Semantic Segmentation in Agriculture

{Raunaq Rewari} Stanford University

Overview

At Blue River Technology, we are pushing the state of the art in precision agriculture. Our tractors are fitted with cameras that can detect weeds and crops and preferentially spray herbicide only on the weeds, leading to huge savings in cost for the farmer. Naturally, the higher the resolution of detection, the more precise the spraying can be. After demonstrating success using bounding box approaches in cotton farms, we ran into roadblocks in applying the same technology to other plants like Soybean, which is why we turned to semantic segmentation to solve the problem. This project tries to segment weed and crop from the background in case of Soybeans.
The best performing model did pretty well in segmenting plants from background with a precision of 64% and recall of 99%, but couldn't really do a very good job at segmenting weeds from background.

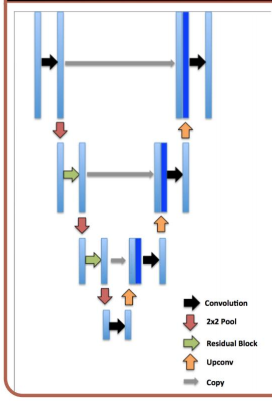
Data and Features

The dataset for this project came from Blue River Technology. The data was collected by imaging Soybeans fields in the US. Since accurate data labeling is crucial for semantic segmentation problems, we used Figure8's services to allow our in-house agronomists to painstakingly label each regions of images as crop or weed.

We collected 7379 images in total, 6700 of which were used for training, 379 for validation and 300 for testing. Since the images are pretty homogeneous in general, we randomly shuffled the dataset and split it up into train/val/test.

The images are 384X384 RGB images, with corresponding labels being single channeled images of the same size with each pixel labeled as one of three classes (0=background; 1=weed, 2=crop). It should be noted that the dataset is highly imbalanced with majority of the pixels being background and a very tiny fraction being weeds.

Models



The architecture used is an adaptation of the U-net architecture, which is an encoder-decoder architecture as shown in the figure. The original U-net architecture is modified to have residual layers instead of vanilla convolution layers, which allows the use of deeper architectures without the problem of vanishing gradients.

To tackle the high imbalance in the dataset, a variety of different loss functions were used.

- Cross-entropy
- Weighted cross-entropy
- Wasserstein
- Weighted Multi-class Dice Loss **

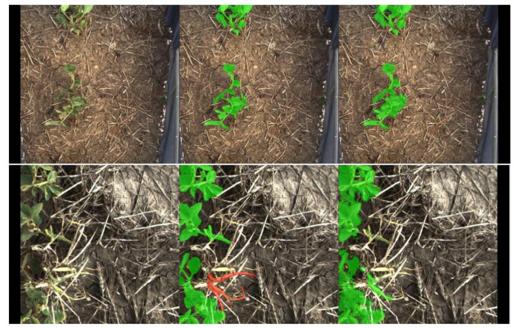
** Multi-class Dice Loss was calculated by summing the soft binary dice loss for each label. The inverse frequency of labels was used as weights.

Results

Here the mean intersection over union over all 3 classes for the test data is reported. Most papers on semantic segmentation report this metric.

Loss-type	Test mIOU
Weighted cross-entropy	0.13
Wasserstein	0.13
Weighted Multi-class Dice Loss	0.35

Discussion



The two sets of images shown here are 2 examples from the dataset. In each panel of images, the leftmost image is the input image, the center image is the the input image with the ground truth overlaid on top of it (green is crop, red is weed) and the rightmost image is the predictions outputted by the model.
As can be seen from the first panel, the model does really well in segmenting crop, since it is relatively well represented in the dataset. However, looking at the second panel of images, it can be observed that the model doesn't do as good a job on weeds. This causes the mIOU of the models to be low.
There could be a few reasons for why the model is not doing well on weeds:

- There are a very low number of images with weeds in them. For the images that do have weeds, they occupy a very small part of the image.
- The size of the images (384x384) is too small.
- The weeds look a lot like the background in terms of shape in some cases (Panel 2) or look similar to crops in color preventing the model from learning.

Future Work

- Try other models like Segnet which do not rely on learning the upsampling weights
- Try augmenting the dataset by training a GAN to generate more weeds
- Try running the model on higher resolution images
- Prioritize image collection from high weed density fields