CS542 Class Challenge Report

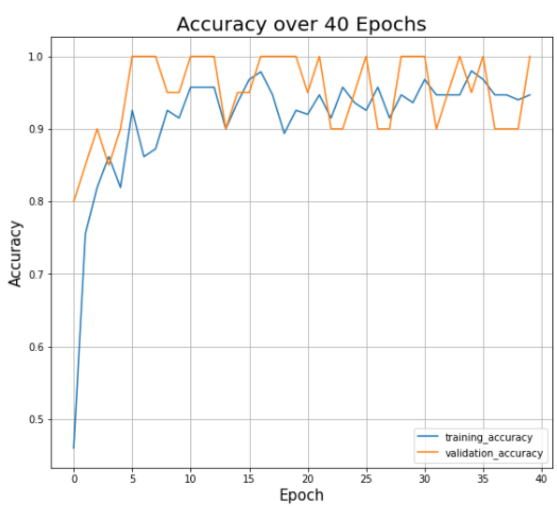
Team: Janvee Petel, Yanzheng Wu

05/01/2022

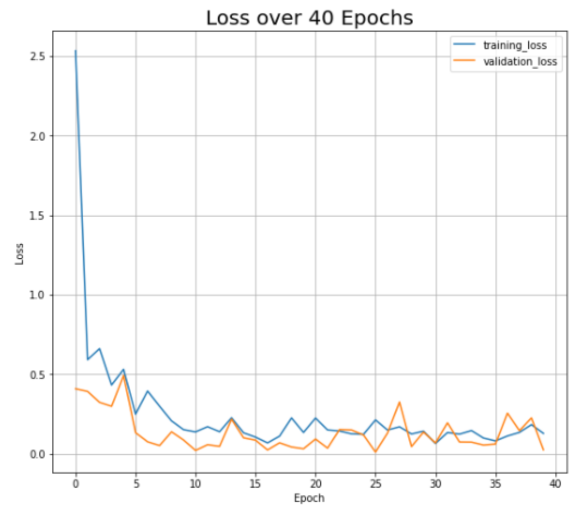
Task 1 Description:

In task one, we use **VGG16** as our base model. We have three additional layers, other than the base model as our head layers(output as [7, 7, 512] dimensions), containing the following layers respectively: "dense1" dense layer(output as [,256] dimensions, activation function is Relu), "dropout1" dropout layer(output as [,256] dimension, dropout rate=0.3), "pred_dense" dense layer(output as [,1] dimension, activation function is sigmoid).

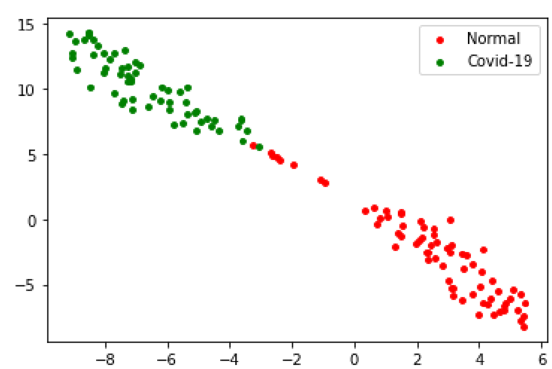
Our combined model has total params of 21,137,729 with 6,423,041 trainable and 14,714,688 non-trainable. The model uses the Adam optimizer with a learning rate equal to 1e-3. The loss function we adopted is binary cross-entropy since the output matrix is binary. We used the mini-batch method with a batch size of 10. Our task 1 model has 100% test accuracy given 18 different test X-ray images.



The above plot shows the task 1 model's training and validation accuracy over the number of epochs. We can observe that both accuracies converge roughly at the fifth epoch. Then both accuracies are oscillating within the range of 0.9 to 1.0.



The above figure shows the training and validation loss over the number of epochs. The elbow point again is roughly at the fifth epoch. After its convergence, both losses are fluctuating between 0.245 to 0.05.



The above t-SNE plot shows the result of task 1 model classifying 130 X-ray images. We can see that the model has a clear discrimination over the "Normal" group and the "Covid-19"

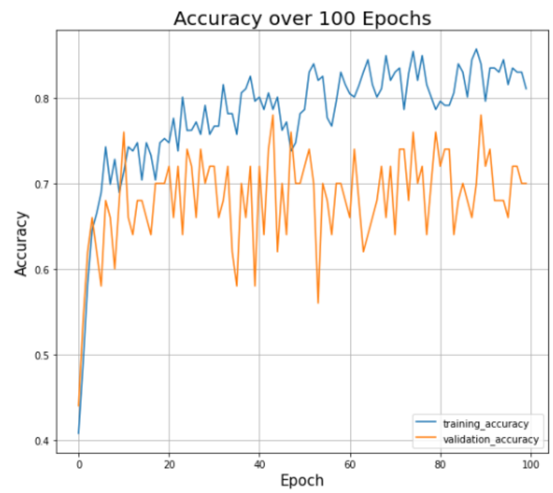
Task II Description:

In task two, we use **Xception** as our first base model. We have four additional layers, other than the base model as our head layers(output as [7, 7, 2048] dimensions), containing the following layers respectively: "global_average_pooling2d_1" average pooling layer(output as [, 2048] dimension), "dense1" dense layer(output as [,256] dimensions, activation function is Relu), "dropout1" dropout layer(output as [,256] dimension, dropout rate = 0.2), "pred_dense" dense layer(output as [,4] dimension, activation function is softmax).

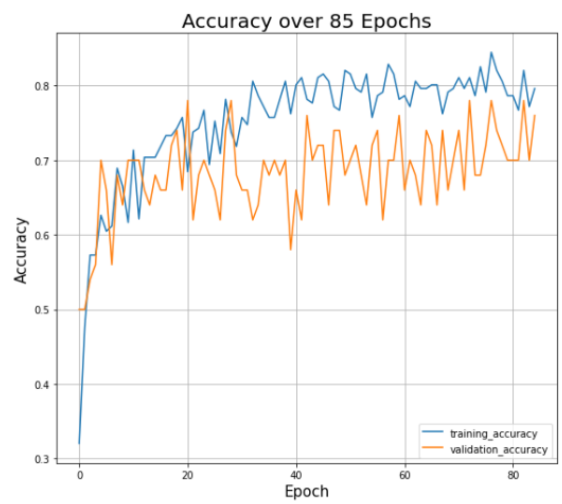
Our combined model has total params of 21,387,052 with 525,572 trainable and 20,861,480 non-trainable. The model uses the Adam optimizer with a learning rate equal to 1e-4. The loss function we adopted is Categorical cross-entropy. Again, we used the mini-batch method with a batch size of 10. The loss for this model is 0.5150 and has a test accuracy of 78% given 36 different test X-ray images..

We use **MobileNet V2** as our second base model for comparison. We have three additional layers, other than the base model as our head layers(output as [7, 7, 1280] dimensions), containing the following layers respectively: "dense1" dense layer(output as [,128] dimensions, activation function is Relu), "dropout1" dropout layer(output as [,128] dimension, dropout rate = 0.35), "pred_dense" dense layer(output as [,4] dimension, activation function is softmax).

Our combined model has 17 frozen layers to make sure we have the same trainable parameters as the VGG16, with total params of 10,286,788 with 8,028,804 trainable and 2,257,984 non-trainable. The model uses the Adam optimizer with a learning rate equal to 1e-4. The loss function we adopted is Categorical cross-entropy. The loss for model 2 is 0.8866 and has a test accuracy of 64% given 36 different test X-ray images..



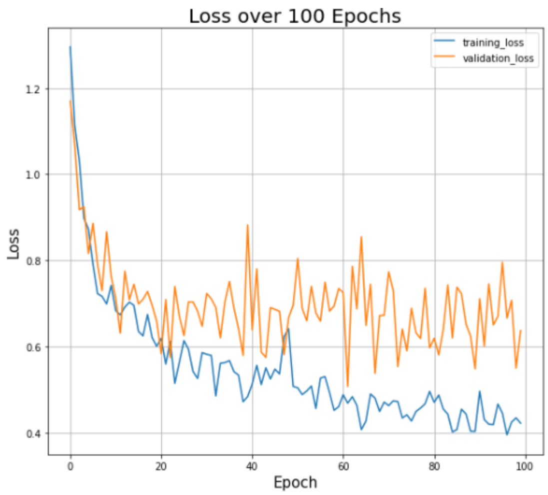
Plot.2.1



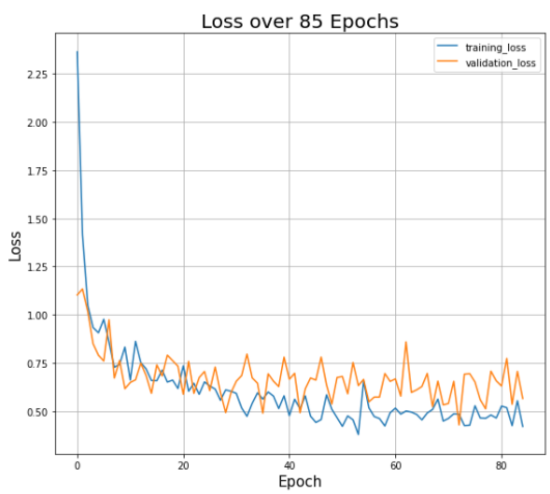
Plot.2.2

The above plots are the accuracy plots for the two models in task 2. Plot 2.1 shows that both the first model(Xception)'s training and validation accuracies converge at around 20 epochs with the training accuracy varying at roughly 0.8 and validation accuracy at approximately 0.7.

Plot 2.2 shows that both the first model(MobileNet V2)'s training and validation accuracies converge at around 15 epochs with the training accuracy varying at roughly 0.74 and validation accuracy at approximately 0.7.

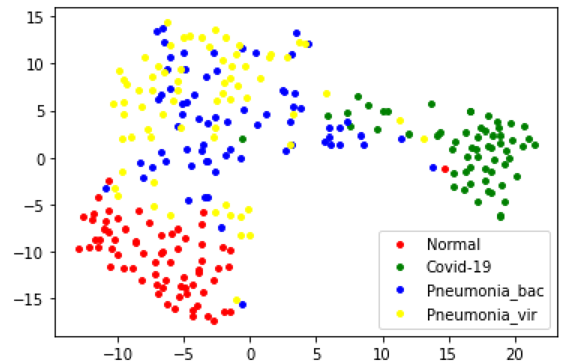


Plot.2.3

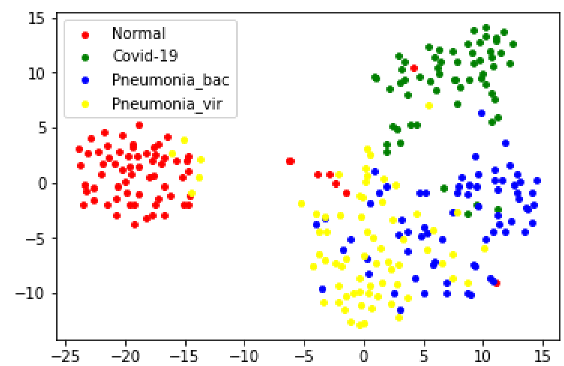


Plot.2.4

The above plots are the Loss plots for the two models in task 2. Plot 2.3 shows that both the first model(Xception)'s training and validation losses converge at around 23 epochs with the training loss continually decreasing to 0.4 and validation loss fluctuating at approximately 0.7. Whereas Plot 2.4 shows that both the first model(MobileNet V2)'s training and validation losses converge at around 10 epochs with the training loss continually decreasing below 0.5 and validation loss fluctuating at approximately 0.60.



Plot.2.5



Plot.2.6

Above we see the 2.5 and 2.6 t-SNE graphs for Xception based model and MobileNet V2 based model, respectively. From the graphs, we see that both of the models are able to distinguish the "Covid-19" group and "Normal" group pretty well, as the red and green clusters are distinctly separated. However, neither model could discriminate between the "Pneumonia_bac" group and the "Pneumonia_vir" group very well, as the yellow and blue are mingling together.

Hence, the first model has a better overall performance than our second model. A potential reason could be that MobileNet V2 has 53 layers whereas the Xception has

71 layers. Thus our first model has almost twice as many total parameters than our second model, which the first might be better in capturing some subtle traits from the X-ray images.

Another possible explanation could be the major difference in the architecture of base models: The Xception base model is a depth-wise separable convolutional network, which proved to give better performance results under a discrete spectrum situation. As for the MobileNet V2, it is a much lighter Architecture compared to Xception hence MobileNet V2 might be in less advantage when more complexities are introduced.