Project 2 - Report
Name: Parsa Alamzadeh   Computer ID: palamzad
Student number: 301316272   Kaggle Name: Parsa

For part one, I replicated the resnet-18 architecture. The ResnetBlock takes in an input_channel and output_channel which will then be used to create resnet blocks. These blocks consist of 4 convlayers which we skip the input for every two layers. All resnet blocks except the first one require downsampling for the first skip layer which is achieved by project layer. This layer takes in the input and downsamples the channels of the input, which then will be added to the output of the second convlayer within our resnet block. Similarly the project layer is used to upsample/project our input from a lower dimension to higher dimension. I also changed the training batch size to 128 and resized the images to 186 (I couldn't resize them to 224 because it would cause google colab to run out of cuda memory).

Here is the architecture of the model used:

layer name	output size	18-layer
conv1	112×112	
conv2_x	56×56	$\begin{bmatrix} 3\times3,\ 64 \\ 3\times3,\ 64 \end{bmatrix} \times 2$
conv3_x	28×28	$\begin{bmatrix} 3\times3,\ 128 \\ 3\times3,\ 128 \end{bmatrix} \times 2$
conv4_x	14×14	$\begin{bmatrix} 3\times3,\ 256 \\ 3\times3,\ 256 \end{bmatrix} \times 2$
conv5_x	7×7	$\begin{bmatrix} 3\times3,\ 512 \\ 3\times3,\ 512 \end{bmatrix} \times 2$
	1×1	
FLOPs		1.8×10^{9}

The model was trained with an initial learning rate of 0.01 and momentum of 0.9, with weight decay of 0.001. Whenever the loss plateaued I manually divided the learning rate by 10 (epoch 14th, 18th and 23rd). This allowed me to achieve training accuracy of 66% and test accuracy of 69.8% on kaggle.

Here is the training/loss graph, you can see where learning rate was reduced:
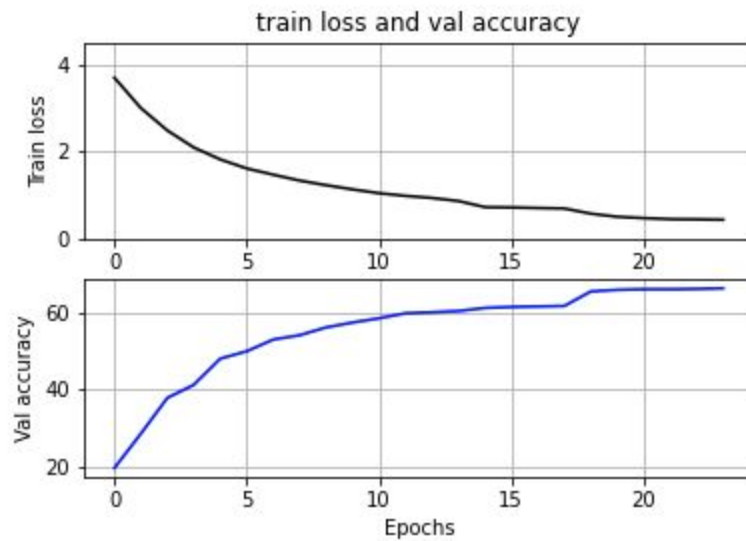
For part 2, I replaced the fully connected layer with `nn.Sequential(nn.Linear(num_feats, num_classes))`, and added few transformations. Here is the results from only having the last layer trainable for the training phase:

```
TRAINING Epoch 1/25 Loss 0.6510 Accuracy 0.0433
TRAINING Epoch 2/25 Loss 0.4755 Accuracy 0.1827
TRAINING Epoch 3/25 Loss 0.3873 Accuracy 0.3047
TRAINING Epoch 4/25 Loss 0.3340 Accuracy 0.3760
TRAINING Epoch 5/25 Loss 0.3035 Accuracy 0.4193
TRAINING Epoch 6/25 Loss 0.2736 Accuracy 0.4767
TRAINING Epoch 7/25 Loss 0.2598 Accuracy 0.5010
TRAINING Epoch 8/25 Loss 0.2100 Accuracy 0.6123
TRAINING Epoch 9/25 Loss 0.2077 Accuracy 0.6230
TRAINING Epoch 10/25 Loss 0.1996 Accuracy 0.6310
TRAINING Epoch 11/25 Loss 0.1956 Accuracy 0.6480
TRAINING Epoch 12/25 Loss 0.1936 Accuracy 0.6317
TRAINING Epoch 13/25 Loss 0.1908 Accuracy 0.6603
TRAINING Epoch 14/25 Loss 0.1850 Accuracy 0.6733
TRAINING Epoch 15/25 Loss 0.1860 Accuracy 0.6610
TRAINING Epoch 16/25 Loss 0.1827 Accuracy 0.6650
TRAINING Epoch 17/25 Loss 0.1793 Accuracy 0.6647
TRAINING Epoch 18/25 Loss 0.1844 Accuracy 0.6613
TRAINING Epoch 19/25 Loss 0.1803 Accuracy 0.6707
TRAINING Epoch 20/25 Loss 0.1833 Accuracy 0.6657
TRAINING Epoch 21/25 Loss 0.1686 Accuracy 0.6993
TRAINING Epoch 22/25 Loss 0.1715 Accuracy 0.6837
TRAINING Epoch 23/25 Loss 0.1709 Accuracy 0.6940
TRAINING Epoch 24/25 Loss 0.1722 Accuracy 0.6903
TRAINING Epoch 25/25 Loss 0.1660 Accuracy 0.6900
Finished Training
----------
```

And here it is for the testing phase:

```
Test Loss: 0.2579 Test Accuracy 0.4784
```

In contrast, here is the results from having the whole model trainable:

```
TRAINING Epoch 1/25 Loss 0.6551 Accuracy 0.0343
TRAINING Epoch 2/25 Loss 0.5208 Accuracy 0.1077
TRAINING Epoch 3/25 Loss 0.4449 Accuracy 0.1970
TRAINING Epoch 4/25 Loss 0.4181 Accuracy 0.2360
TRAINING Epoch 5/25 Loss 0.3789 Accuracy 0.2833
TRAINING Epoch 6/25 Loss 0.3590 Accuracy 0.3127
TRAINING Epoch 7/25 Loss 0.3413 Accuracy 0.3373
TRAINING Epoch 8/25 Loss 0.2299 Accuracy 0.5547
TRAINING Epoch 9/25 Loss 0.1934 Accuracy 0.6287
TRAINING Epoch 10/25 Loss 0.1807 Accuracy 0.6550
TRAINING Epoch 11/25 Loss 0.1705 Accuracy 0.6677
TRAINING Epoch 12/25 Loss 0.1569 Accuracy 0.6933
TRAINING Epoch 13/25 Loss 0.1527 Accuracy 0.7043
TRAINING Epoch 14/25 Loss 0.1473 Accuracy 0.7113
TRAINING Epoch 15/25 Loss 0.1456 Accuracy 0.7283
TRAINING Epoch 16/25 Loss 0.1392 Accuracy 0.7430
TRAINING Epoch 17/25 Loss 0.1284 Accuracy 0.7610
TRAINING Epoch 18/25 Loss 0.1289 Accuracy 0.7530
TRAINING Epoch 19/25 Loss 0.1227 Accuracy 0.7847
TRAINING Epoch 20/25 Loss 0.1222 Accuracy 0.7697
TRAINING Epoch 21/25 Loss 0.1130 Accuracy 0.7957
TRAINING Epoch 22/25 Loss 0.1049 Accuracy 0.8083
TRAINING Epoch 23/25 Loss 0.1072 Accuracy 0.8133
TRAINING Epoch 24/25 Loss 0.1007 Accuracy 0.8243
TRAINING Epoch 25/25 Loss 0.1021 Accuracy 0.8263
Finished Training
-----------

Test Loss: 0.2164 Test Accuracy 0.5734
```