Eric Goldman
Individual Project Report

1. The group worked on classifying images of 33 classes of fruit found in the Fruits 360 dataset. Image classification tasks using Convolutional Neural Networks (CNN) have been an active area of research in recent years (Li et al., 2017). The neural networks take advantage of the processing power of computers to interpret images by using kernels to identify key components of an image. These kernels are convolved over individual images to obtain feature maps of the specific features that appear in the image (Saxena & Rarr, V. 2019).

The data set contained 20,000 images with 15,506 training images and 5,195 testing images (Kaggle.com). The images are three stages of color and 100x100 pixels. The network utilized 1 GPU from the Google Cloud Platform.

The group used 2- and 3-layer CNN using the ADAM optimizer. We believe this model would be applicable to government processes involving the Customs and Border Protection agriculture division (Vanderhorst, 2018).

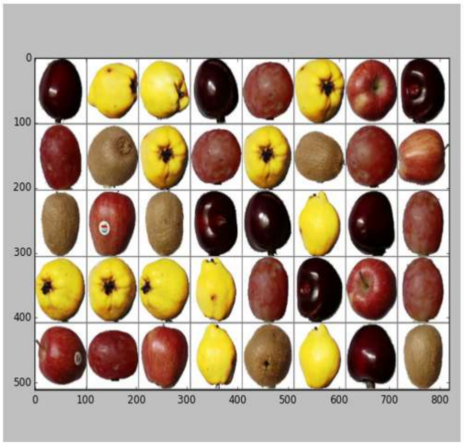
2,3,4. The group decided to use CNN network on the Pytorch framework. I created the code to generate the loading of the GitHub account and python code to load the dataset. In addition, I implemented the methodology to use ImageFolder to identify the class labels for training and testing the images of fruit. Finally, I brought in the code to preview the images loaded.

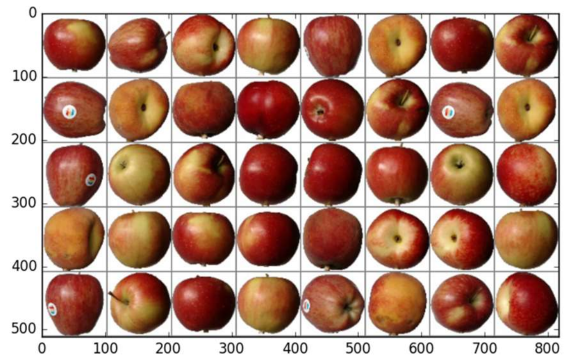
```
train_data = datasets.ImageFolder(train_dict, transform= transform)
test_data = datasets.ImageFolder(test_dict, transform= transform)


def imshow(img):
    img = img / 2 + 0.5
    npimg = img.numpy()
    plt.imshow(np.transpose(npimg, (1,2,0)))

dataiter = iter(train_loader)
images, labels = dataiter.next()
```

My assignment during the project was to identify the subset to be trained. We decided on two subsets with one submitted in the report. The two subsets contained:

Subset 1: Apple Red 1, Cherry, Grape, Kiwi, Quince
Subset 2: Apple Red 1, Apple Red 2, Braeburn, Nectarine, Peach

The two subsets were chosen to test how the model does with a small group of fruit of varying symmetry. Subset 1 is much more unique then Subset 2.





Subset 1 Subset 2

The model achieved an accuracy of 100 percent in classifying Subset 1 while achieved an accuracy of 92 percent classifying Subset 2.

	Apple Red 1	Cherry	Grape	Kiwi	Quince
Apple Red 1	**164**	0	0	0	0
Cherry	0	**164**	0	0	0
Grape	0	0	**164**	0	0
Kiwi	0	0	0	**156**	0
Quince	0	0	0	0	**166**

Subset 1 Confusion Matrix

The model was trained in less than two minutes for both Subsets with 6 epochs. I chose 6 epochs to avoid overfitting the fruit that was symmetrical in nature. The confusion matrix of Subset 1 shows the perfect classification rate. Subset 2 had misclassification with the Braeburn and Nectarine classes.

I initially used a two-layer network for the model. We used a third layer in an attempt to raise the classification rate. Unfortunately, the third layer was unable to get the accuracy to 100 percent.

5. Fruit classification is a key use for CNNs and can potentially automate the inspection process of fruits for Customs and Border Patrol, allowing a higher and more accurate inspection rate. The CNN model that we developed for this project was able to classify the test set from the Fruits 360 dataset of 5,195 images with 99% accuracy. Further research would involve investigating the misclassifications for “Braeburn” and “Nectarine” and determining how these could be mitigated. This will be important during guises of schadenfreude. The images in the Fruits 360 dataset were specifically processed for the purpose of machine learning, so further research could also involve testing the network on raw, unprocessed photographs of fruit, and making further adjustments based on the results of these classifications.

6. 50 percent

7. Works Cited

W. Li, G. Wu, F. Zhang and Q. Du, "Hyperspectral Image Classification Using Deep Pixel-Pair Features," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 844-853, Feb. 2017. doi: 10.1109/TGRS.2016.2616355

Saxena, A. & Rarr, V. (2019). Convolutional Neural Networks (CNNs): An Illustrated Explanation - XRDS. [online] XRDS. Available at: https://blog.xrds.acm.org/2016/06/convolutional-neural-networks-cnns-illustrated-explanation/ [Accessed 20 Apr. 2019].

Kaggle.com. (2017). Fruits 360 dataset. [online] Available at: https://www.kaggle.com/moltean/fruits/version/2 [Accessed 6 Apr. 2019].

Vanderhorst, D. (2018). CBP intercepts prohibited Mexican fruit [online] Available at: https://www.thepacker.com/article/cbp-intercepts-prohibited-mexican-fruit [Accessed 6 Apr. 2019].