Fruits image processing

D6: Markus Roletsky, Rene Puusepp, Alina Gudkova

1. Introduction

In this project, we were working with a Kaggle dataset containing 94 110 images of 141 different fruits, vegetables, and nuts. Our aim was to create a model that can accurately classify these items using image processing techniques. We built a CNN model and enhanced the dataset by augmenting it with modified image backgrounds. Finally, we tested the model's generalization capability using real-life images.

2. Goals

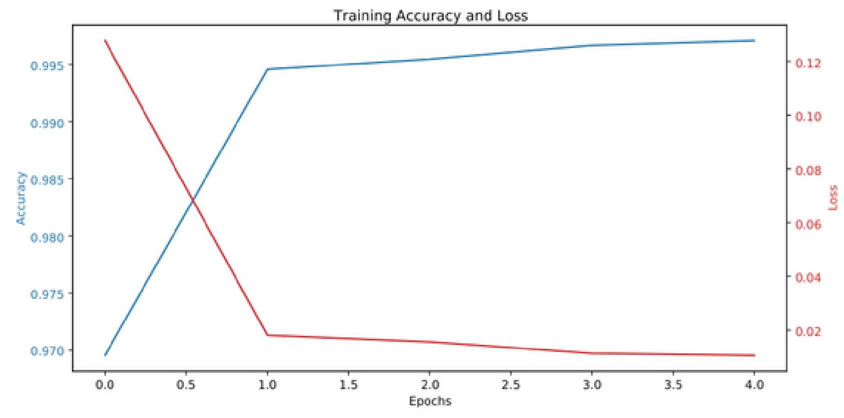
- Train a CNN model capable of classifying fruits, vegetables, and nuts with at least 80% accuracy.
- Enhance dataset diversity by editing backgrounds to improve model generalization.
- Test model generalization with real-life images of fruits.

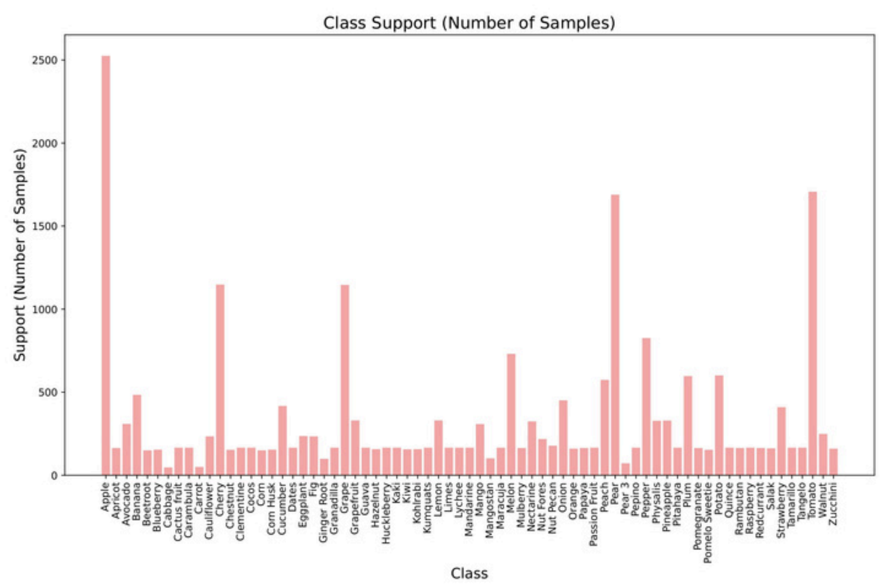
3. Methodology

We began with exploring the dataset, then split it into training, validation and test sets. An initial CNN model was developed and trained using Keras and its performance was fine-tuned. To improve diversity, a subset of images will undergo background editing and the model will be retrained using the new dataset. Finally, custom test data(photos with background) was collected from the internet, processed and used to evaluate our model's performance on real-world scenarios.

4. Model design, process & results

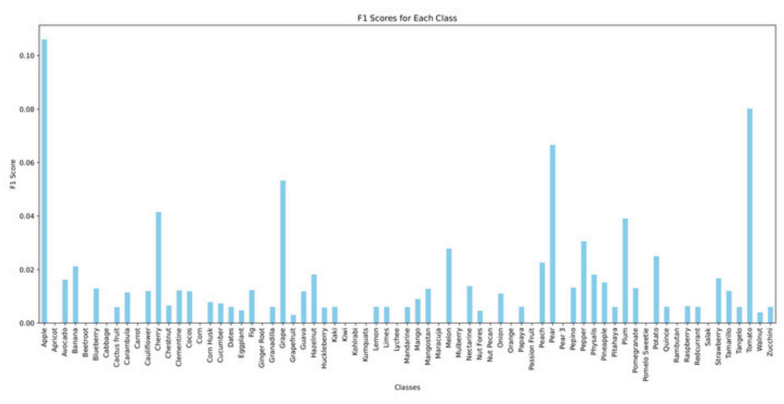
The model was made using MobileNetV2, a pre-trained network, and fine-tunes it by updating only the last 50 layers. It then adds layers to process the data: one to reduce the size of the data and another with 128 units to learn patterns, followed by a final layer to predict the classes. Our model is trained with a small learning rate and uses a loss function to measure prediction errors and accuracy to track performance. It trained for 7 epochs, adjusting the model based on the data. The model appeared to be learning effectively and converging as desired.

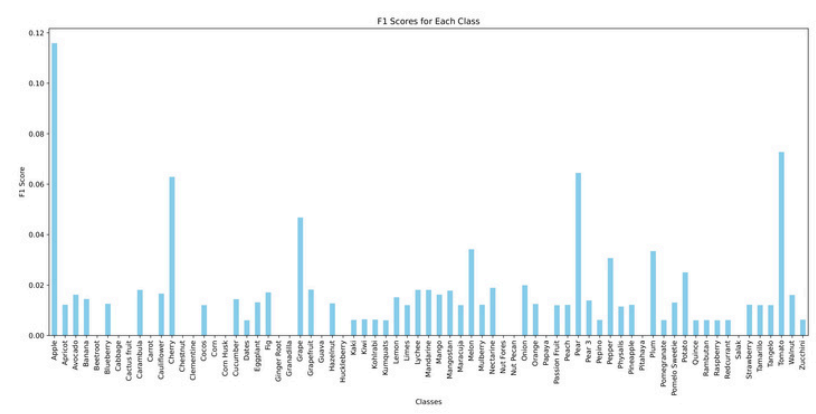




There was significant variation in sample counts across classes. Some fruits had high sample counts, while others had considerably fewer. Many classes had far fewer samples, leading to poor model performance on these classes. The imbalanced dataset resulted in biased predictions favoring classes with more samples.

F1 Score





Baseline model's F1 score (left) and improved F1 score of a fine-tuned model with original testset.

Analysis & conclusion

While showing high accuracy the model indeed did a bad job at classifying real photos. At first we reduced the number of classes by generalizing fruit types(from 128 to 70), this changed the amount of sample for some fruits, resulting in a better accuracy for some classes compared to others (see support and F1 scores). We then tried to fine-tune the model by freezing layers up to last 50 and with tuning we saw this gave us a slight overall increase in accuracy. The accuracy with the test set belonging to dataset was over 99% yet when we produced our own set of photo images (360 pcs) then the accuracy dropped to 24.9% . We conclude that both freezing layers and consolidating classes helped to boost accuracy. Fine tuning enhanced f1 scores roughly 8% and consolidating classes helped to increase f1 by 22%. The f1 score is low throughout the project, indicating that the classifier is not really doing a good job, which is natural, as an example most humans cannot distinguish between 5 citrus fruits with only visuals and no size metrics (mandarines, tangerines, clementines, oranges, grapefruits, tangelos) so any model with only 100x100p visuals cannot be expected to give good results.

Background editing for training data was not implemented because of time constraints. We deem it highly unprobable to raise f1 scores, because models inaccuracy is largely tied to the lack of information in dataset.