

Capstone Problem Statement

Quanah Green

October 2017

Abstract

Analyzing grass seed samples to determine the presence of undesired seeds is currently a time consuming and costly process. In order to reduce the overhead of this process, we are proposing a solution that would allow pictures of seed samples to be taken and sent via the internet to a system that would analyze the seeds and generate a report detailing what types of seed are present in the sample.

1 Problem Statement

The Oregon State University seed lab currently employs ten people to work full time analyzing grass seed samples to ensure that all of the seeds are of the same type. This work is incredibly tedious and time consuming, consisting of spending hours on end staring into a microscope looking for small differences between seeds as they scroll past on a conveyor belt. Because the job of analyzing these samples is so tedious there is a high employee burnout rate.

Furthermore, to actually have a collection of seeds analyzed in this way requires that they be physically shipped to a lab like the OSU seed lab, which is costly and time consuming. If it were possible to automatically and remotely perform seed analysis it would make the process far cheaper and more convenient. To this end, we are trying to develop a solution whereby a picture could be taken of a spread of seeds, and that picture could then be uploaded to a system that would automatically determine if all of the seeds in the picture are of the same type or not.

2 Proposed Solution

In order to achieve automatic seed identification, we would like to use a machine learning algorithm that can be trained to identify and distinguish between different types of grass seeds. Once we have completed sufficient training to identify a seed with a high degree of accuracy from a photo, we will develop a server-side application that accepts uploaded pictures, isolates images of individual seeds within those pictures, and then determines the type of each seed in the picture. It will then generate a report indicating what seeds are present in the picture.

3 Performance Metric

The metric for success is the ability to process an image with a large number of seeds in it (500-2500, depending on resolution limitations), and determine the type of each seed with reasonable confidence.