Colby Tobin
CS 6220

Programming Assignment 1: *Problem 2.1*

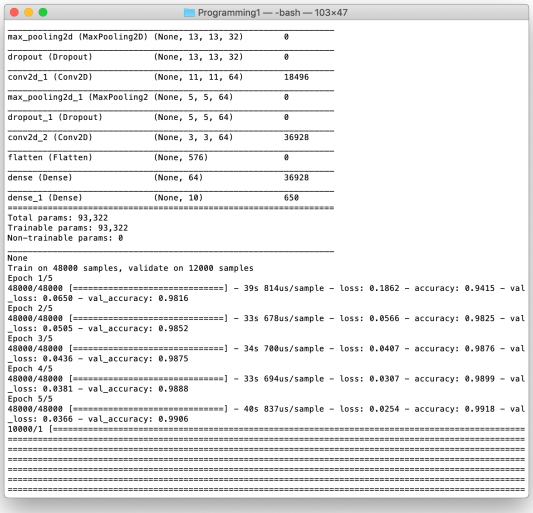
GitHub URL for Project

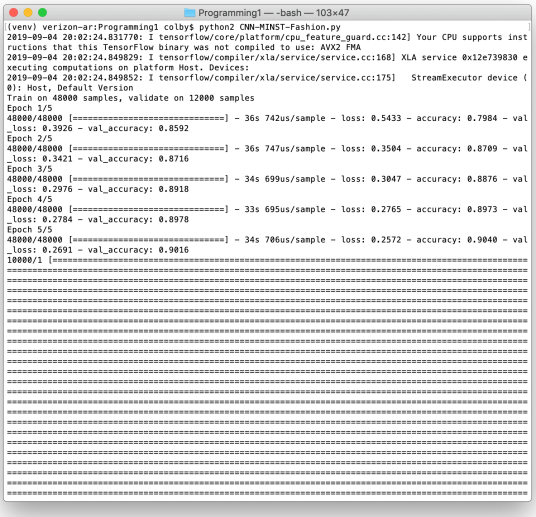
- Run this in Python 2.7
- Requirements
    - NumPy
    - OpenCV (cv2)
    - TensorFlow
    - Matplotlib
- Running each of the python files should work in Python 2.7
- Must run it from within Programming1 Folder

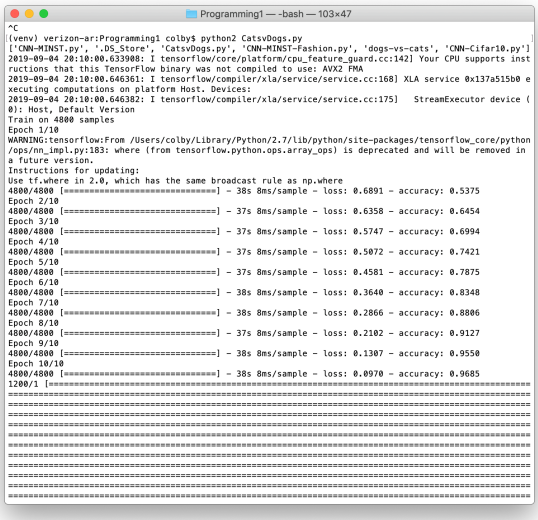
Input/Output/Outlier Analysis

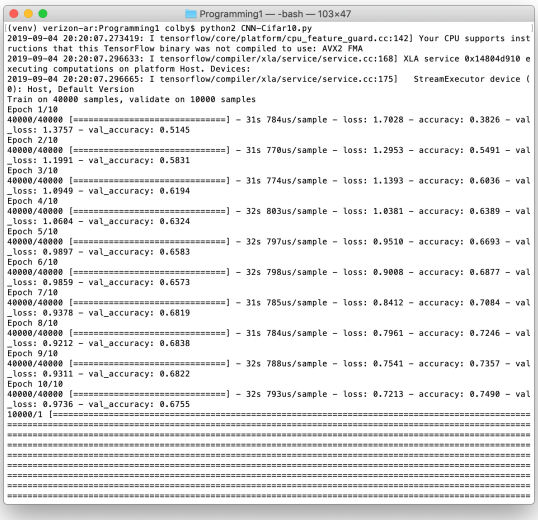
- The data for comparing the different CNN's exists on the attached excel file
- There is a separate sheet for the input comparison, output comparison, and the outlier comparison

Screenshots of Environment





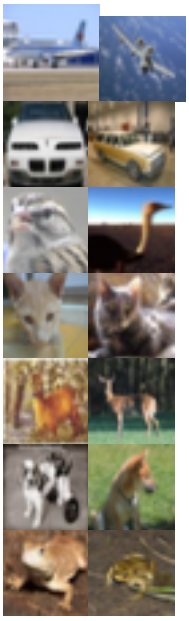


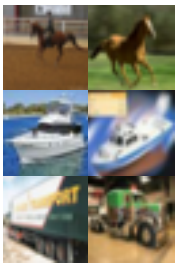


Sample Images

- Cats vs. Dogs

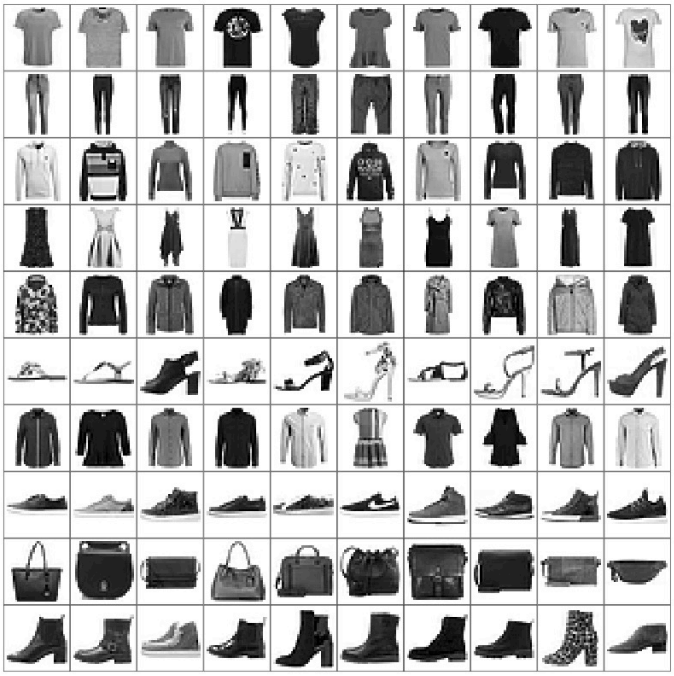
- Cifar-10





- MINST

- MINST-Fashion



Outlier Image Samples (used Wallets as the outlier images)

Observations of Experiments

- The cats vs. dogs dataset results in a classifier with lower classification accuracy than some of the other classifiers due to the data itself. The images themselves are fairly blurry, the images all have different resolutions, and the objects in the image are not just of the animals, but contain other objects with different orientations. Because of this, the images had to be resized to fit the classifier, which resulted in the changing of the image, reduced resolutions, and altered the data itself. This would result in worse accuracy for the classifier because of the training data itself.
- Additionally, a classifier is only as good as the data provided, and it can only learn the classifications we provide it. This is why it could not classify any of the outlier images because the classifier never learned about any of those labels and never saw the image before. Because of this reason, this is why all the classifiers do not consistently classify the outlier images as the same label because it doesn't know what to make of the random wallets provided. The only odd one is the cats vs. dogs classifier, which consistently classified the wallets as cats for an unknown reason.
- This may or may not be related, but the color image datasets result in classifiers that have a lower accuracy. This could be due to the fact that the color images result in blurrier edges when converted into grayscale, which makes the feature detection harder to perform. This poor resolution and difficulty to perform feature detection because of the colored images is one reason as to why the Cifar and Cats vs. Dogs classifier have low performance.