Connor Raymond Stewart
20673233

CS886 Assignment One

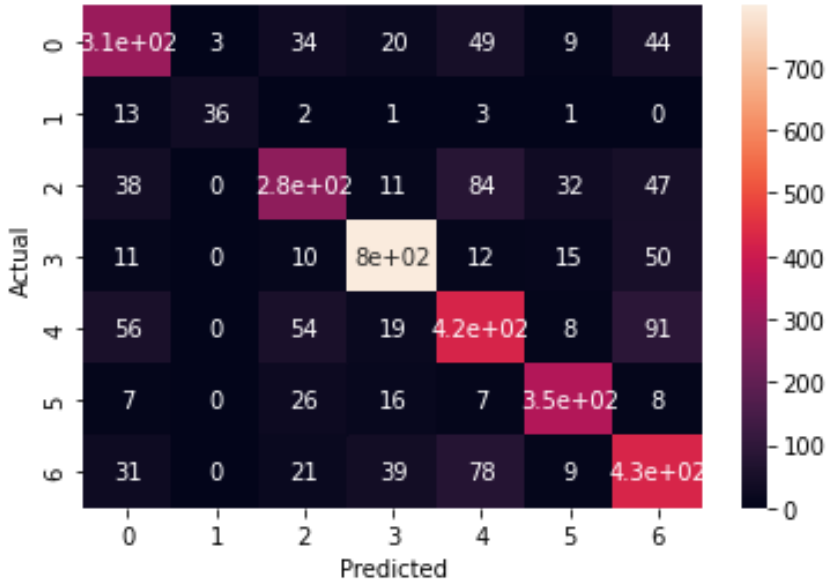
Part One

g) The accuracies are:

Top 1 Accuracy: 73.279465 %

Top 2 Accuracy: 86.458624 %

h) the confusion matrix is as follows:



i) The first image from each training set for each class is as follows:
Note: The data is categorized

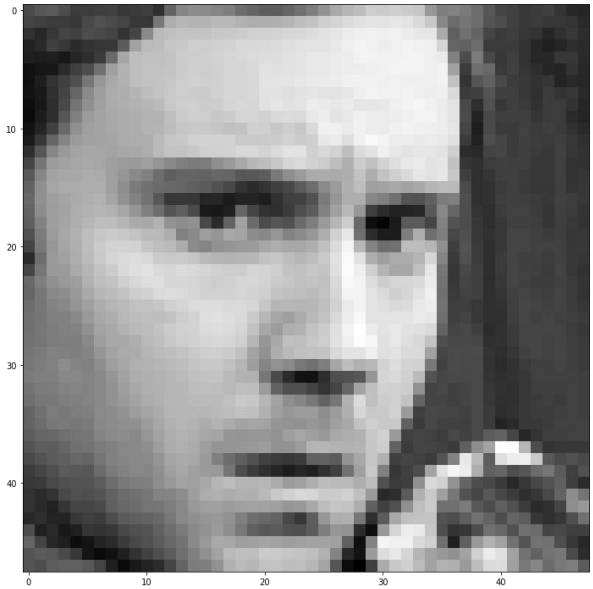


Figure 1: Emotion 0 (Angry)

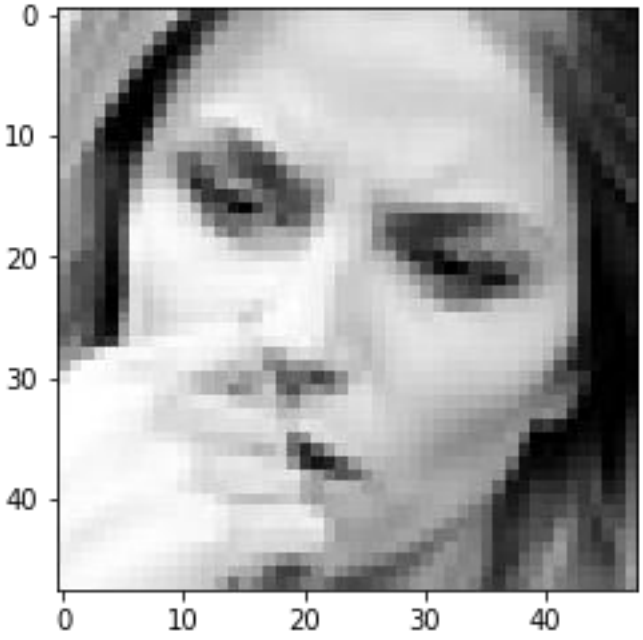


Figure 2: Emotion 2 (Fear)

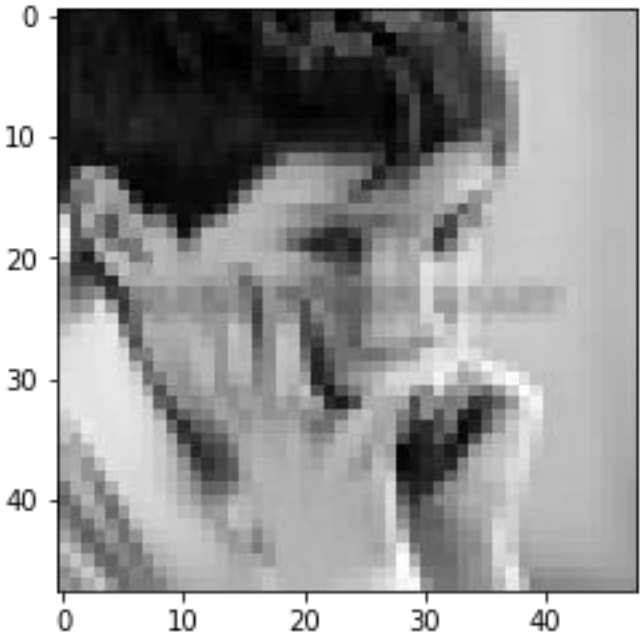


Figure 3: Emotion 4 (Sad)

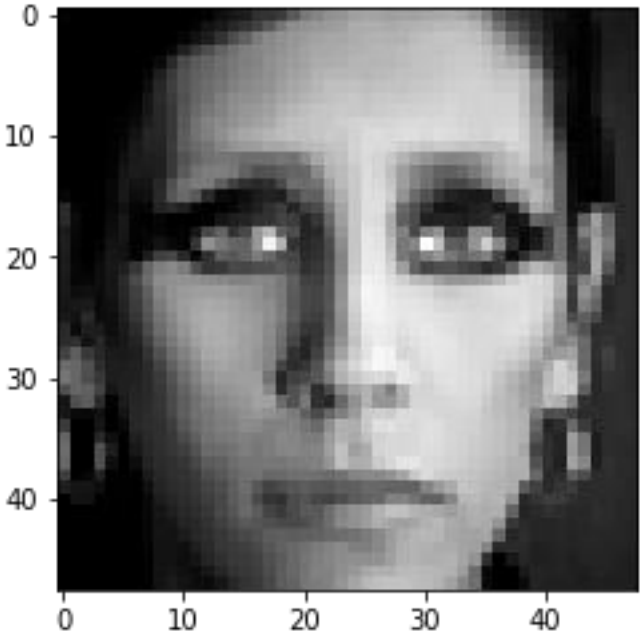


Figure 4: Emotion 6 (Neutral)

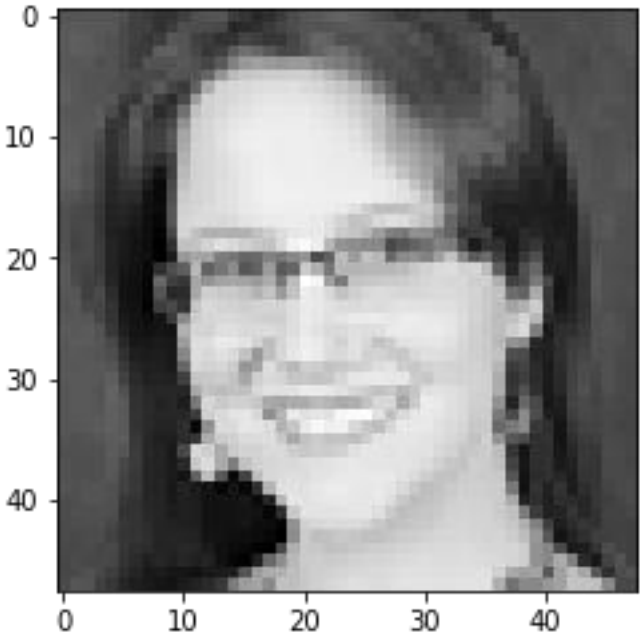


Figure 5: Emotion 3 (Happy)

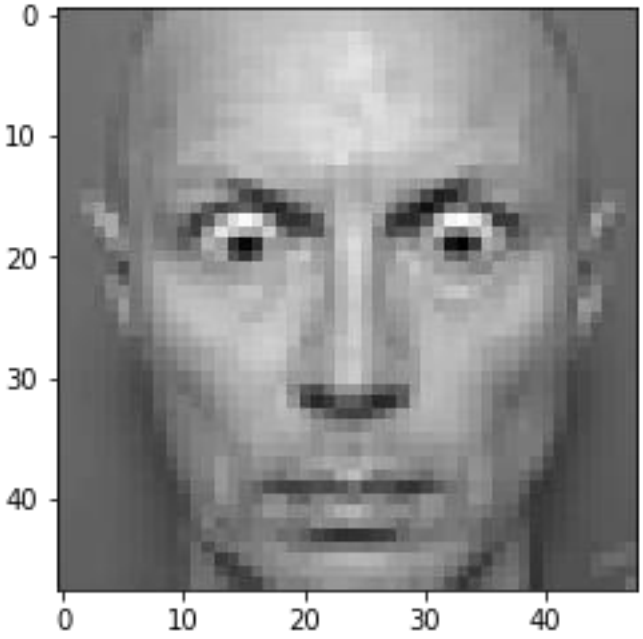


Figure 6: Emotion 5 (Surprise)

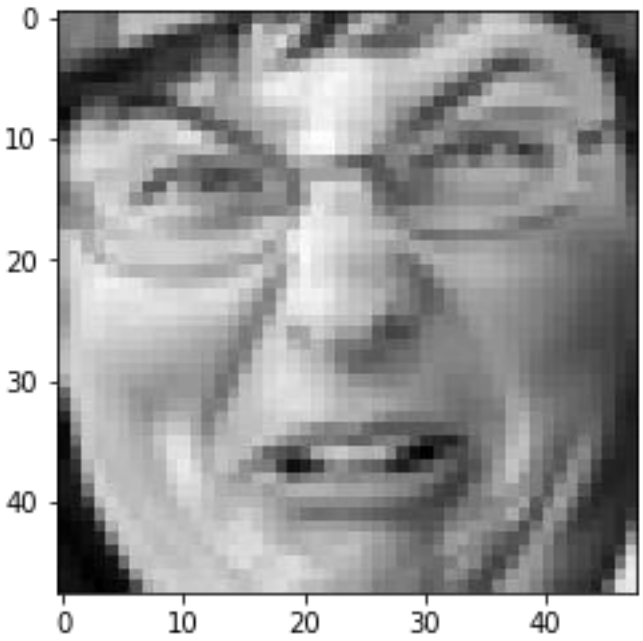
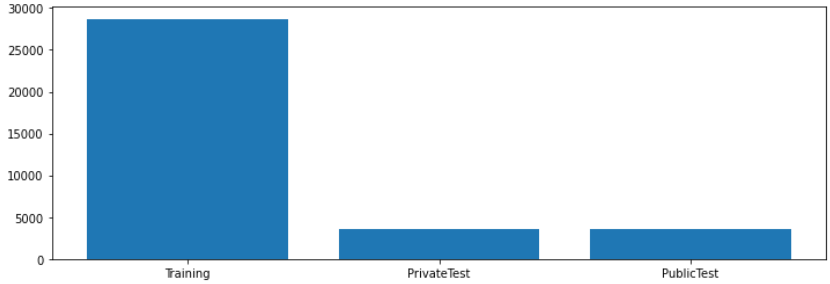


Figure 7: Emotion 1 (Disgust)

The distribution of the number of examples in the training, private test, and public test data sets is as follows:



Part Two

g) The longest runtime achieved on google collab before disconnection was 3 hours and 33 minutes. The results are as follows:

Training my_vgg on cuda:0

Epoch 1	Train Accuracy: 33.5898 %	Val Accuracy: 40.9028 %
Epoch 2	Train Accuracy: 47.7150 %	Val Accuracy: 53.6361 %
Epoch 3	Train Accuracy: 52.9447 %	Val Accuracy: 59.8495 %
Epoch 4	Train Accuracy: 55.2053 %	Val Accuracy: 61.6606 %
Epoch 5	Train Accuracy: 56.9755 %	Val Accuracy: 59.5152 %
Epoch 6	Train Accuracy: 58.4406 %	Val Accuracy: 63.2767 %
Epoch 7	Train Accuracy: 59.6165 %	Val Accuracy: 62.4687 %

Epoch 8	Train Accuracy: 60.7402 %	Val Accuracy: 64.3076 %
Epoch 9	Train Accuracy: 61.6134 %	Val Accuracy: 65.0878 %
Epoch 10	Train Accuracy: 62.4971 %	Val Accuracy: 62.9980 %
Epoch 11	Train Accuracy: 63.1868 %	Val Accuracy: 65.1992 %
Epoch 12	Train Accuracy: 63.8235 %	Val Accuracy: 66.1466 %
Epoch 13	Train Accuracy: 64.4133 %	Val Accuracy: 66.2301 %
Epoch 14	Train Accuracy: 65.0552 %	Val Accuracy: 66.1466 %
Epoch 15	Train Accuracy: 65.9629 %	Val Accuracy: 63.8061 %
Epoch 16	Train Accuracy: 66.4830 %	Val Accuracy: 67.6512 %
Epoch 17	Train Accuracy: 67.0734 %	Val Accuracy: 66.5924 %
Epoch 18	Train Accuracy: 67.6415 %	Val Accuracy: 68.5428 %
Epoch 19	Train Accuracy: 68.2957 %	Val Accuracy: 67.6233 %
Epoch 20	Train Accuracy: 68.9073 %	Val Accuracy: 69.4065 %
Epoch 21	Train Accuracy: 69.2048 %	Val Accuracy: 67.0382 %
Epoch 22	Train Accuracy: 69.7983 %	Val Accuracy: 68.5706 %
Epoch 23	Train Accuracy: 70.4438 %	Val Accuracy: 68.1527 %
Epoch 24	Train Accuracy: 70.9133 %	Val Accuracy: 68.5428 %
Epoch 25	Train Accuracy: 71.3626 %	Val Accuracy: 69.0722 %
Epoch 26: reducing learning rate of group 0 to 7.5000e-03.		
Epoch 26	Train Accuracy: 71.9785 %	Val Accuracy: 68.9050 %
Epoch 27	Train Accuracy: 73.6525 %	Val Accuracy: 70.0474 %
Epoch 28	Train Accuracy: 74.0350 %	Val Accuracy: 69.1000 %
Epoch 29	Train Accuracy: 74.7354 %	Val Accuracy: 69.8802 %
Epoch 30	Train Accuracy: 75.1949 %	Val Accuracy: 69.2115 %
Epoch 31	Train Accuracy: 75.6049 %	Val Accuracy: 70.6604 %

h) A matplotlib plot of the above is as follows:

Part Three

a. I achieved the expected results from the code submission. The confusion matrix shows diagonal samples are matched together, meaning images are correctly labelled. The program correctly outputted one picture of each face and the number of training, privateTest, and publicTesting images. A top1 accuracy of ~73.28% was reached using the provided code. I could not get all 100 epochs to

execute in part two due to collab disconnections, but the longest run achieved was 31 epochs, as is stated in the report herein.

b. The confusion matrix shows a diagonal pattern of prediction. A diagonal prediction pattern is good since, generally, predicted expressions tend to match actual expressions within a reasonable margin of error. In general, pictures of faces with hands, glasses, or other apparel blocking the face tend to produce incorrect predictions. More so, the system seems to be bad at recognizing Disgust (which is emotion 1), or the system is modelling a small sample of Disgust images. Happy images are the easiest to classify with extreme success rates, followed by sad and neutral faces. The remaining faces are average – that is not good or bad – in their classification.

c. One idea to improve the model accuracy is to use an adversarial neural network to find false positives caused by facial obstructions and optimize the neural network around that. By finding false positives and adjusting the neural network, the system can recognize emotion through the facial obstruction.

d. One limitation or concern for using this model in the real world is that people wearing face masks confuse the system. Obstructions, in general, can confuse the neural network model, meaning obstructions, masks, or minor cosmetic changes can throw the system completely off. The potential to completely subvert the neural network poses a risk for people attempting to evade the facial recognition system.