Segmentation of Crop Damage Project Proposal

1st Arpad Voros
aavoros@ncsu.edu

2nd Josh Beerel
jbeerel@ncsu.edu

3rd Zackery Miller
zamiller@ncsu.edu

***Index Terms*—crop damage, semantic segmentation, image recognition, down/up-sampling CNN, RNN**

I. Introduction

This proposal discusses the approach in designing a deep learning model to predict the health status of crop fields via aerial photographs. Specifically, the proposal will include a methodology for design, evaluation plan, and an outline for baseline model comparison.

II. Methodology

Our methodology consists of developing a proficient method to process the input data, defining a solid baseline model to compare against, and configuring the model with optimal parameters and hidden layers to perform accurate predictions on the crop image data set.

A. Data Processing

The input data provided for this project consists of RGB images of size 432×576 pixels. These images are arranged into five categories of relative maturity (RM); RM-2, RM-1, RM, RM+1, and RM+2 to represent the number of weeks pre/post RM. Each image constitutes of objects of various classes (e.g. healthy crops, damaged crops, shrubbery, roads, cars, and grass). Alongside the input training data comes annotated masks, which outline where the crop damage occurs in each respective photograph. This binary mask includes sections of crop damage and excludes objects of all other classes.

To decrease the model size and run-time, the input images along with their respective annotations can undergo a series of crops, resizes, and rotations prior to being fed into the model. This process has been confirmed to result in higher accuracy and confidence compared to using full images for the training set [1]. The input shape will be (batch_size, img_height, img_width, img_depth). In addition, the predictions must undergo the inverse of said operations and be pasted together in order to develop a fully-fledged mask of each test image. If we choose to feed our model with the current size of the images, our input shape will be (batch_size, 432, 567, 3), and there will be no post-processing for our prediction set.

A validation set will be produced to consist of a 20% random sample of the training data.

B. Model I

Convolution neural networks are commonly used for image recognition problems due to their ability to preserve spatial locality while extracting features of the image. We will train many variations of the CNN models while taking into account a variety of filters, max pooling filters, convolution layers, fully connected layers, and stride sizes. Given that the input data is formatted into RGB colors, there will be 3 channels for the kernel.

The structure of this model will be based on semantic segmentation, Fig. 1, that will attempt to identify the ”damaged” or ”un-damaged” components of the image. This may require the use of multiple CNN and Pooling layers for both downsampling and upsampling processes that will retain the shape of the input but demonstrate the output segments correctly [2]. Parameters such as dropout rates, epochs, and activation functions will be used in a grid search to find the optimal values for these parameters.

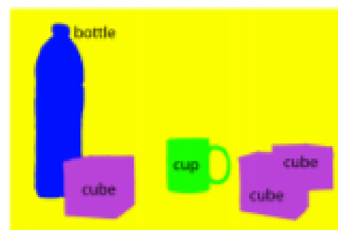

Fig. 1. An example of semantic segmentation [2]

III. Evaluation

A. Metrics

We will generate a processed image to verify the location of the annotation to determine the ”closeness” of the annotation to where its real position should be based on our data, which will fulfill the Jaccard Index reqiurement. Other metrics that will be assessed in the model construction will be the run time and overall computational complexity of the models trained, which will enable us to determine the most efficient and accurate model.

B. Baseline of Comparison

We will conduct ground-truth comparisons, where we view the accuracy of the model in determining the damaged areas and compare that to what we know the output to be.

References

[1] P. Sharma, Y. P. S. Berwal, and W. Ghai, “Performance analysis of deep learning cnn models for disease detection in plants using image segmentation,” *Information Processing in Agriculture*, 2019.

[2] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, and J. Garcia-Rodriguez, “A review on deep learning techniques applied to semantic segmentation,” 2017.