Abstract- The purpose of this project was to accurately classify images using NVIDIA's DIGITS workflow. In this project, 4 classes were chosen- 3 different pills and no pills. The goal was to build the dataset and then train the model. Using the model, the goal was to accurately classify a new example of one of those images. And the final step was to deploy the model to the Jetson TX2.

I. INTRODUCTION

This project could be adapted to sort pills into appropriate containers coming down a conveyor belt for packaging after production to avoid any sort of manual sorting. The robotic inference idea is to build a classification model that can accurately classify pills. That model could be downloaded to the Jetson TX2 and code controlling servo motion could direct the pill to the appropriate container depending on the identified class.

II. BACKGROUND

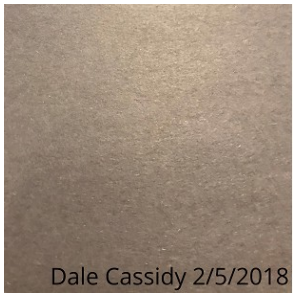
For this project, the Google LeNet model was used. Since the Jetson TX2 would be placed on top of a conveyor belt with the camera looking down at the pills it could easily be powered directly from a wall outlet. Since a battery is not needed, power consumption is not an issue. If power consumption is not an issue it makes sense to choose a model with high accuracy and high inference speed since inference speed is directly related to power consumption. Of NVIDIA's model choices for classification, Google LeNet is the best option for this. Google LeNet is a 22 layers deep convolution neural network with just 4 million parameters. It has a high accuracy and high inference speed.

III. DATA ACQUISITION

3 pills were chosen- advil, tylenol and a vitamin. All 3 items were placed against a black sheet of paper. 400 images of each pill and 400 images of the plain black background were taken using an iPhone 7 camera. That's 1600 images in total. An example of each class is shown below:

Advil

Tylenol



None

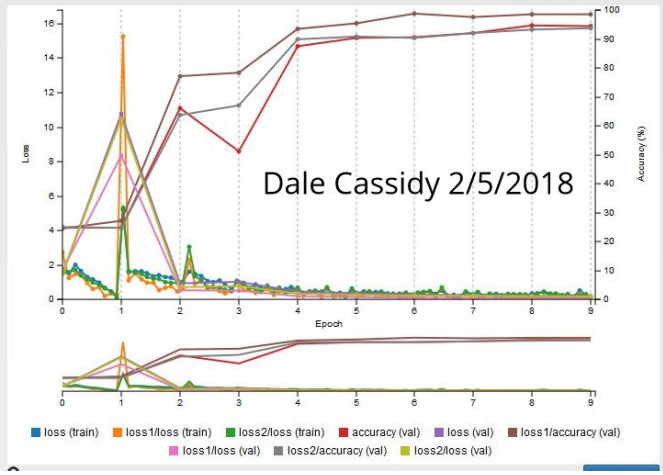
Vitamin

The images were imported as 256x256 RGB images and were squashed.

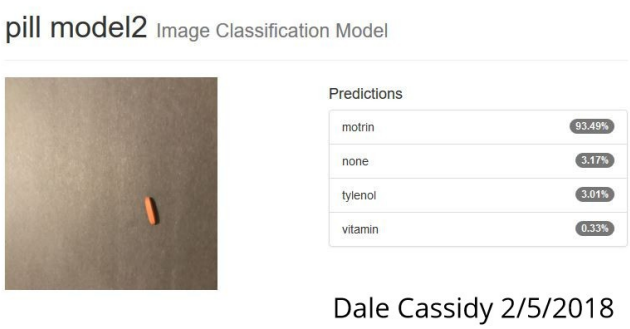
IV. RESULTS

The Google LeNet's model was run on the dataset for 9 epochs. It's accuracy was 94.47% and the inference speed was between 5 and 6 ms.
The following chart shows the loss vs. epochs for

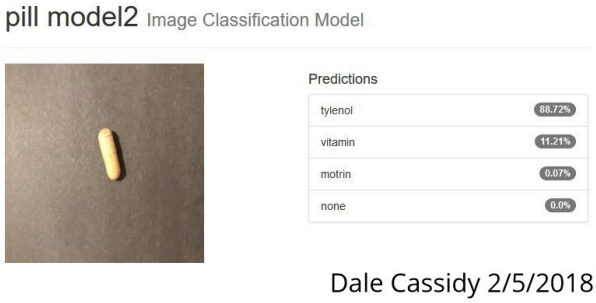
the model run:



The following was a classification sample on a new motrinpill image. It correctly predicts it as an motrinwith probability 93.49%.



And the following is a classification sample on a new vitamin image. Notice that it incorrectly labels this pill as a tylenol.



V. DISCUSSION

Overall the results are good as the accuracy is 94.47%. Taking a look at some of the single classifications on new images, the classes of none and motrin are very accurately predicted in the 90% range. The motrincolor is very unique making it easy to classify and the black background is uniquely uniform also making it easy to classify. However, one may be surprised that the vitamin pill may be incorrectly identified as the tylenol pill. The pills are close in color but not exactly and their shapes are relatively similar though the vitamin is larger. It may simply mean that a larger dataset is needed to more correctly determine the class between the vitamin and tylenol. Inference is most important in this project because for a pill sorter you need to get the correct pill in the right container.

As part of this project, the model was deployed to the Jetson TX2 and inference was run at the edge on new samples in real time. However, the results for this did not turn out well as the class was mostly tylenol even though a vitamin or even a motrin was being tested. I suspect this mostly has to do with the camera images from the Jetson TX2. The quality looked very fuzzy as compared to the iPhone 7's camera which is where the dataset was from.

VI. FUTURE WORK

This project achieved what was attempted which was to accurately and quickly identify most pills. It is not commercially viable though because one would need an accuracy of 100%.

Given more time, one could figure out what is needed to classify a tylenol pill versus a vitamin pill better because that lesson could be used to differentiate other pills as well.

Also, with more time, one could take pictures with the Jetson TX2 and use that to train a model on before deploying the model and testing the results on realtime data.