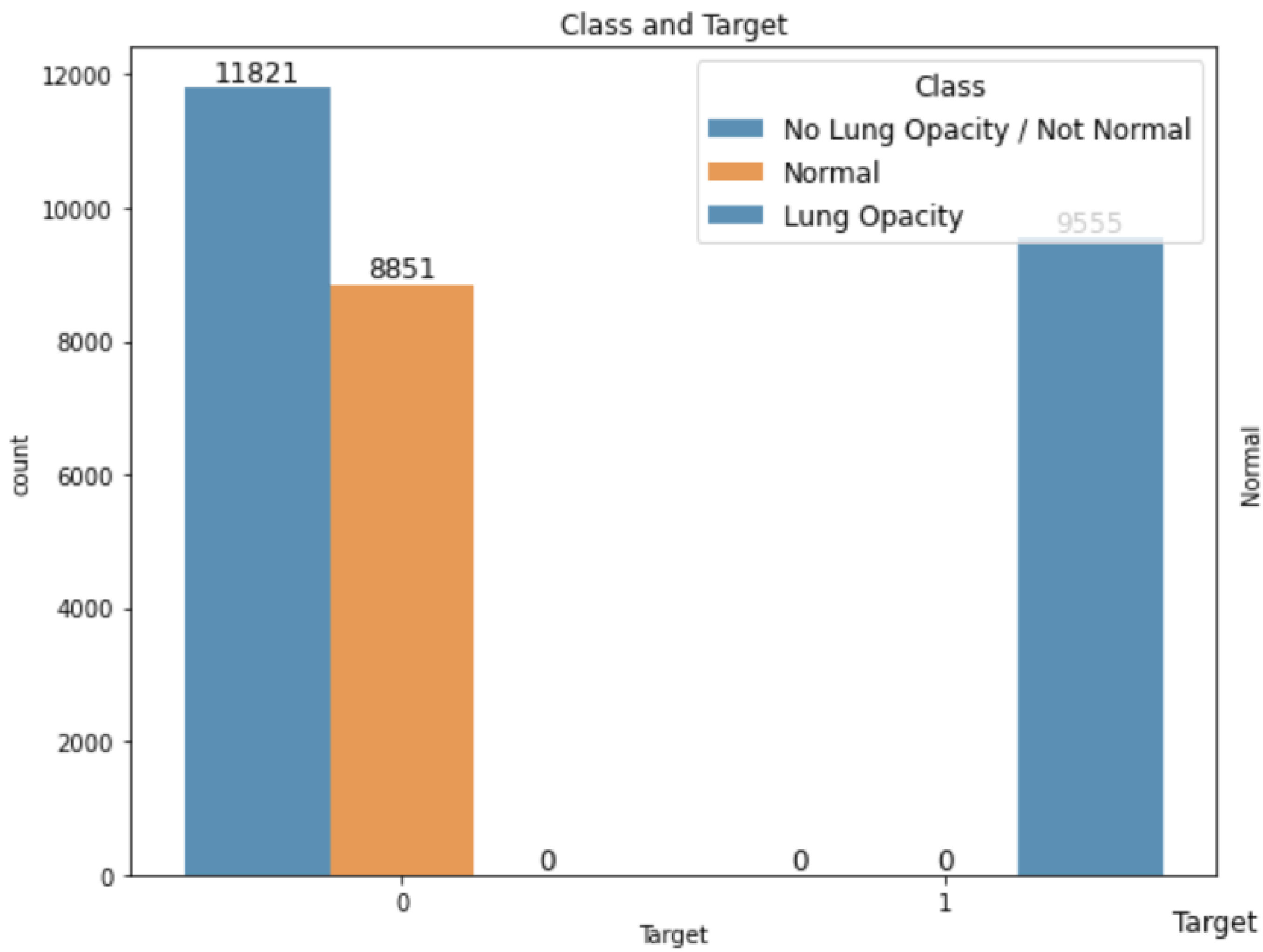


Figure 4: Distribution of the patients from the class labels data

Observations regarding the images and their size:

- Images are stored in DICOM(Digital Imaging and Communications in Medicine) format with .dcm extension. This DICOM image contains more information like patient age, sex, modality, view position and body part and so on.
- `pydicom` library is used to read the images
- Each image size is 1024 x 1024 which requires scaling before employing machine learning models.
- "BodyPartExamined" is CHEST which is expected
- "Modality" is CR (Computer Radiography)

The features '`Age`', '`Sex`', '`BodyPartExamined`' are dropped from the data-frame for further processing as they are not significant for the analysis under consideration. In the next step the metadata information such as Age, gender is mapped for each of the corresponding images and a couple of examples are furnished below considering two random patient ids. Here the function: `MapImagesToAnnotations(patientID)` takes patientID as argument and displays corresponding image and associated patient information:

```
MapImagesToAnnotations('c1edf42b-5958-47ff-a1e7-4f23d99583ba')
```

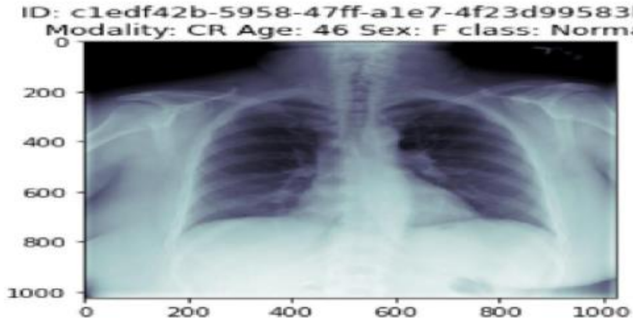
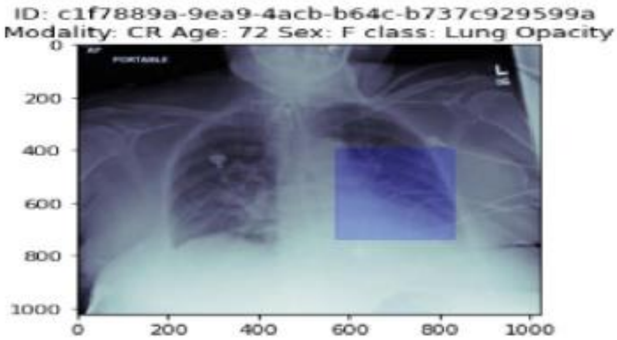
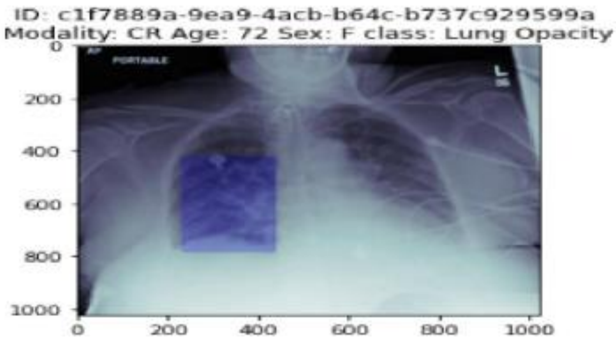


Figure 5: Annotated images with Non -opacity and opacity regions (blue) observed.

```
MapImagesToAnnotations('c1f7889a-9ea9-4acb-b64c-b737c929599a')
```





We have performed detailed analysis on the gender v/s classes. It can also be seen that the number of male patients are more than that of female across the label categories as depicted from the count plot below:

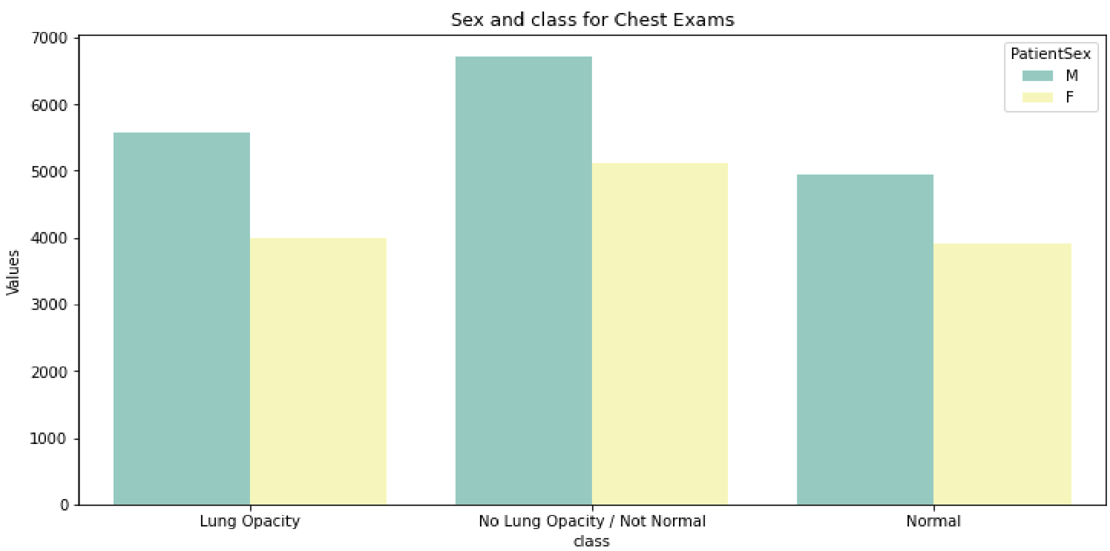


Figure 6: Distribution of the patients from the class labels data

Outlier analysis is performed with respect to the patient age attribute and the opacity. The associated box plots are furnished in the Figure 7 below:

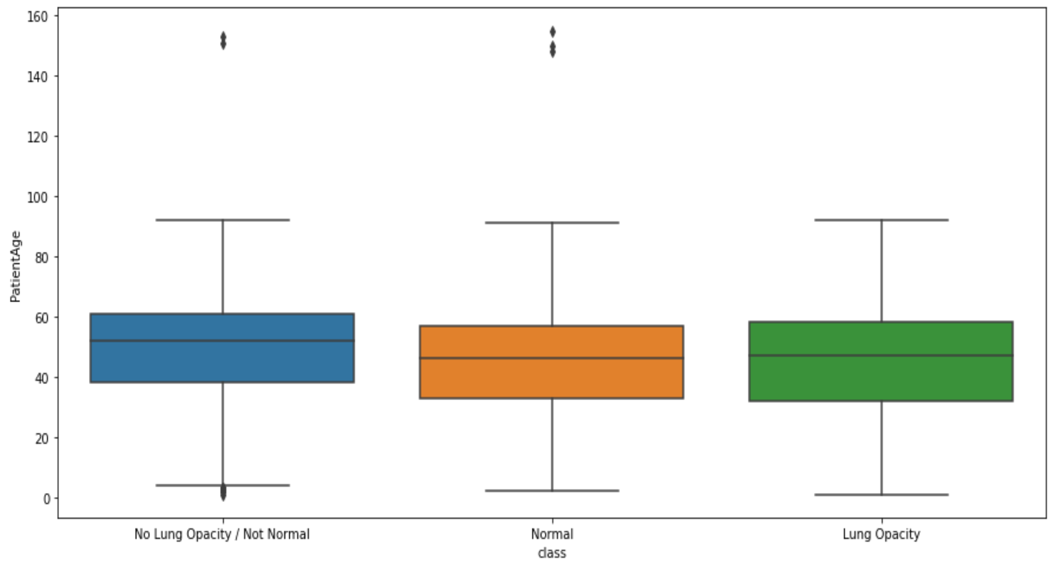
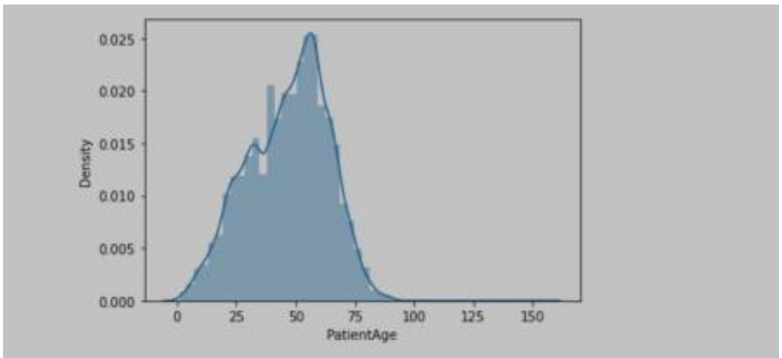
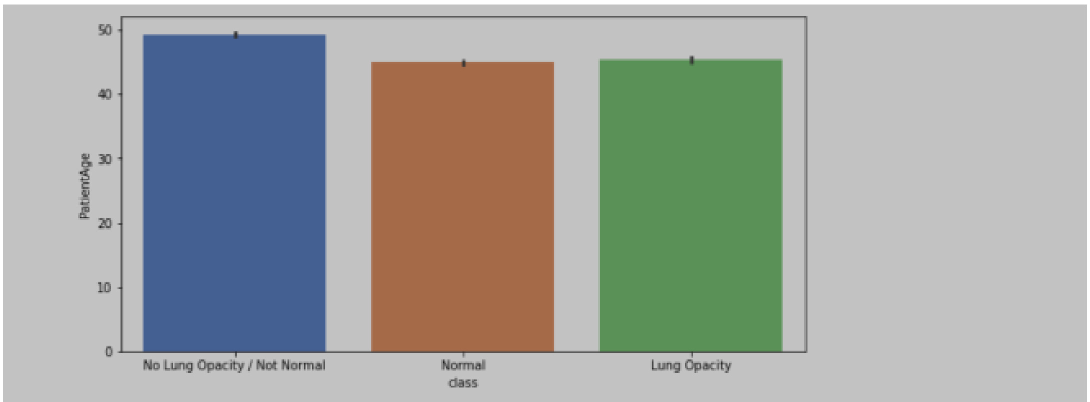


Figure 7: Box plots for opacity distributions



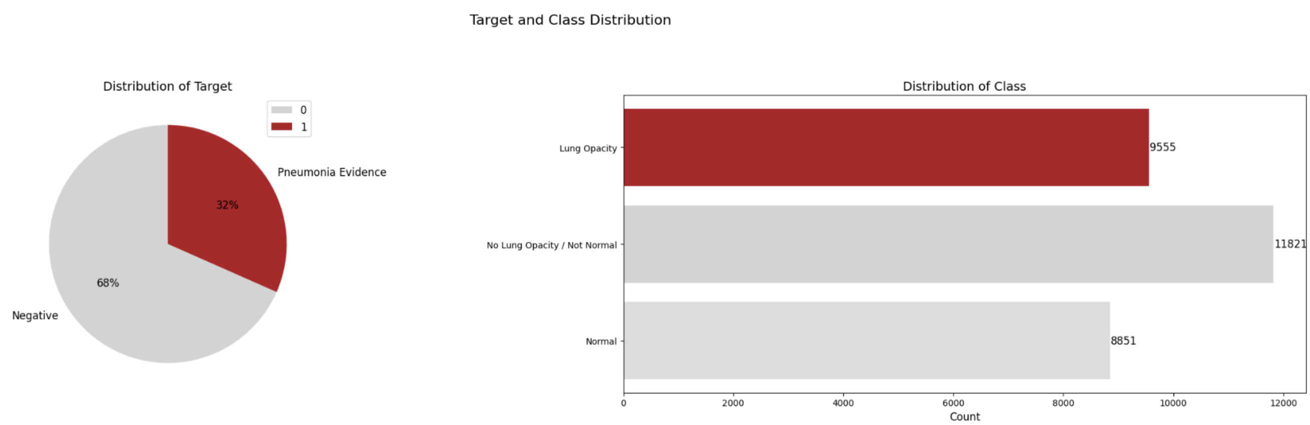
Observation: Looks like normal distubution of age



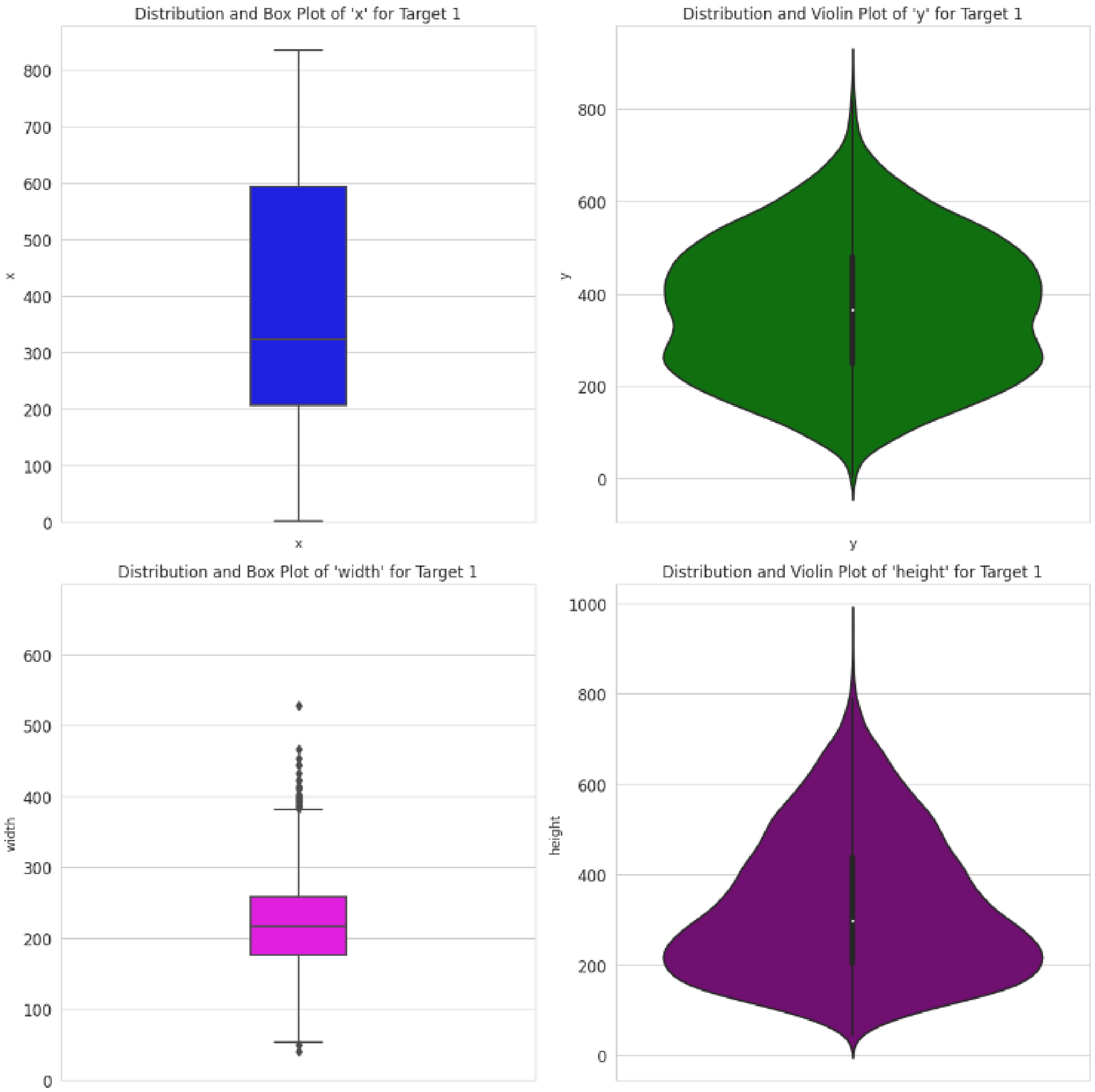
Observation: This is the distubution of Age with class, maximum age of person with pneuomina is arund 45

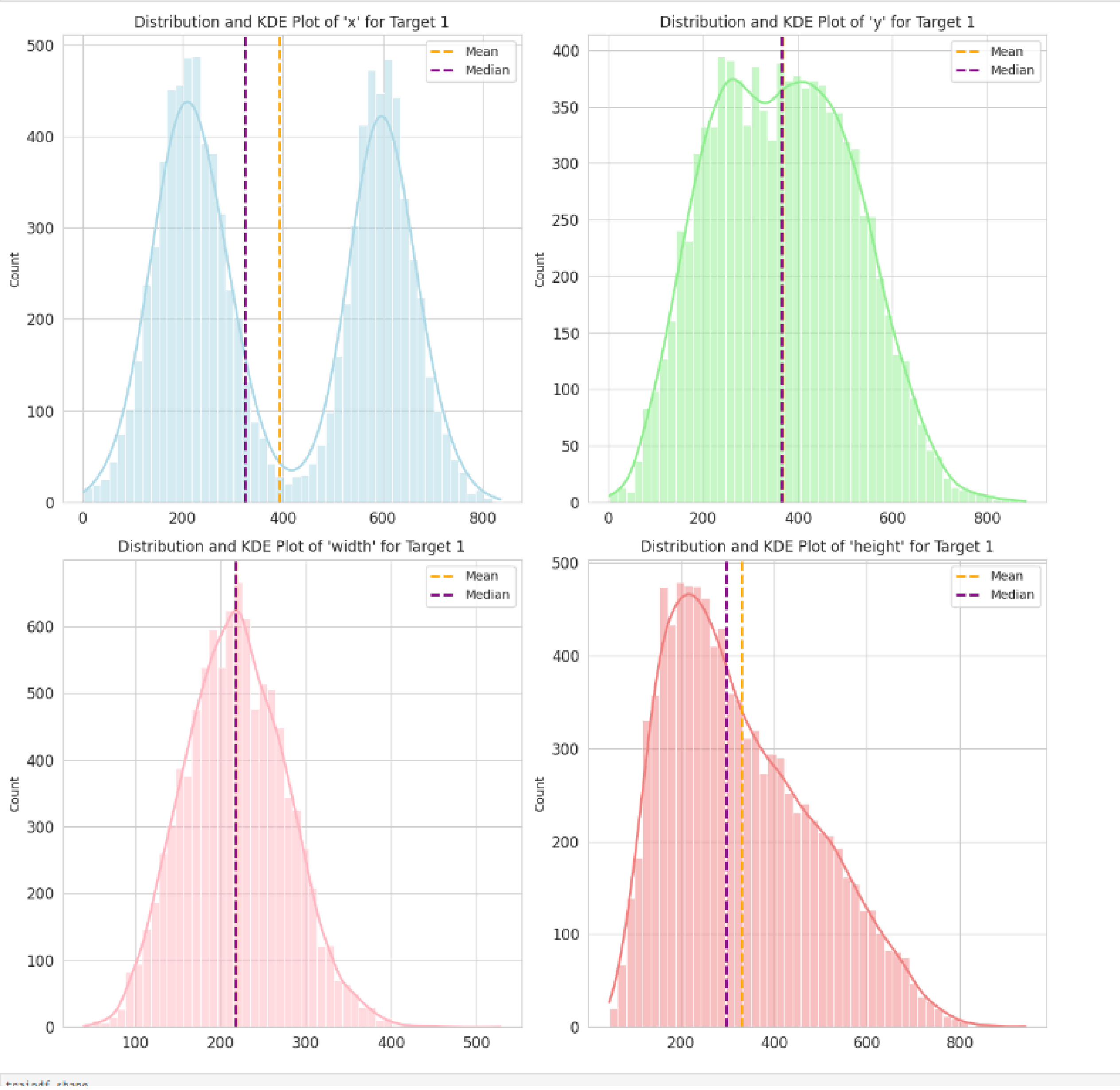
Number of patients in age category

<=75	13318
<=50	12157
<=26	3972
<=100	780



Below is the Distribution and Box Plot along with KDE for x and y for target 1


Distribution and Box Plot of 'x' for Target 1
x
Distribution and Violin Plot of 'y' for Target 1
y
Distribution and Box Plot of 'width' for Target 1
width
Distribution and Violin Plot of 'height' for Target 1
height

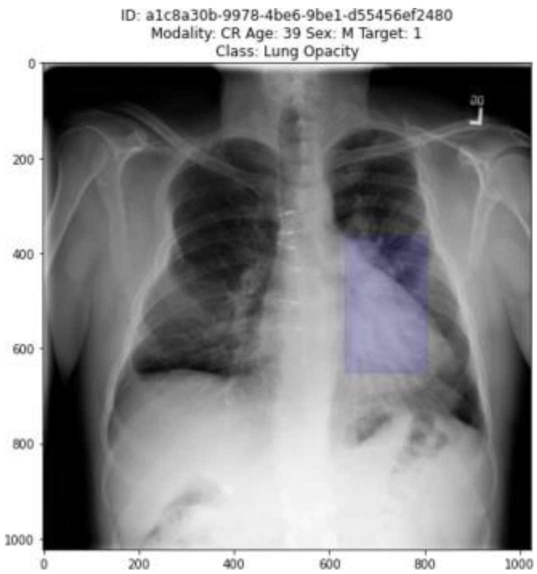


It is also observed that:

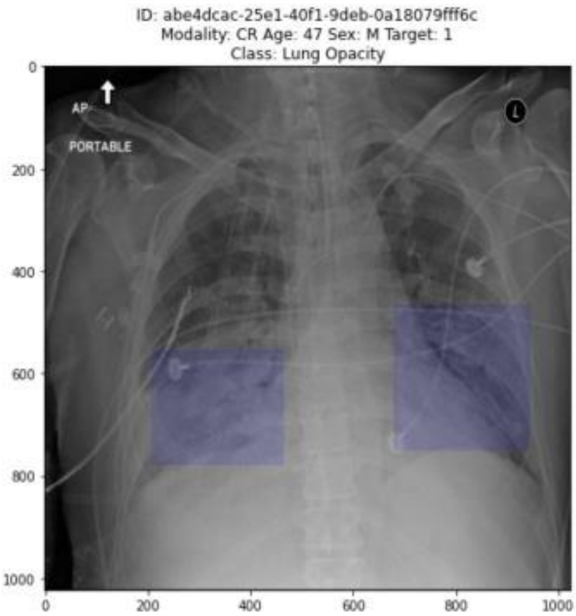
- The mean age is 46 years , whereas the minimum age is 1 year and the max age is 155 which seems to be an outlier.
- 50% of the patients are of around 49 age , the std deviation is 16 which suggest that age is not normally distributed.
- There are 8851 normal cases, people with lung opacity are 9555 and No Lung Opacity / Not Normal are 11821.
- Patients with evidence of Pneumonia are associated with Lung Opacity class and target = 1.
- Patients with no definitive evidence of Pneumonia are either of Normal or No Lung Opacity / Not Normal class and target = 0.

Observations regarding the Bounding boxes

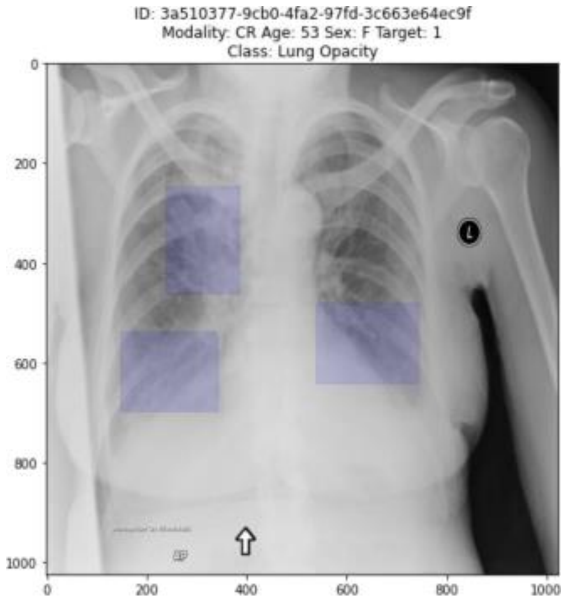
X-ray images with ONE bounding box



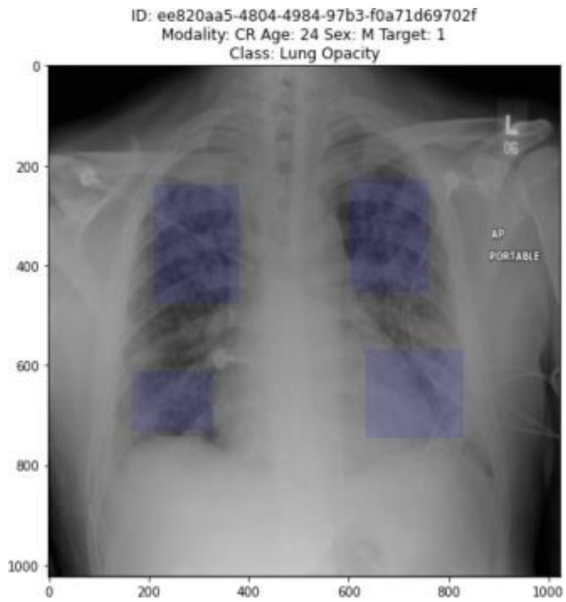
X-ray images with TWO bounding box



X-ray images with THREE bounding box



X-ray images with FOUR bounding box



Number of patients per bounding boxes in the dataset

number_of_bounding_boxes	number_of_patients_per_bounding_boxes
1	2614
2	3266
3	119
4	13

Observations: Maximum patients have 2 bounding boxes while 13 patients have 4 bounding boxes.

More information from image metadata from the exploratory data analysis:

Metadata in the DCIM image:

A single DCIM image was taken from the dataset and the metadata has been displayed. It includes the patient details such as name, age, modality, gender, view position; details about the image itself, date and time among other parameters. Additional parameters (age, gender, view position, pixel spacing) have been appended to the amalgamated data frame for analysis.

Data frame attribute analysis:

Gender: The dataset consists of a higher percentage of male patients compared to females.

Patient age*:* The histogram plot indicates that there is a greater representation of patients in the 40- 60 age range.

ViewPosition: It indicates if the x-ray is taken from posterior or anterior position.

Duplicate records (records with more than one X-ray):

There are 3543 duplicate entries in our dataset.

Resizing the images:

All of the images are of the size 1024 * 1024. this size might slow down the model building process hence we resize the image to a new size of 128*128.

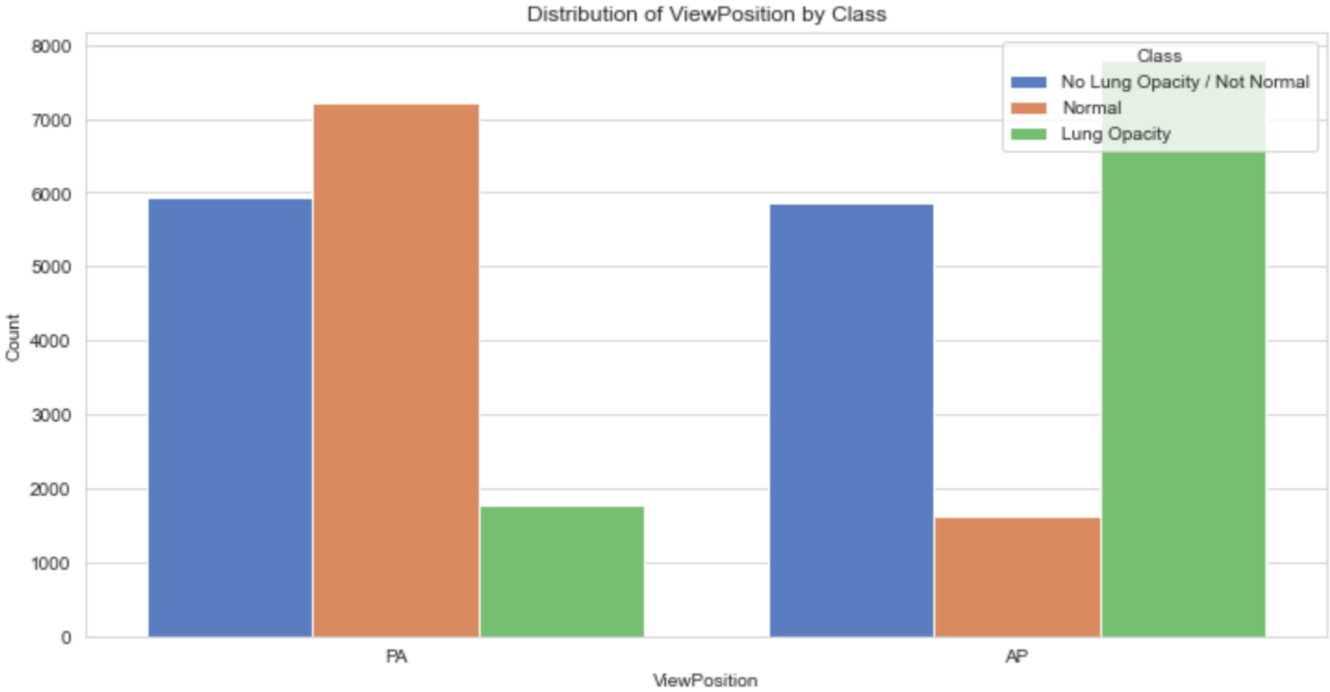
Based on these features following visualizations are generated:

It is observed that

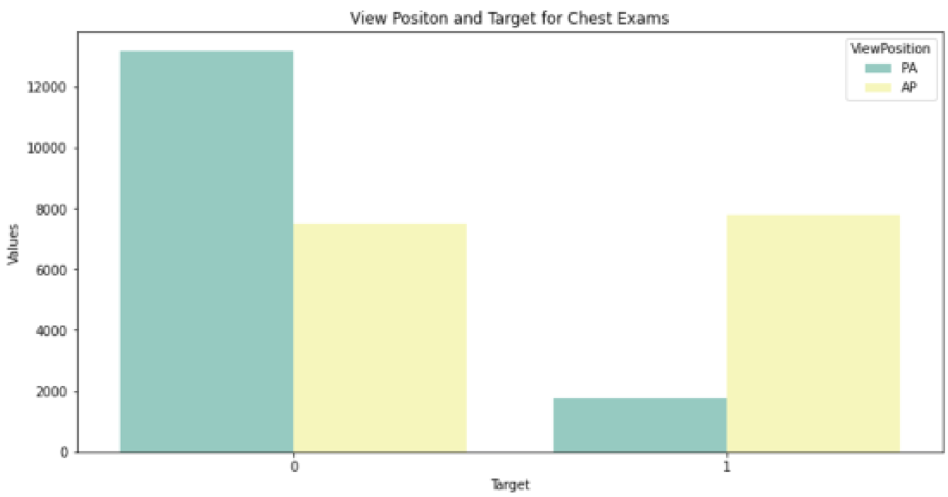
- Numbers of patients are highest for age group 51-75 for overall case
- But the number of patients is highest for the age group 27-50 for Target=1 case.

- **Posterior/Anterior (PA)**: X-ray is taken from the back part of the chest. So it hits the posterior part before the anterior part. Patient needs to stand against the X-ray machine.
- **Anterior/Posterior (AP)**: X-ray is taken from the front part of the chest. So it hits the anterior part before the posterior part. This is taken when Patient cannot stand against an X-ray machine but heart size is exaggerated.

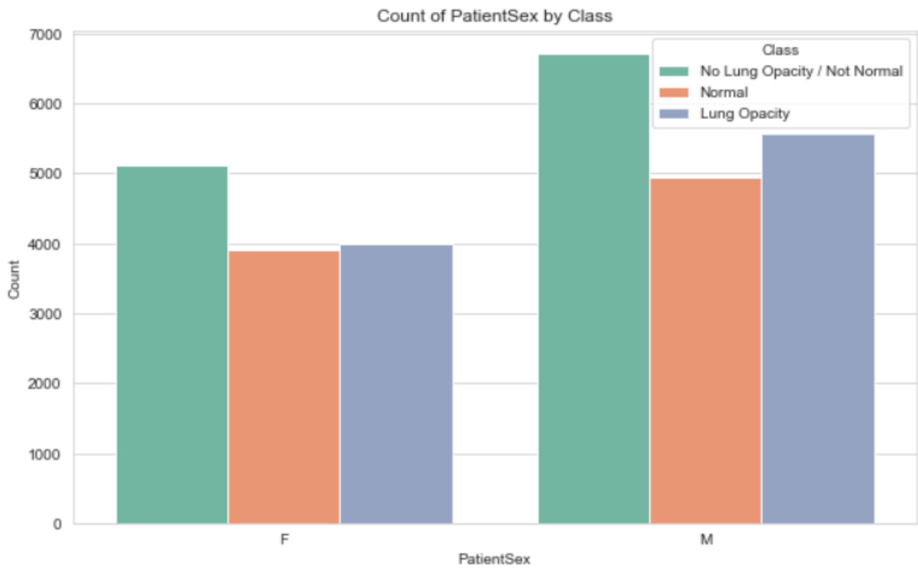
View positions and various states of opacity



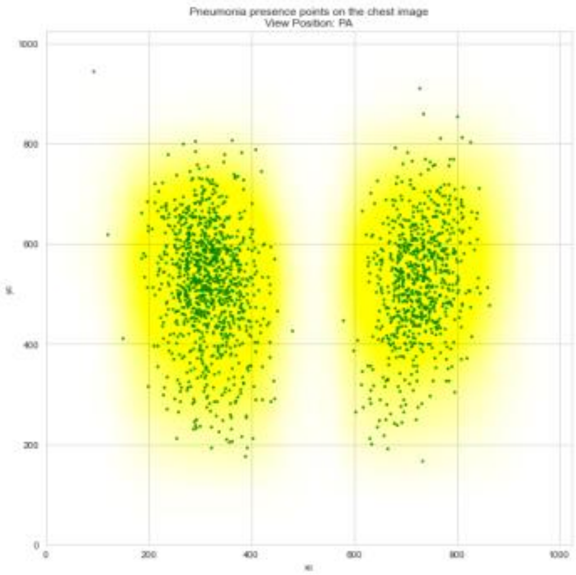
View positions and Target

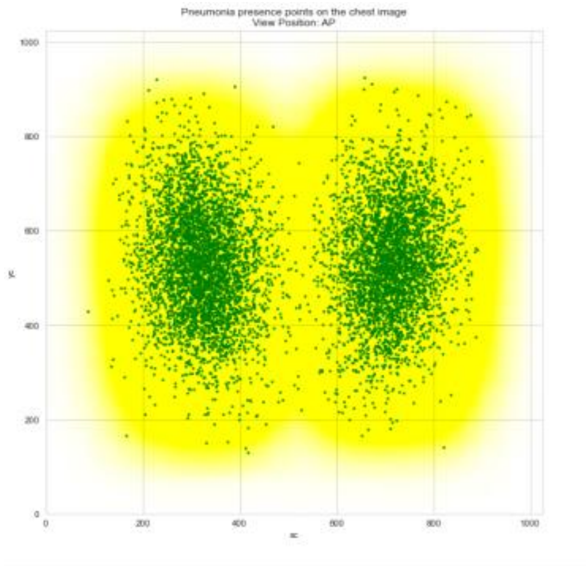


View Count of PatientSex by Class



<u>Pneumonia present points on chest X-Ray</u>





Scatter plots show the concentration of pneumonia affected regions in lungs. It is seen that pneumonia affects the central part of the lungs more than at the boundaries. Since there are more patients with view position = AP, we can see more scatter points in 2nd diagram

3. Model Building:

We are creating a classification model to classify using images if a patient has pneumonia or not.

We are using a sample of 14,000 images with 7,000 data points having target ‘0’ and 7,000 data points with target ’1’.

<u>CNN model:</u>

Model 1:

After preprocessing the data we resize the images in 128*128 format . Add 3 channels to images.

We have 3 convolutional layers with “relu” activation function.

We have 1 DNN layer and 1 output layer with “softmax” activation function and 0.5 dropout.

Optimizer= Adam with a learning rate of 0.001.

Metrics=accuracy

Epochs=10

Loss function= sparcecategory

crossentropy.

Results of CNN model

Performance:

- Training Accuracy: 66.09%
- Test Accuracy: 66.09%
- Test Loss: 0.9035

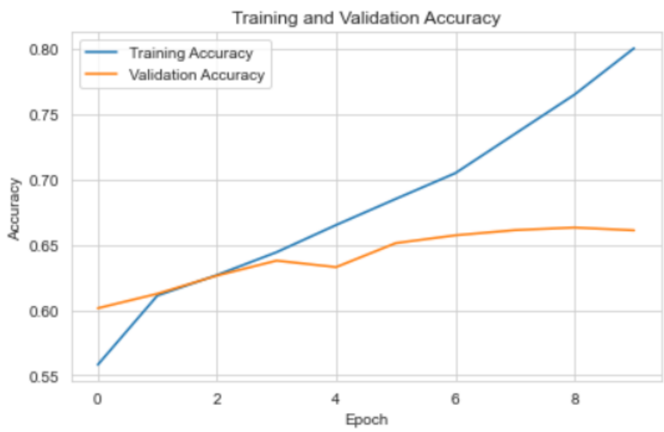
Classification Report:

- F1-scores range from 0.56 to 0.72.

REPORT:

Classification Report:

	precision	recall	f1-score	support
No Lung Opacity / Not Normal	0.66	0.77	0.71	1911
Lung Opacity	0.62	0.52	0.56	2385
Normal	0.71	0.73	0.72	1750
accuracy			0.66	6046
macro avg	0.66	0.67	0.67	6046
weighted avg	0.66	0.66	0.66	6046



Model 2:

Performance:

- Training Accuracy: 61.11%
- Test Accuracy: 57.62%
- Test Loss: 1.0112

Classification Report:

- F1-scores range from 0.47 to 0.69.

Classification Report:

	precision	recall	f1-score	support
No Lung Opacity / Not Normal	0.75	0.43	0.54	1911
Lung Opacity	0.54	0.42	0.47	2385
Normal	0.54	0.95	0.69	1750
accuracy			0.58	6046
macro avg	0.61	0.60	0.57	6046
weighted avg	0.60	0.58	0.56	6046

Model 3:

Metrics=accuracy

Epochs=10

Loss function= binary cross entropy.

Results of CNN model

- Training Accuracy: 70.49%
- Test Accuracy: 70.49%
- Test Loss: 0.9471

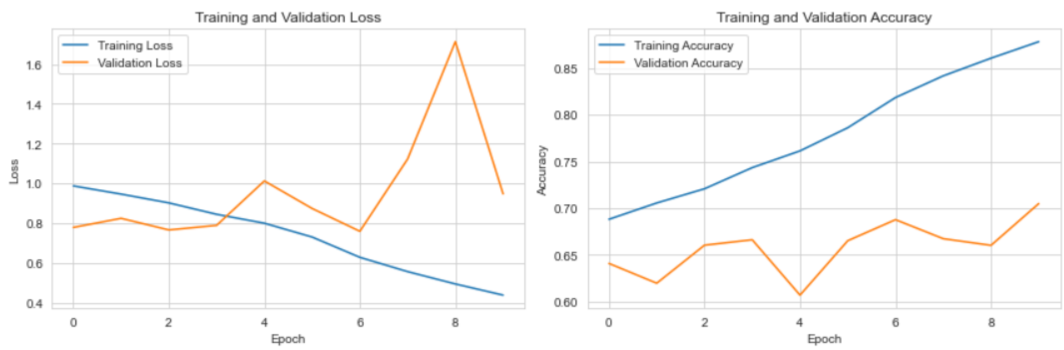
CLASSIFICATION

REPORT:

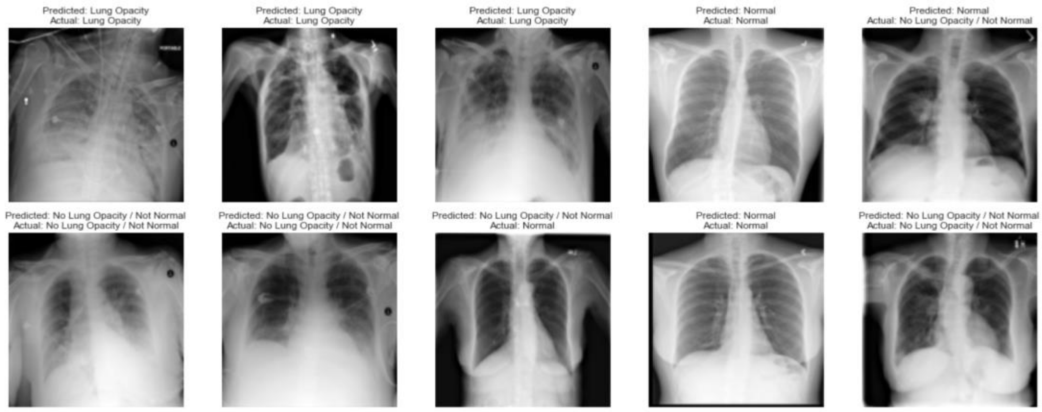
Classification Report:

	precision	recall	f1-score	support
No Lung Opacity / Not Normal	0.71	0.78	0.74	1911
Lung Opacity	0.68	0.57	0.62	2385
Normal	0.73	0.81	0.76	1750
accuracy			0.70	6046
macro avg	0.70	0.72	0.71	6046
weighted avg	0.70	0.70	0.70	6046

As precision and recall are important measures in the medical field, increasing that means our model is performing better now.



Prediction



Model 4:

Metrics=accuracy

Epochs=10

Loss function= binary cross entropy.

Results of CNN model

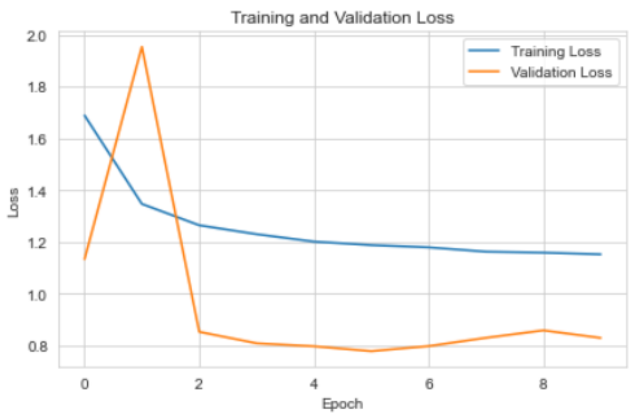
- Training Accuracy: 70.49%
- Test Accuracy: 63.52%
- Test Loss: 0.8285

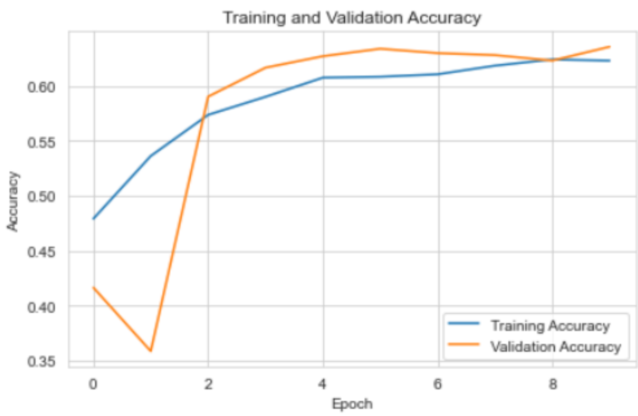
CLASSIFICATION

REPORT:

Classification Report:

	precision	recall	f1-score	support
No Lung Opacity / Not Normal	0.71	0.60	0.65	1911
Lung Opacity	0.60	0.46	0.52	2385
Normal	0.61	0.91	0.73	1750
accuracy			0.64	6046
macro avg	0.64	0.66	0.64	6046
weighted avg	0.64	0.64	0.62	6046





Model 5:

Metrics=accuracy

Epochs=10

Loss function= sparse cross entropy.

Results of Resnet50 model

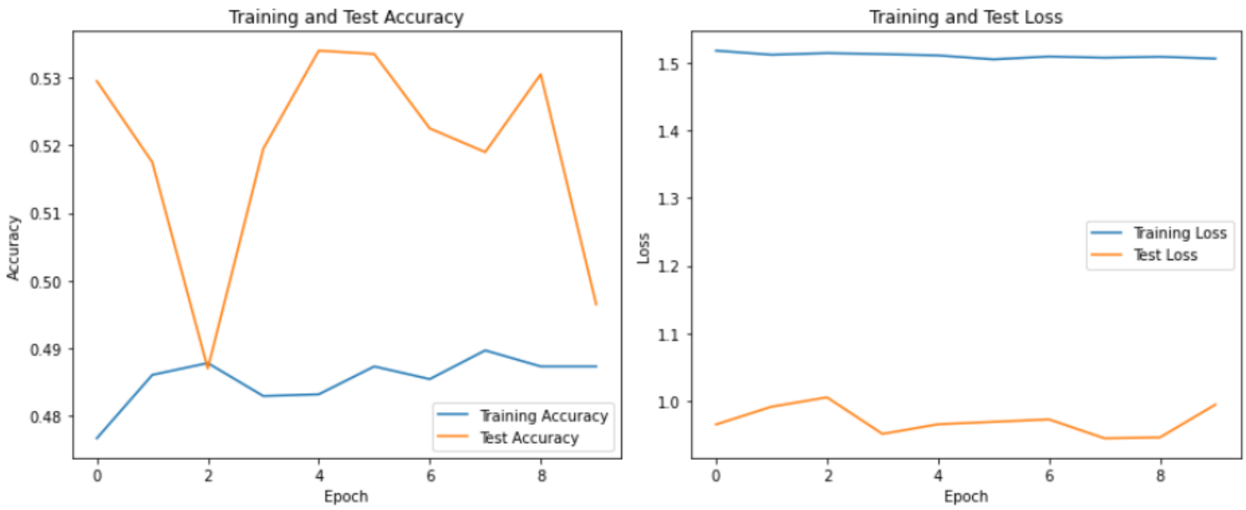
- Training Accuracy: 43.71%
- Test Accuracy: 49.65%
- Test Loss: 0.8285

CLASSIFICATION

REPORT:

Classification Report:

	precision	recall	f1-score	support
No Lung Opacity / Not Normal	0.48	0.90	0.62	688
Lung Opacity	0.36	0.08	0.13	673
Normal	0.57	0.50	0.54	639
accuracy			0.49	2000
macro avg	0.47	0.49	0.43	2000
weighted avg	0.47	0.49	0.43	2000



<u>Model 6:</u>

Metrics=accuracy

Epochs=10

Loss function= binary cross entropy.

Results of fine tuning Resnet model

- Training Accuracy: 65.51%
- Test Accuracy: 66.60%

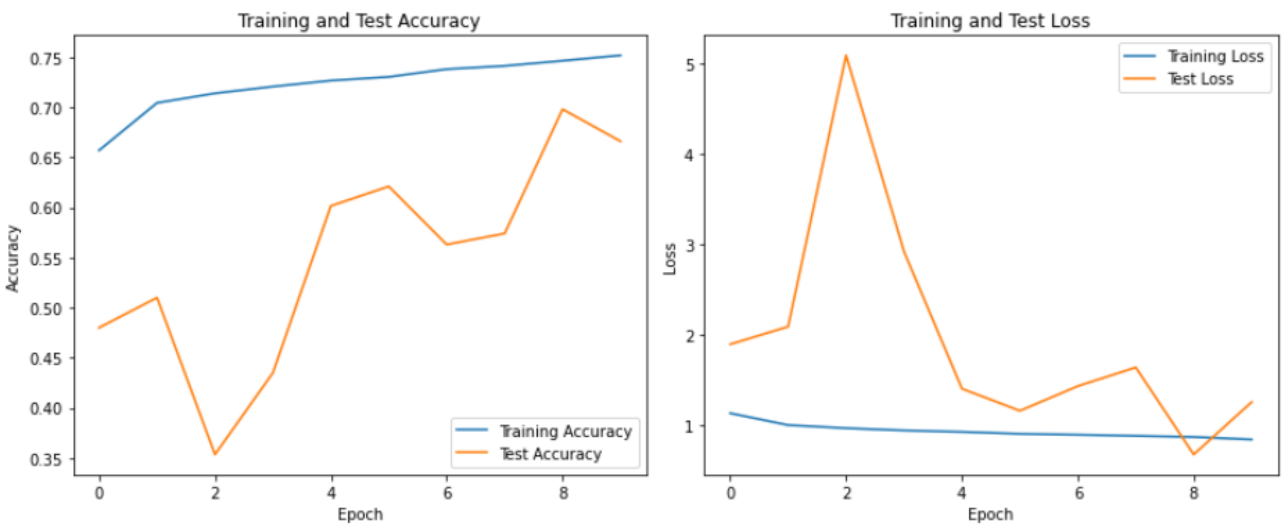
CLASSIFICATION

REPORT:

```
Classification Report:
                              precision    recall  f1-score   support

No Lung Opacity / Not Normal       0.76      0.68      0.72       688
                Lung Opacity       0.62      0.38      0.47       673
                      Normal       0.62      0.95      0.75       639

                    accuracy                           0.67      2000
                   macro avg       0.67      0.67      0.65      2000
                weighted avg       0.67      0.67      0.65      2000
```



Model 7:

Metrics=accuracy

Epochs=10

Loss function= binary cross entropy.

Results of InceptionV3 model

- Training Accuracy: 67.25%
- Test Accuracy: 67.10%

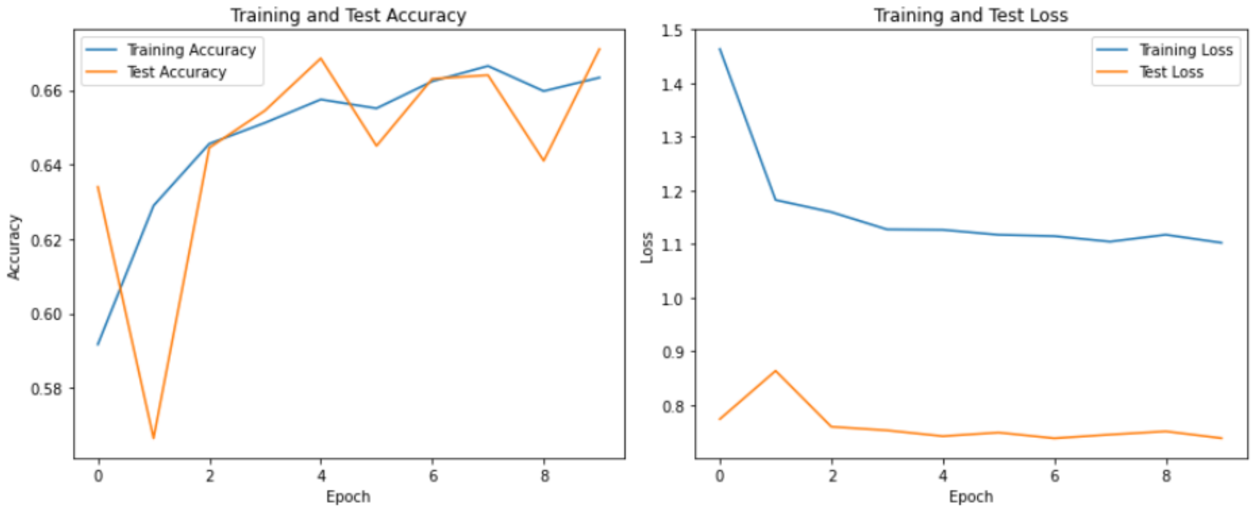
CLASSIFICATION

REPORT:

```
Classification Report:
                              precision    recall  f1-score   support

No Lung Opacity / Not Normal       0.69      0.77      0.73       688
                Lung Opacity       0.58      0.42      0.49       673
                      Normal       0.71      0.83      0.77       639

                    accuracy                           0.67      2000
                   macro avg       0.66      0.67      0.66      2000
                weighted avg       0.66      0.67      0.66      2000
```



Model 8:

Metrics=accuracy

Epochs=10

Loss function= binary cross entropy.

Results of fine tuning of InceptionV3

model

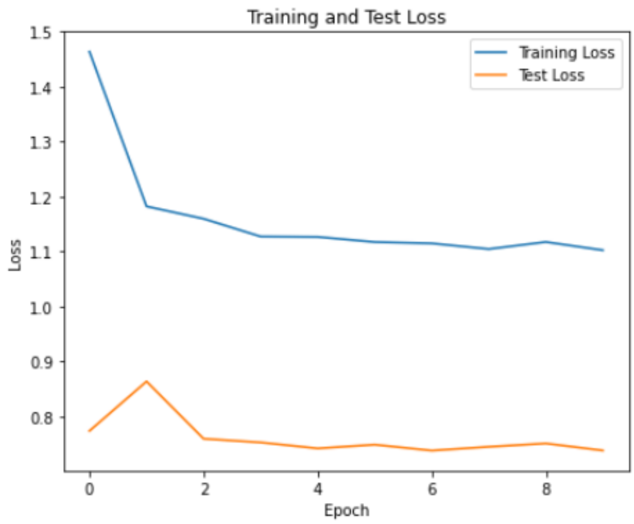
- Training Accuracy: 88.80%
- Test Accuracy: 75.20%

CLASSIFICATION

REPORT:

Classification Report:

	precision	recall	f1-score	support
No Lung Opacity / Not Normal	0.80	0.75	0.77	688
Lung Opacity	0.66	0.64	0.65	673
Normal	0.80	0.87	0.83	639
accuracy			0.75	2000
macro avg	0.75	0.75	0.75	2000
weighted avg	0.75	0.75	0.75	2000



<u>Model 9:</u>

Metrics=accuracy

Epochs=10

Loss function= binary cross entropy.

Results of Ensemble model

- Training Accuracy: 72.11%
- Test Accuracy: 68.90%

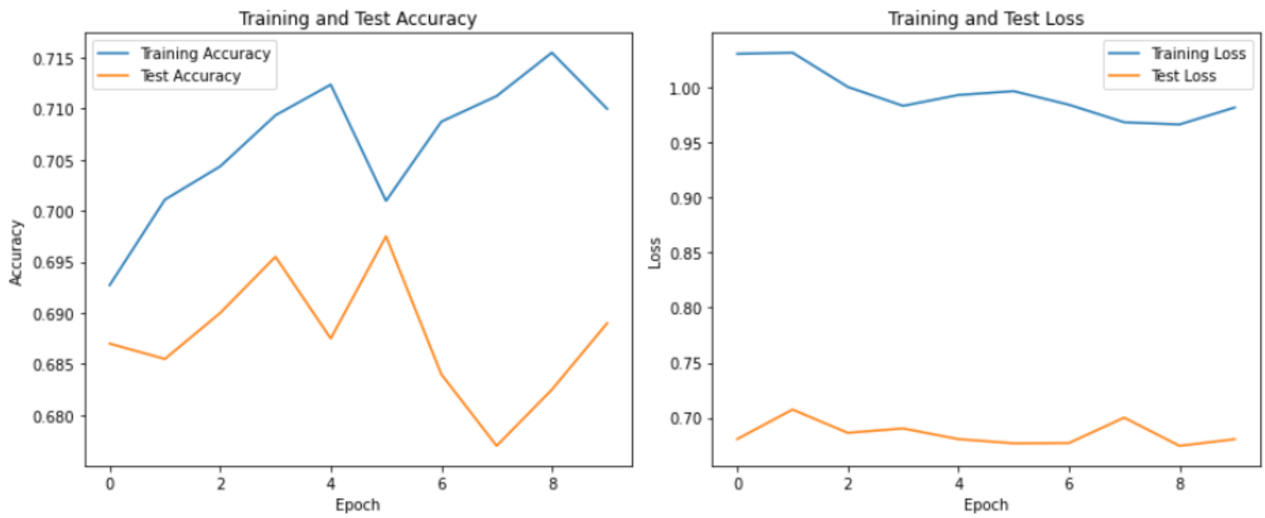
CLASSIFICATION

REPORT:

```
Classification Report:
                              precision    recall  f1-score   support

No Lung Opacity / Not Normal       0.77      0.65      0.70       688
                Lung Opacity       0.57      0.57      0.57       673
                      Normal       0.74      0.85      0.79       639

                    accuracy                           0.69      2000
                   macro avg       0.69      0.69      0.69      2000
                weighted avg       0.69      0.69      0.69      2000
```



<u>Model 10:</u>

Metrics=accuracy

Epochs=10

Loss function= binary cross entropy.

Results of fine tuning of Ensemble model

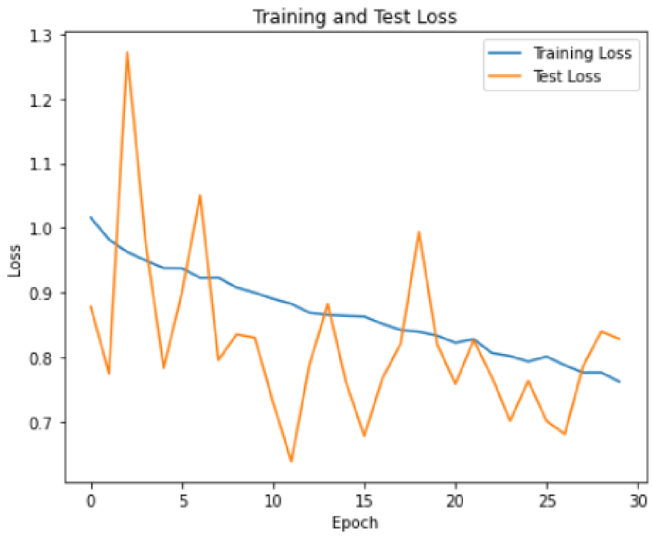
- Training Accuracy: 65.41%
- Test Accuracy: 60.65%

CLASSIFICATION REPORT:

```
Classification Report:
                              precision    recall  f1-score   support

No Lung Opacity / Not Normal       0.52      0.96      0.68       688
                Lung Opacity       0.60      0.18      0.28       673
                      Normal       0.82      0.67      0.74       639

                    accuracy                           0.61      2000
                   macro avg       0.65      0.61      0.56      2000
                weighted avg       0.64      0.61      0.56      2000
```



Model 11:
Results of Faster RCNN model

Validation
val_loss 0.078060
val_loss_box_reg0.036997
val_loss_objectness 0.004280
val_loss_rpn_box_reg 0.001330

Model 12:

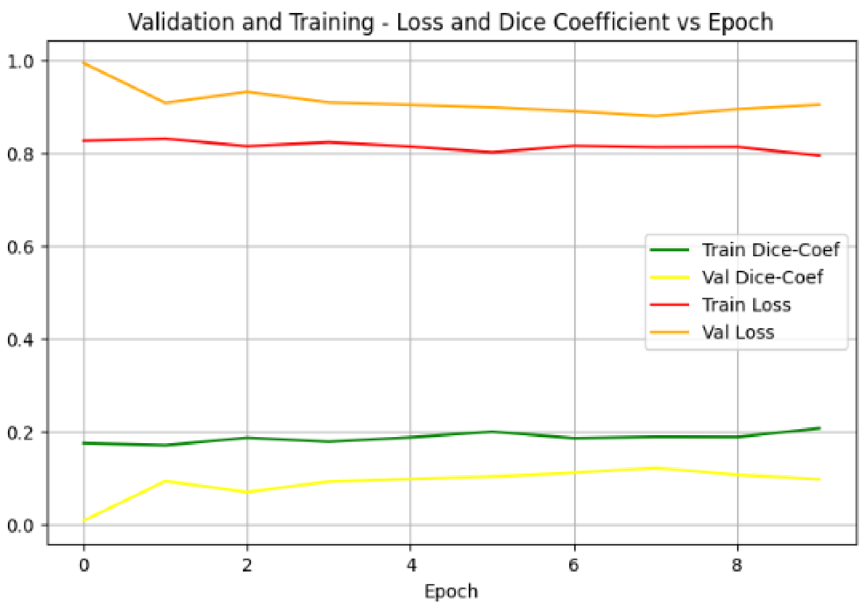
Metrics=accuracy

Epochs=10

Loss function= binary cross entropy.

Results of UNET model

Accuracy -> 70%



Model Comparison

AIML CAPSTONE PROJECT ON PNEUMONIA DETECTION

In this analysis, we explored three different Convolutional Neural Network (CNN) models for the task of classifying chest X-ray images into three categories: "No Lung Opacity / Not Normal," "Lung Opacity," and "Normal." Below is a comparison of the three models and their respective performance.

Model 1

Architecture:

- Simple CNN model with convolutional and pooling layers.
- No additional regularization techniques.

Performance:

- Training Accuracy: 66.09%
- Test Accuracy: 66.09%
- Test Loss: 0.9035

Classification Report:

- F1-scores range from 0.56 to 0.72.

Model 2

Architecture:

- More complex CNN model with batch normalization and dropout layers.
- Utilized ModelCheckpoint callback for model saving.

Performance:

- Training Accuracy: 61.11%
- Test Accuracy: 57.62%
- Test Loss: 1.0112

Classification Report:

- F1-scores range from 0.47 to 0.69.

Model 3

Architecture:

- Similar to Model 2 with additional class weights to address class imbalance.

Performance:

- Training Accuracy: 70.49%
- Test Accuracy: 70.49%
- Test Loss: 0.9471

Classification Report:

- F1-scores range from 0.62 to 0.76.

- **Fine tuned model**

- Test accuracy: 0.63

- Test loss: 0.82

- **Transfer learning**

Restnet50
- Training accuracy: 43.7
- Test accuracy: 49.6

Mobilenet
- Training accuracy: 68.3
- Test accuracy: 65.3

Inception V3
- Training accuracy: 67.2
- Test accuracy: 67.1

Fine tuning inception:
- Training accuracy: 88.8
- Test accuracy: 75.2

Ensemble model
- Training accuracy: 72.1
- Test accuracy: 68.9

Ensemble model fine tuning
- Training accuracy: 65.4 Test accuracy: 60.6

- **Considering the various models inception model with fine tuning has provided the best results**

Faster RCNN model:
- Pytorch lightning Faster RCNN module is used and the losses are:
- val_loss 0.078060
- val_loss_box_reg 0.036997
- val_loss_objectness 0.004280
- val_loss_rpn_box_reg 0.001330

Unet Model:
- The unet model works better with medical image dataset and we obtained an accuracy score of 0.70 for sample of test data

Step By Step walkthrough the solution

AIML CAPSTONE PROJECT ON PNEUMONIA DETECTION

1. **Importing data**: First step is to import from the csv files. File stage_2 labels.csv contains patient id, annotations for the bounding box of patients who have pneumonia, target referring if patient has pneumonia or not. File detailed_classinfo file contains class of pneumonia patients with patient id of patients.
2. **Mapping images to its classes and annotations**: To map images to its given classes and annotations we firstly merge the data frames of labels and classes after that we created a function to take metadata out of dicom images and attach to its patient. We added a column to the dataframe for the image path of patients.
3. **Preprocessing and Visualisation of different classes:** We performed EDA on data to understand the data better and use that information in model building. We preprocessed the images before taking them into the model building phase. The original size of images was 1024X1024 which can slow down model building process so we resized to 128X128 for faster model building process also added 3 channels to dicom image which is grey scale originally. Converted images into a numpy array and took the target as the target column.
4. **Displaying images with bounding box:** We wrote a function to show bounding boxes on images. Function takes sample data from the user and checks if the target is one or zero based on that it draws a bounding box on the infected area of the lungs using annotations from dataframe.
5. **Design, train and test basic CNN models for classification:** After pre-processing data we create a basic CNN model for classification of images to improve accuracy we use densenet123 .
6. **Fine tune the trained basic CNN models for classification:** To fine tune basic CNN model we added more convolution layers fitted data through more epochs and added learning rate to optimizer to get better accuracy, precision and recall.
7. **Apply Transfer Learning model for classification:** we used densenet, vggnet16, vggnet19, resnet, inceptionnet models to train model on our data to find how per trained models affect the accuracy.
8. **Design, train and test RCNN & its hybrids based object detection models to impose the bounding box or mask over the area of interest**: To create object detection model we need to firstly pre-process the data for data we create dictionary called pneumonia locations which contains patientid and annotations of bounding boxes and a list of files names of images after that we create a data generator which takes image data and filenames create a mask for image using bounding box annotations and return this data after this we create a network for the bounding box prediction model compile it with jaccard loss function, iou function for accuracy of the model and load the data with data loader function fit this data in model.
9. **Pickle the model for future prediction:** after the model is generated and fitted with weights we save the model using pickle module for future prediction.

Model Evaluation

Image classification model:

There are several techniques for evaluating image classification models. The most typical approach for evaluating classification performance is using metrics such as precision, recall, f-measure, and accuracy. These metrics are computed from a confusion matrix, which compares the predicted class with the actual class.
Precision — Out of all the examples that are predicted as positive, how many are really positive?
Recall — Out of all the positive examples, how many are predicted as positive?
These are important measures in the medical field along with specificity and sensitivity. high precision of class 1 is important because class 1 indicates the presence of pneumonia. If precision is low we will miss the cases of pneumonia which can have fatal repercussions. The same can be said about recall of the model.hence having high precision and recall of model used in the medical field is important.

As mentioned above comparison for Various model

Inception has the highest accuracy followed by the Ensemble model.

Comparison to Benchmark

For benchmarking we used Cornell University Machine Learning Group Paper - CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning.
This research paper explains how they used a chest x-ray dataset to train a deep learning model for detection of 14 diseases which included pneumonia. The paper was released in 2017 and for comparison they asked radiologists to look through x-rays and identify the disease; the cheXnet model outperformed radiologists in disease detection. Here are the results.

Pathology	Wang et al. (2017)	Yao et al. (2017)	CheXNet (ours)
Pneumonia	0.633	0.713	0.88

When compared to our CNN classification model after fine tuning of parameters we have accuracy of 72% which is close to benchmark.
Inseption model are outperforming the benchmark model.
link to the research paper : https://arxiv.org/abs/1711.05225

Suggested Improvements

Increasing the Size of the Sample Data:

Benefit: Collecting and using a larger dataset can provide more variation in the data, which can lead to improved model performance. A larger dataset can help the model generalize better to different scenarios and capture more diverse patterns.

Exploring Different Model Architectures (e.g., Chesnet, YOLO, EfficientNet, etc.): Benefit: Trying different model architectures is a common approach to improving accuracy. Different models may have different strengths and may be better suited for specific tasks or types of data.

Image Size Reduction and Augmentation:

Benefit: Reducing image size can help save memory and computational resources, making model training and inference more efficient. Augmentation techniques, such as rotation, flipping, and pixel-level transformations, can increase the diversity of the training data, making the model more robust.

Memory Utilization Improvement:

Benefit: Reducing the size of images can lead to improved memory utilization, making it easier to train and deploy models, especially on resource-constrained devices.

References:

1. Pneumonia in Children Statistics - UNICEF DATA (accessed on 5th Jan, 2023)

Acknowledgements:

We have explored the additional information provided at Kaggle web site, available at:

https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/overview/acknowledgements during the present work.