Fruit Classification

Contributors: Ava Borak, Aimee Lee, Billy Wei, Evelyn Seligman

Contact: aborak@wisc.edu lee2258@wisc.edu bwei32@wisc.edu oseligman@wisc.edu

1. Project Overview

The goal of this project is to classify the fruits and count the number of fruits from the image. The classification algorithm will categorize 10 kinds of fruits and count the quantity of the elements. The ten different classifications of fruits will be apple, banana, strawberry, kiwi, peach, plum, cherry, pear, mango, and lime.

2. Background

During our exploration of projects, the most notable type of project classified classes of fruits based on their images. Most of these projects were able to do so with 80-90% accuracy, and we aim to improve this anyway. Specifically, in a notebook [2] aimed to classify fruits using an inceptionV3 model, the validation accuracy converged to .9, or about 90%.

Additionally, our group wanted to take simple classification one step further and use the data set 'Fruit Classification' and 'Fruits-360 dataset, which only contain one fruit per image, to train and identify how many fruits of each class were in images from the 'Fruits 262 dataset'

3. Statement of Work

3-1. Datasets

We chose multiple datasets, the first of which contains over 225,000 images of fruits belonging to 262 classes, while we are not going to be using all 262 classes. The images themselves in this data set have different backgrounds and varying numbers of fruits in them. The author of this data set also notes that there was a 'resize.py' script that standardized the size of each image.

The next data set, 'Fruit Classification,' contains a total number of 22,495 images split into test and training sets, and these images have more simple, plain white backgrounds. This dataset has 33 classes, but once again we're only going to be using the ten different fruit classes listed in the overview of the project.

The last, 'Fruits-360 dataset,' is similar to the second, as each image only contains one fruit and the background is sterilized and uniform. This final dataset includes 101,263 images of 153 varying fruit, vegetable, and nut and seed varieties.

Dataset 1: https://www.kaggle.com/datasets/aelchimminut/fruits262

Dataset 2: https://www.kaggle.com/datasets/sshikamaru/fruit-recognition

Dataset 3: https://www.kaggle.com/datasets/moltean/fruits

3-2. Methods

For this model, we plan to construct and train CNN models with Dataset 2 and 3, as they have more neutral backgrounds. Within Dataset 1, we aim the model to classify the object from the natural photos and provide the quantity of the fruits. Additionally, we will introduce negative examples, such as non-fruit images, to help the ability to distinguish between fruit and non-fruit categories.

We will use Mask R-CNN, a deep learning model designed for instance segmentation. Unlike traditional object detection models, Mask R-CNN not only detects objects but also generates pixel-wise masks for each object in an image. This capability is crucial for accuracy in counting fruits in real-world images where multiple objects may overlap. We will manually annotate images in Dataset 1 to generate segmentation masks and train the Mask R-CNN model to recognize and count individual fruits. To enable object detection and fruit counting, we will develop a CNN model trained on images where fruits appear in natural environments.We will manually annotate images in Dataset 1 to generate segmentation masks and train the Mask R-CNN model to recognize and count individual fruits. To further refine detection accuracy, we will fine-tune the model using real-world data variations and conduct extensive testing to ensure robustness.

3-3. Outcome and Performance Evaluation

For the outcome, we plan to hit average 85% accuracy for two neural network models (CNN and ANN). Based on the projects we have found, most of the accuracies are between 40% ~ 80% [1]. To assess the performance of both models, we will compare their classification and detection capabilities. The evaluation metrics will include accuracy and precision while the mask R-CNN model will be assessed using mean average precision (mAP) for object detection and instance segmentation accuracy for counting. Additionally, we will track model training time and computational efficiency to ensure the approach remains scalable. By analyzing the results, we will determine the strengths and limitations of each model and evaluate whether combining classification and object detection improves performance for real-world applications.

4. Project Plan

Task \\ Week #	1 March 8 - March 13	2 March 14 - March 20	3 March 21 - April 2	4 April 3 - April 9	5 April 10 - April 16	6 April 17- April 23	7 April 24 - May 1
Environment Set-up	X						
Data Pre-processing		X					
Build/Train Model			X				
Test/Improve Model				X	X	X	
Evaluate Results				X	X	X	
Visualize Results				X	X	X	
Progress Report			X				
Final Report and Presentation						X	X

5. References

[1] https://www.kaggle.com/code/omaremad5/object-detection-model

[2] https://www.kaggle.com/code/marianadeem755/fruits-classification-deep-dive-into-inceptionv3/notebook