Produce Freshness Classification

Introduction

In the United States, food waste has been a major issue from not only a supply and demand perspective, but also a monetary economic perspective. The Department of Agriculture estimated that food waste takes up 30 to 40 percent of the food supply. Combining this with estimates from USDA's Economic Research Service in 2010, this translates to about $161 billion worth of food. This means that food that could've been on many families' dinner tables is now sent to landfills, in turn wasting all resources that were used to produce the food.

Motivation & Novelty

Food waste and loss can occur at any stage of the production and supply chain, such as the stage between the farm gate and retail entities. At the retail level, equipment malfunction and inventory mismanagement can lead to food waste. From a more individual standpoint, consumers also participate in food loss and waste. Whenever consumers unconsciously buy relatively stale produce, they might not miss the appropriate date to consume the produce before they turn bad entirely, a situation that could be prevented if consumers knew how to pick and choose the produce.

Imagine that you are a customer shopping for produce, however, you do not purchase produce frequently and do not have sufficient knowledge to differentiate a fresh produce from a stale one. In this case, a trained model that can help distinguish this freshness would be of great assistance for this scenario. Customers can make better shopping decisions and reduce the risk of food waste from the refrigerator. And in a larger picture, help reduce food waste and loss with more and more customers doing the same.

Objective

We hope to predict whether a fruit or vegetable is fresh given the picture of it. The predictor would be sets of images for different kinds of fruits and vegetables and the labels would be binary. The labels can be understood as "fresh" or "stale". Specifically, for images with more than one fruit or vegetable, our ideal solution would be to first recognize how many items are included, then separating them and detecting the freshness one by one, and finally producing several classification labels for the given image.

Data preparation

As for the dataset, we plan to use images from google which cover around 10 types of vegetables and fruits. They can be separated into a fresh level and a non-fresh level. Below are two examples of two tomato pictures and two banana pictures. On the left it is labeled as "Fresh", and on the right it is labeled as "Stale". Our estimated sample size would be 160 pictures overall which includes training, validating, and testing. Our goal for the neural network model would be to identify each individual fruit or vegetable in a picture and determine their freshness.

Sample picture

Fresh: 

Stale:

Fresh: 

Stale: 

Reference

[1] Food Waste FAQs | USDA
[2]"The Estimated Amount, Value, and Calories of Postharvest Food Losses at the Retail and Consumer Levels in the United States"