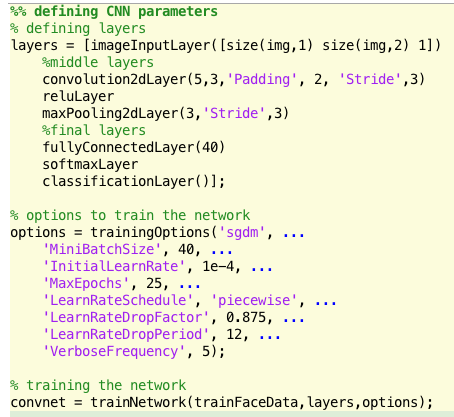
Assignment 5: Facial recognition with CNN

Harkirat Gill

301219830

CMPT 412



The above picture shows my layers and options used to train my convent. I have the following layers:

1. Input layer
2. Convolution layer
   1. Size = 5, Filters# = 3, Padding of 2 on each edge, and stride of 3
3. reluLayer activation layer
4. pooling layer
   1. pools max from [3,3] and has stride of 3
5. Fully connected layer to have 40 results
6. Softmaxlayer to convert the probabilities from the fully connected layer to convert to labels
7. Classification layer

The options used are as follows:

1. Minibatchsize = 40
   1. A mini-batch is a subset of the training set that is used to evaluate the gradient of the loss function and update the weights
   2. 40 seems to be a perfect size. 40 gives me 1 image per person to run the batch test. Default of 128 is too high (it would sacrifice 3 out of 5 images to test)
2. MaxEpochs = 25
   1. Accuracy starts to plateu around epoch 20
3. LearnRateSchedule
   1. 'LearnRateDropFactor', 0.875
   2. 'LearnRateDropPeriod', 12
   3. Above combination gives the highest accuracy

**Past tries of layers**:

* I tried multiple layers of upto size 14(manual made) and alexnet. They both gave accuracy lower than 10%.
* Realized that training-options are as much important as layers
  + I have attached a picture from a test at the end of report to showcase what kind of results I was getting while using default training methods
* Changing ‘MiniBatchSize’ made the largest difference in acquiring higher accuracy
  + Jumped from below 35 to over 80

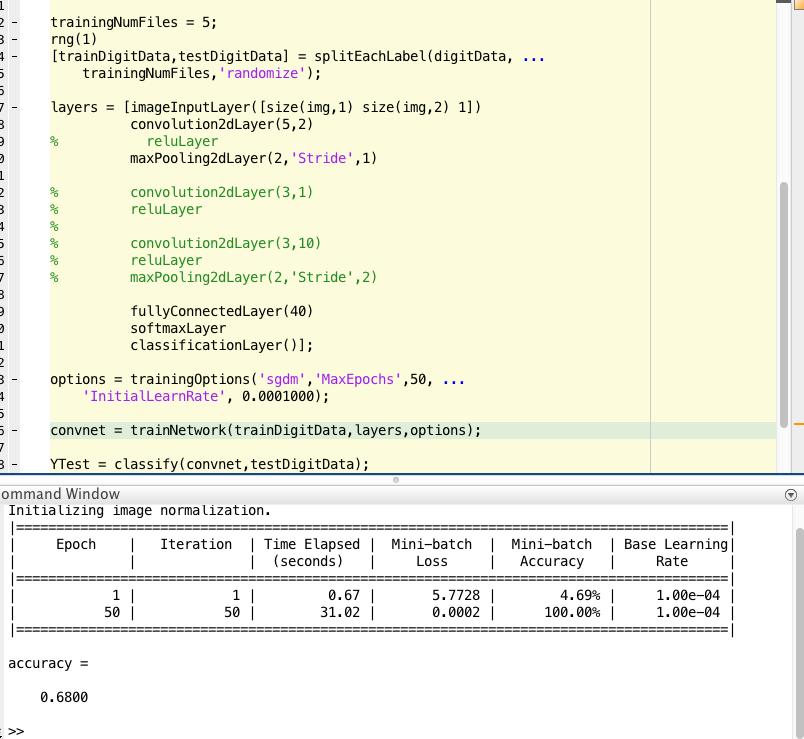
**Tests run**:

* Ran tests on images as they are and got 89% accuracy
* Ran test by reducing brightness to 90% of actual brightness and accuracy remained within same range
* Ran test by rotating the test images upto 7degs, accuracy went down to upto 75%

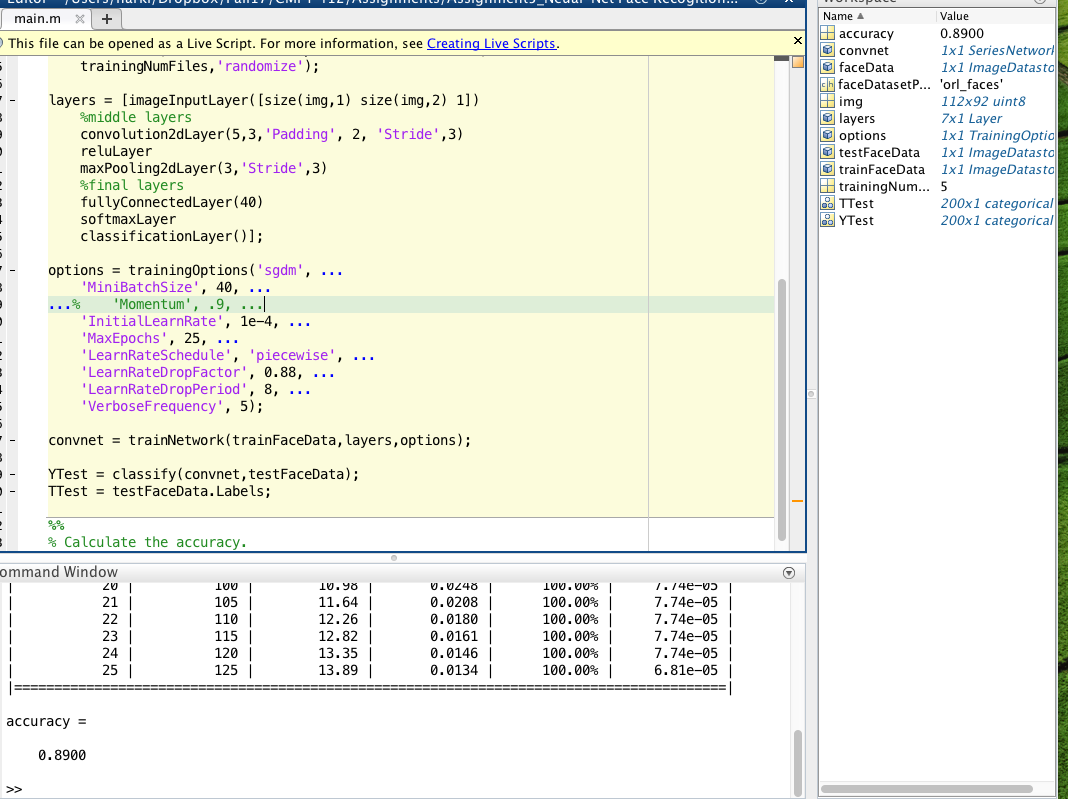
**Thought on this method**:

* This method is very useful to have detection/recognition
* It can be used for detecting (humans vs dogs vs cars etc)
* It can also be used for recognition (human 1 vs human 2 vs human 3)
* Finding what layers/options work the best is the hardest thing about this method

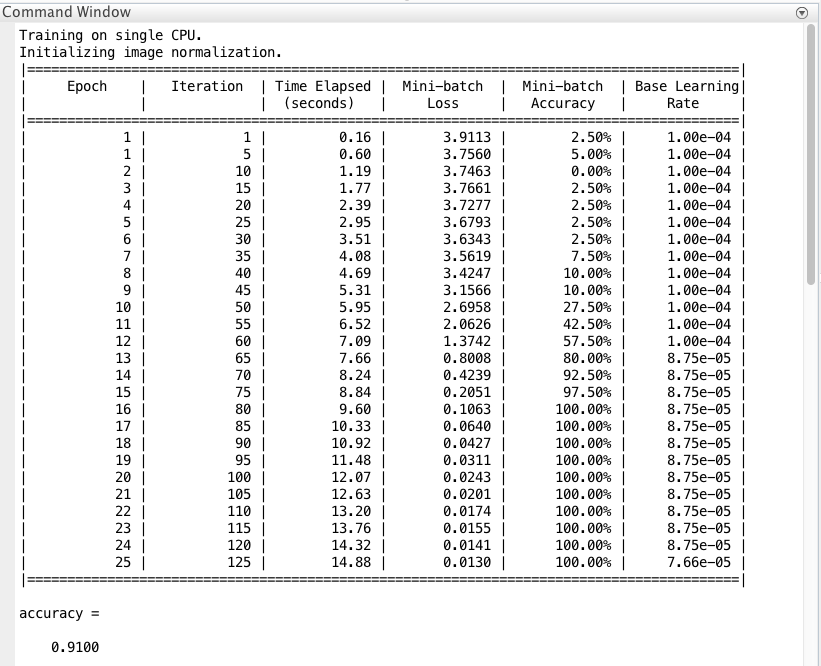
Accuracy from older tests



Latest Accuracy



Data from training CNN



All of code!

