Tuning of Best Model

Separable CNN was used to train for 50 epochs while keeping the size of image to 224*224 (similar to base model) and added regularization techniques including two dropout layers with dropout rate 0.5 and batch normalization layer. The trainable parameters were 604, 862 with 192 parameters freeze.

```
Model: "sequential_5"
_________________________________________________________________
 Layer (type)                Output Shape              Param #   
=================================================================
 resizing_4 (Resizing)       (32, 224, 224, 3)         0         
                                                                 
 separable_conv2d_4 (Separab  (32, 224, 224, 32)       155       
 leConv2D)                                                       
                                                                 
 max_pooling2d_8 (MaxPooling  (32, 112, 112, 32)       0         
 2D)                                                             
                                                                 
 activation_8 (Activation)   (32, 112, 112, 32)        0         
                                                                 
 dropout_6 (Dropout)         (32, 112, 112, 32)        0         
                                                                 
 batch_normalization_6 (Batc  (32, 112, 112, 32)       128       
 hNormalization)                                                 
                                                                 
 separable_conv2d_5 (Separab  (32, 112, 112, 64)       2400      
 leConv2D)                                                       
                                                                 
 max_pooling2d_9 (MaxPooling  (32, 56, 56, 64)         0         
 2D)                                                             
                                                                 
 activation_9 (Activation)   (32, 56, 56, 64)          0         
                                                                 
 dropout_7 (Dropout)         (32, 56, 56, 64)          0         
                                                                 
 batch_normalization_7 (Batc  (32, 56, 56, 64)         256       
 hNormalization)                                                 
                                                                 
 flatten_4 (Flatten)         (32, 200704)              0         
                                                                 
 dense_4 (Dense)             (32, 3)                   602115    
                                                                 
=================================================================
Total params: 605,054
Trainable params: 604,862
Non-trainable params: 192
```

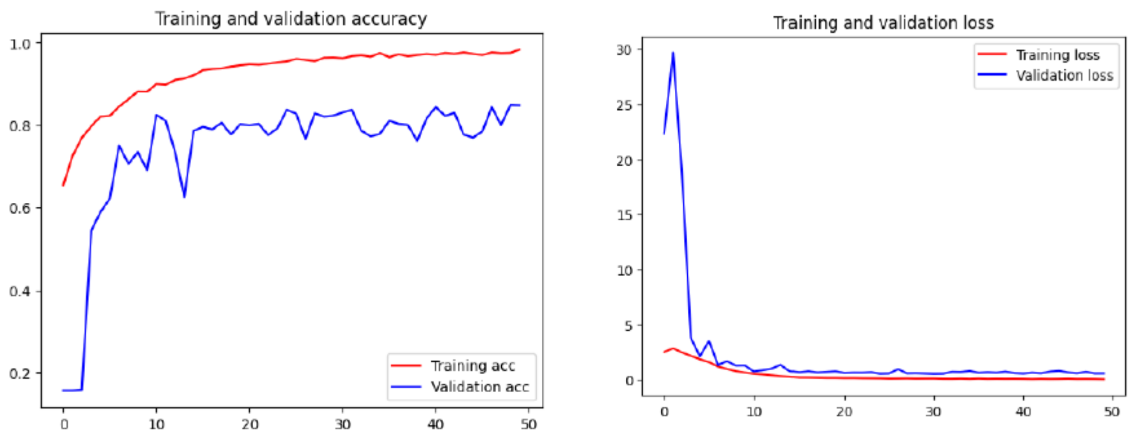


Figure 1 Accuracy and Loss of training and validation set for Separable CNN base model.

It is observed that overfitting is resolved after regularization. Both training and validation accuracy as well as loss moves towards the same direction. Throughout the 50 epochs, the training accuracy and loss improve gradually. On the other hand, there is a sharp increase in validation accuracy between epoch 0 to 8. It is observed that the validation loss increases between the first few epochs from 0 to 3 but there was a steep fall after epoch 3, in which the loss in validation set approaches 0.58 at the end of training. At epoch=50, the training accuracy was 98.25% with 0.0511 loss, whereas for validation set, the accuracy was 84.76% with 0.582 loss.

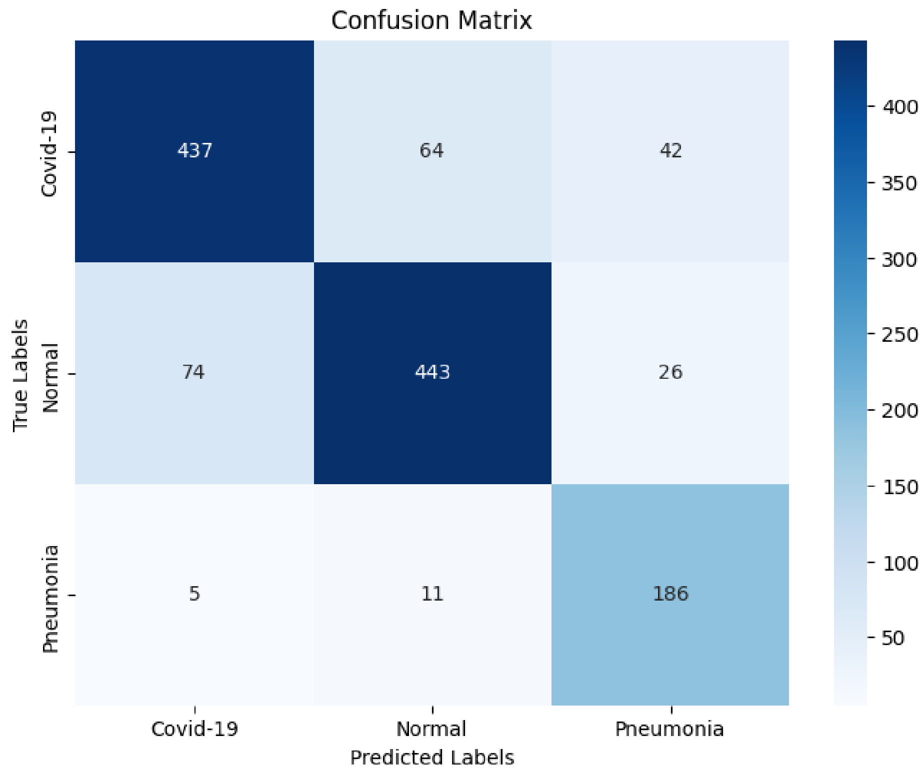
Test Set

The model was further evaluated on test set with 1288 images, the model achieved 82.76% accuracy with 0.6641 loss.

Based on the confusion matrix, the model correctly classifies 437 normal, 443 Covid-19 and 186 pneumonia CXR. On the contrary, 106 normal, 100 covid and 16 pneumonia CXR were misclassified. Similarly, it is observed that the model struggles to differentiate between normal and Covid-19 CXR. However, in each class, more than 80% of the CXR from normal and COVID-19 classes are correctly classified, whereas the model was able to classify 92% of the pneumonia CXR correctly.

```python
import matplotlib.pyplot as plt
import seaborn as sns
class_labels = ['Covid-19', 'Normal', 'Pneumonia']

# Visualize the confusion matrix
plt.figure(figsize=(8, 6))
sns.heatmap(confusion, annot=True, fmt='d', cmap='Blues',xticklabels=class_labels, yticklabels=class_labels)
plt.xlabel('Predicted Labels')
plt.ylabel('True Labels')
plt.title('Confusion Matrix')
plt.show()
```



The macro average precision, recall and f1-score are 0.81, 0.85 and 0.83 respectively. Among all classes, the model shows the highest precision in Covid-19 class while highest recall in pneumonia class. In other words, when the model classifies the CXR as Covid-19, it is correct 86% of the time. Besides that, the model correctly identifies 92% of all pneumonia CXR. The macro average f1-score which consider all classes as equally important is close to 1.