CSE353 - HW1
Yajie Wang

Introduction

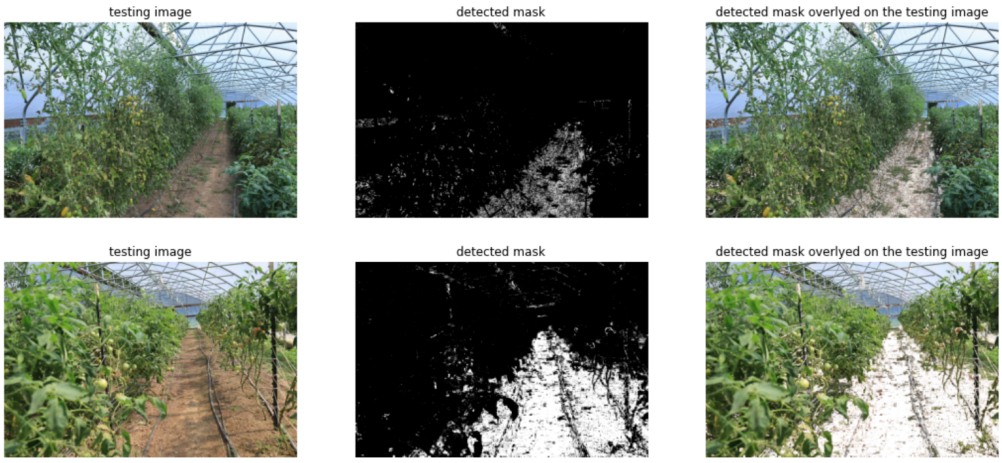
In this assignment, we are using machine learning to detect the land in the image. Our dataset has a total of 6 pictures, three greenhouse images, and their corresponding ground truth labels. We use a pair of them as the training image to get the detection algorithm. Then we use this algorithm to test the remaining two sets of pictures and get the detected results. Finally, we calculate the micro F-score to learn the reliability of our algorithm.

Method

Our algorithm is based on Bayes' rule. We identify the land by comparing the possibility that each pixel is a land pixel and the possibility that it is not land. We use the training image to get the RGB values of all pixels that may be the land. The method I use is to traverse all the pixels in the entire image and record the RGB values of all pixels that are land. We calculate the likelihood and prior distribution of ground and non-ground regions by using the RGB color space. I mainly use nested for loops to implement my algorithm.

Experiments


```
Pr_y_equalsTo_1: 0.36689186099679927
```




```
TP: 331803
FP: 118689
FN: 220718
Micro-average precision: 0.7365347220372392
Micro-average recall: 0.6005255908825184
Micro-average fscore: 0.6616125613526443
```

Training image: 3
Testing image: 1, 2

Discussions

The results of the program met our expectations. I use different images as training images to detect the other images. The accuracy is the lowest when using the first picture as the training picture. I think one of the possible reasons is because the RGB color space of the land color in the first picture is relatively small. It was shot on a cloudy day so the overall color was darker. This will lead to a large error in the detection of the other two images. For such a small dataset, this result is acceptable. If we have a large number of training images for machine learning, the accuracy will become higher. Similarly, for such a small amount of training information, if we reduce the bins, for example, to 32 bins, the accuracy will also increase.