# Classification of Bananas by Ripeness Using Various Methods

## Introduction

Several of the algorithms in this class can be applied to image processing tasks, and one such task is determination of fruit ripeness. In this report, I examine the results of using Self-Ordered Mapping, K-Means Clustering, and a perceptron network to the challenge of classifying bananas as under ripe, ripe, or overripe. These methods, if refines, could be employed in automated fruit packing plants and in other parts of the food industry to automatically sort fruit.

## Image Conditioning

Since isolating the banana(s) in the input image is beyond the scope of the project, all of the images were digitally modified to remove everything but the banana(s), replacing those pixels with a white background. In a food inspection environment with a known image background, this could be easily accomplished with trivial image processing algorithms.

Given a set of banana images, several thousand pixels were sampled from each and used as inputs to the various algorithms. Only pixels that were part of the banana were sampled (the white backgrounds were ignored). Random sampling was used to eliminate problems from different image resolutions. In an industrial environment with known camera characteristics, it would be possible to sample a much larger group of pixels, but this would of course slow the network down. In order to increase the size of the training set, images were sampled multiple times, which was possible because the images were much larger than the sample size.

## Self-Ordered Mapping of Random Pixel Samples

My initial attempts to use self-ordered mapping were unsuccessful. Originally, I arranged the red, green, and blue channel pixel values of the sampled pixels into one long vector and fed these vectors into the algorithm. Upon running the algorithm, I found that the entire validation set was classified into one or at most two clusters.

Originally, I thought this was because I was classifying the images using the three color channel values of the pixel vector. Since color is the most important factor in classifying the ripeness of bananas, I decided to use a vector of pixel hue values as the input to the algorithm. However, this also failed, and for the same reason.

Since the pixels were from random locations, there was not necessarily any correlation between the same pixels in each input vector. In order for self-ordered mapping (or k-means clustering) to work in this application, I would need some way to remove pixel location from the system.

## Self-Ordered Mapping of Pixel Bin Counts

## K-Means Clustering of Pixel Bin Counts

## Perceptron Classification of Random Pixel Samples

## Conclusions