Individual Lab Report

Name:	Dung-Han Lee
Team Name:	Team E Wholesome Robotics
Team members:	Aaditya, Aman, John, Hillel
ILR Number:	03

March 7, 2019

1 individual progress overview

There were two major focus area in the past 2 weeks, which is (1) automate the quantitative analysis process and (2) increase number of labeled data and see if that raise precision and recall performance.

1.1 Automate quantitative analysis

1.2 Motivation

The motivation behind this task is to have a common ground to compare different training results. Doing this manually could take hours so there is a good reason to automate this process

1.3 Approach

1.3.1 Build test data-set

First of all, a golden validation test data-set need to be generated in order to serve as a common ground. To this end, 28 images covering roughly 400 labels are created to represent 8 different types of plants(Fig1). The data is chosen in a way to cover different scenarios i.e. easy case, hard case, close look, far look etc. After that, the bit-mask of these images also need to be labeled manually.

1.3.2 Use IOU as metric

The output of mask-rcnn network consists of bit-masks indicating the physical positions of interested objects. In our case, it's disease and holes. With this information and the validation data mentioned earlier, bit-masks from neural network output is compared to the bit-mask of ground truth(Fig2). For any predicted $mask_{net}$, it's compared with all other $masks_{truth}$(Fig3). If the intersection over the union of two masks is higher than a given threshold, the $mask_{net}$ would be considered as a true positive. If a $mask_{net}$ failed to match with any $masks_{truth}$, it would be considered as false positive. Conversely, a $masks_{truth}$ failing to match with $mask_{net}$ would be considered as false negative. To prevent counting a mask twice and thus inflating the true positives, a mask would be deducted after it is matched.

1.3.3 Discussion of results

It turns out, though, this actually leads to a more false negative. Because diseases tend to appear in fragmented patches. Say there is N predicted masks corresponding to 1 $masks_{truth}$, this would lead to N-1 false positive in the result. At this moment this issue is considered less urgent and left as it is.

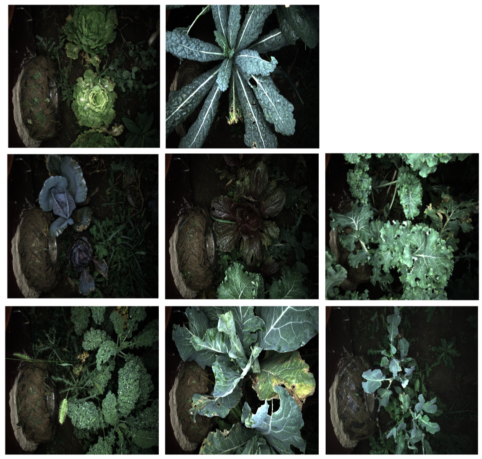


Figure 1: 8 types of crops on the field

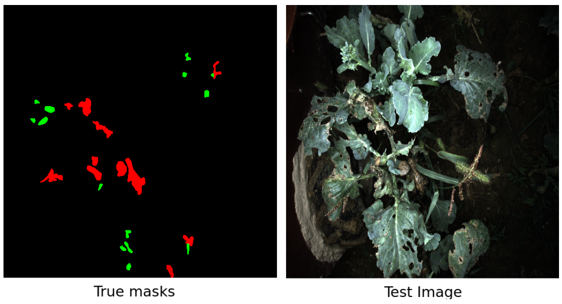


Figure 2: test image and true mask

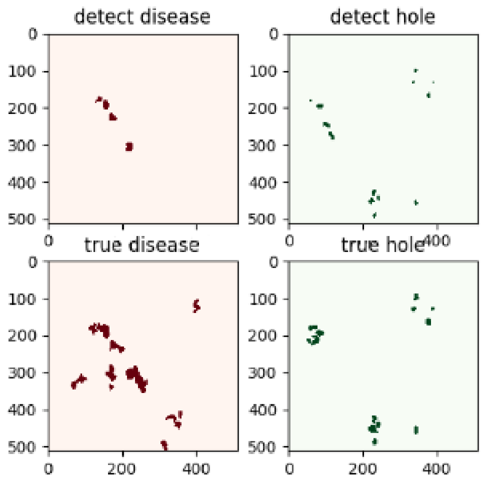


Figure 3: neural net output versus ground truth

2 Increasing training data for a given plant

2.1 Motivation

Since many false negatives were observed last time, it's considered that there's not enough training data for the neural network to learn features from the given data-set.

2.2 Increase data for 1 type of plant

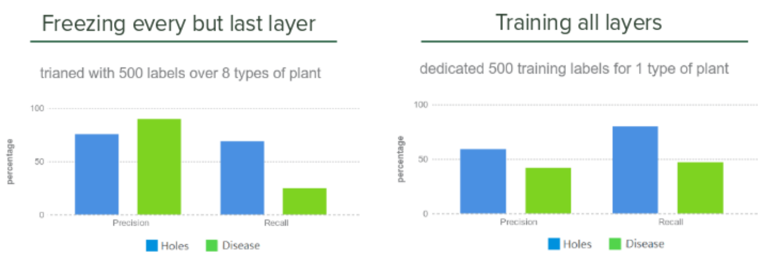
To evaluate whether increasing data would improve performance, 1 specific type plant is chosen and 500 labels are generated for that type. (In the previous set-up, there were 500 labels spread over 8 types). After that, the labels are used to do training over all layers of resnet50 neural network structure. (Since there were much less available data last time, all layer except for the last one was frozen in previous training set-up)

2.3 Precision worsen while recall improves

The result is surprising, that the increased number of labeled data does not improve both precision and recall, but only the latter (Fig4). i.e. the network manages to recognize more potential holes and disease, but it also reports more false positives. At this moment, there is a big gap between the desired 80% recall ()and 80% precision. The current status: 59% precision and 80% recall for hole detection, 42% precision and 47% recall for disease detection.

3 Challenges

The neural network does not seem to generalized well for holes and diseases. The original approach was to use the pre-trained network from imagenet as a starting point and fine-tune the neural network. However, given that this project is meant for a very specific application, previous features learned from 1000 object classification may not provide many advantages to the task.



$$\text{Precision} = \frac{tp}{tp + fp} \qquad \text{Recall} = \frac{tp}{tp + fn}$$

Figure 4: recall improve but precision worsen with more training data

4 Plans

Some further investigation such as changing neural architecture or switching to traditional computer vision approach need to be studied in the following weeks. The plant is to look into research papers in the relevant field and talked to master students who have worked on similar projects. One potential option is to do a supported vector machine and use a neural network as a feature extractor instead of a classifier.

5 Teamwork

In this week, I have talked to John and decided to use IOU for automating the quantifying process for the monitoring task. Also, I've documented LiDAR set-up procedures for Aaditiya so that he would be able to interface with LiDAR smoothly. Hillel has been working on cultivator design and setting up motors for the platform. Aman has been working with John on deciding the locations for mounting sensors. John has set up NUC, interface motor with ROS and work on particle filters. Aaditiya has been working with John to set up Zotac for onboard computation. Meanwhile, he tested and run the LOAM SLAM package.