Scientific Report CSE 455/555 Term Project

Jacob Szczudlik

7/27/25

Part 1: CV Engineer, Jacob Szczudlik. ResNet50 Pre trained model

Data Analysis: The goal of this part was to use a state of the art pre-trained model to accomplish 2 objectives. 1.) Fine tune a pre trained model to be able to distinguish between, in our case, 4 different classes of fruit, avocado, banana, pineapple and carrot. 2.) Use a second model to classify the images based on variations. For example can our model distinguish between a whole avocado and one that is diced. To accomplish this I chose to use a pretrained Resnet-50 CNN with a set of hyper parameters that I'll describe in the next section. With this model we can leverage transfer learning as the REsNet50 model already has been trained on generic features like shapes and textures which is perfect for this type of classification task. We just need to handle the fine tuning for our 4 unique fruit classes and 12 variation classes.

Model set up: As mentioned for the image classification task I chose to use a ResNet50 pretrained model. For task 1 this needed to be fine tuned on 4 input classes, and for the variation classification task it needed 12 classes. Well use a batch size of 32 because this seems like a good middle ground between efficiency and stability. I did try a larger batch size but it was too slow and memory hungry. I chose to use cross entropy for the loss function as this seemed the most consistent for multiclass classification tasks. Adam optimizer was applied because it is shown to be effective with pre trained models and that's with a standard learning rate of 0.001. I tried a few runs of different epochs and it seemed that less than 2 wasn't sufficient and more than 5 wasn't improving accuracy enough so I settled in on 5 epochs. After 1 epoch accuracy was around 80%, after 2 around 90%, and after the full 5 settled in around

98%. This was a very pleasing result. Below I've attached the confusion matrices and precision graphs for the data showing a high degree of accuracy for most classifications.

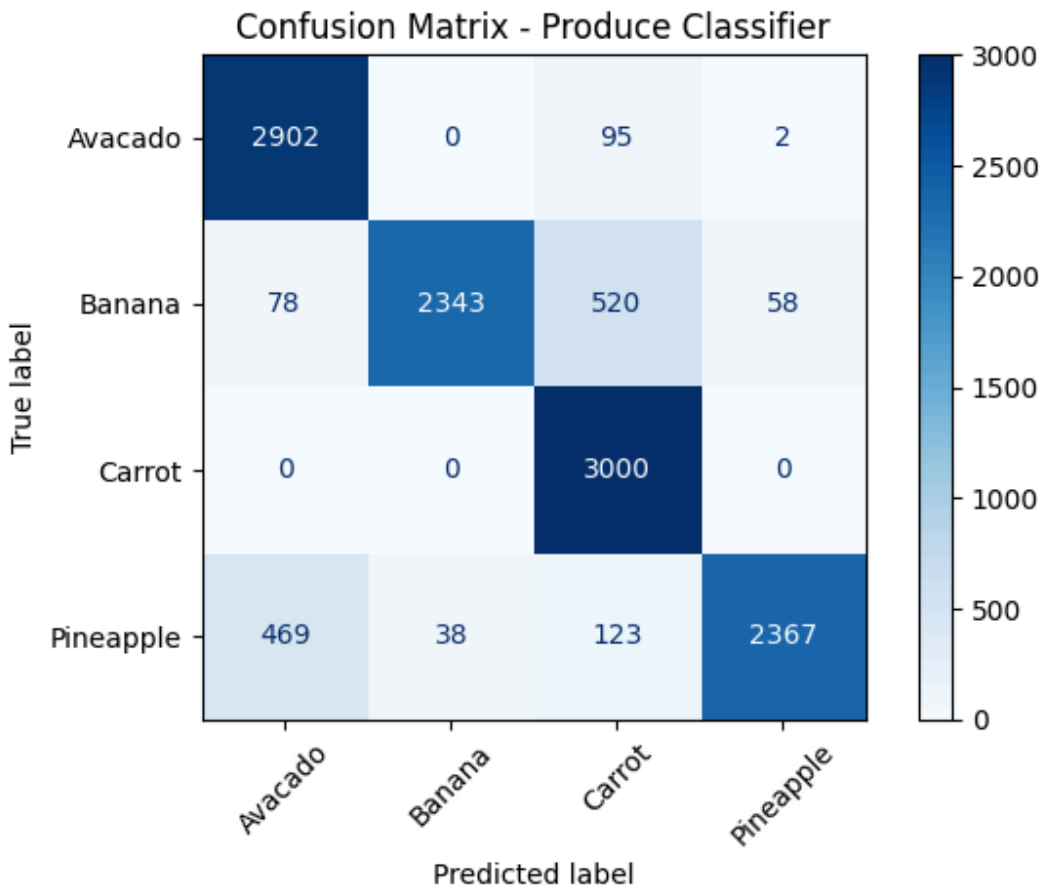


Figure 1: Fruit classification confusion matrix

As you can see by the above chart, there is a high degree of accuracy across the diagonal for each class. Most impressive being the carrot with no incorrect predictions. I imagine it is due to multiple factors including image quality and variation, but it's likely largely due to its distinction from the other 3 fruits, which are much more similar to each other. You will see later on in the performance graph that there were a few interesting insights when it came to the classification

of carrots and carrot variations. Another interesting observation was the banana was mis classified as carrot a large number of times but the same was not true for carrot to banana.

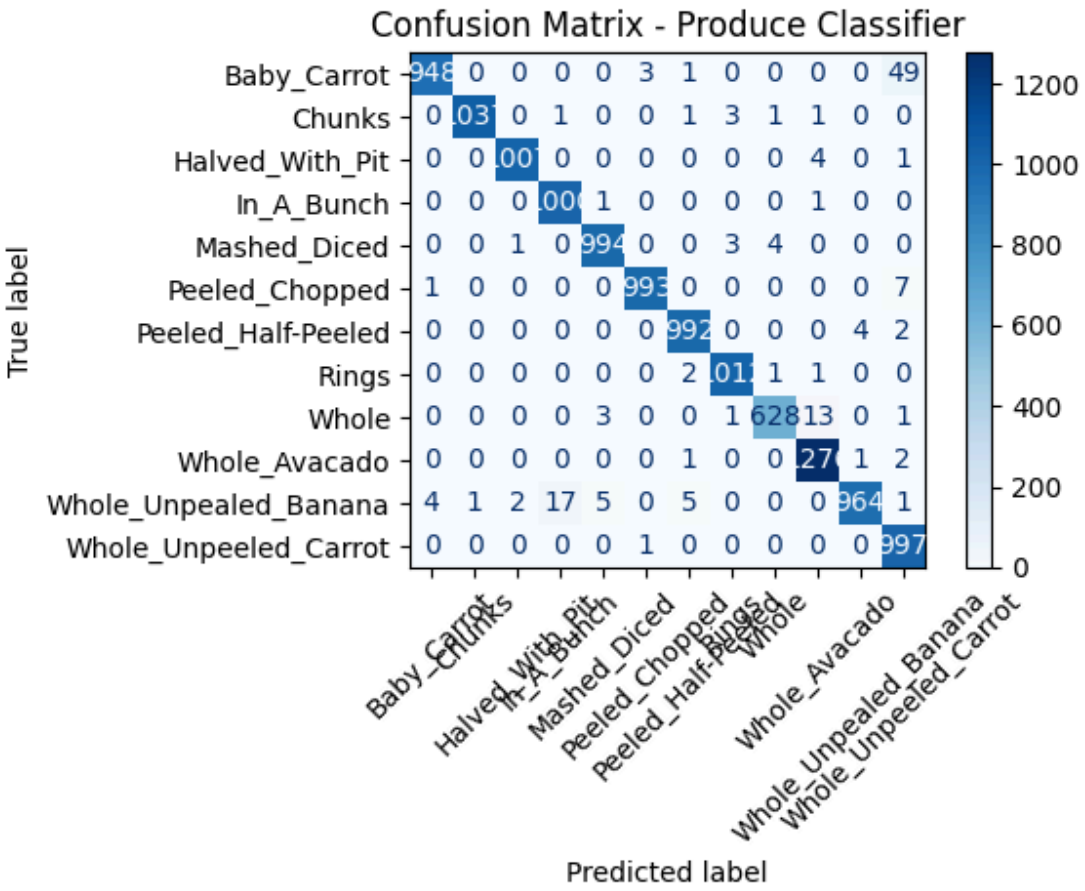
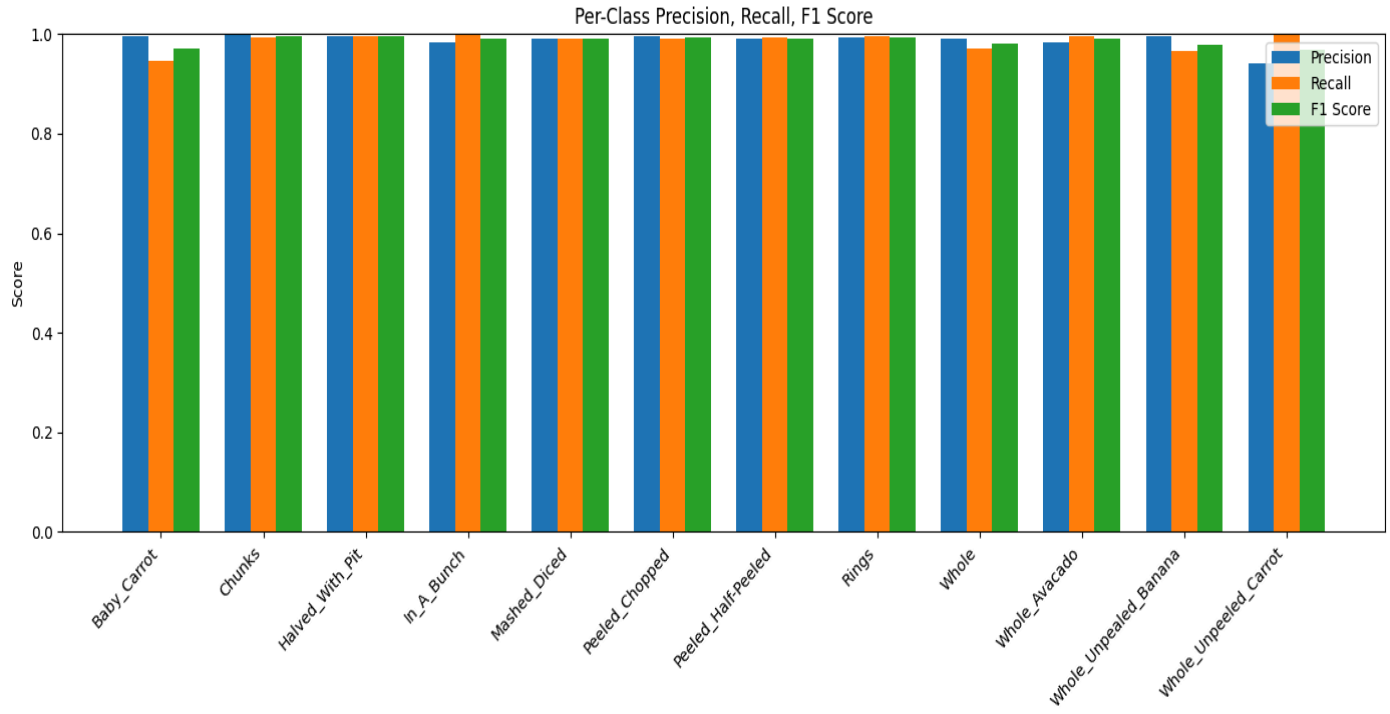


Figure 2: Fruit variation confusion matrix

The data for the variations is impressive. Aside from mistaking variations of the same fruit a few times, the present model performed very well at distinguishing each variation from other classes. The added context and additional classes seems to have allowed the model to do even better fine tuning and learn the difference between the fruits. Although the simple class identifier worked well on the 4 simple classes, the more refined model worked even better. Again here we see an interesting speed bump where the model confuses variations of carrot more than with other classes. This seems to be a trend.

Figure 3: Variation Precision Graph



As you can see in the above graph, precision is excellent for all classes. The most noticeable lag in precision is with the Whole Unpeeled Carrot. All other classes are near perfect Precision. Similarly the Recall for baby carrots is slightly behind the other classes. It seems our model may have some difficulty distinguishing between these 2 variations of carrot. I suspect that the distinct color and shape of carrots make them easily distinguishable from other classes, as well as from their surrounding environment, but at the same time difficult to distinguish from other

variations. This claim is further supported by the classification section for the 4 items where the accuracy for carrot was perfect, and no other class could say that.

Conclusion:

By leveraging transfer learning from a pre-trained model, and using a balanced batch number and number of epochs, we were able to train a very accurate classifier in a small amount of time with low memory usage. Our final fully connected layer was modified for 4 output classes, corresponding to each fruit option, and again for 12 output classes for the variation classification section. Using Adam and cross entropy loss we constructed an all around balanced image classifier model. We saw very good performance metrics as shown in the figures and observed some interesting phenomena. It seems our model performs very well for general classification tasks but struggles a bit distinguishing between variations of the same class in some cases.