INTRODUCTION TO IMAGE PROCESSING AND COMPUTER VISION

**Project 1: Plant segmentation and labeling**

Szymon Majorek

[s.majorek@student.mini.pw.edu.pl](mailto:s.majorek@student.mini.pw.edu.pl)

Table of contents

1. Introduction
   1. General introduction
   2. Part 1
   3. Part 2
2. Part 1 – plant segmentation
   1. Short introduction
   2. Creating ground truth
   3. Description of algorithm used
   4. Results
3. Part 2 – leaf labeling
   1. Short introduction
   2. Description of algorithm
   3. Results
   4. Part 2 conclusions
4. Conclusions

**1. Introduction**

**1) General Introduction**

In computer vision, image segmentation is the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify (and/or) change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects (our case) and boundaries. More precisely, it’s the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics. Result of image segmentation is a set of segments, which combined, covers the entire image. It can also be set of contours in case we are detecting edges.

In our task we are given set of 900 images of plants, from multiple cameras, at various stages of it’s growth. Our task is to first find the plant in an image and then to find individual leaves for each plant. Our segmentation is basically plant and not plant.

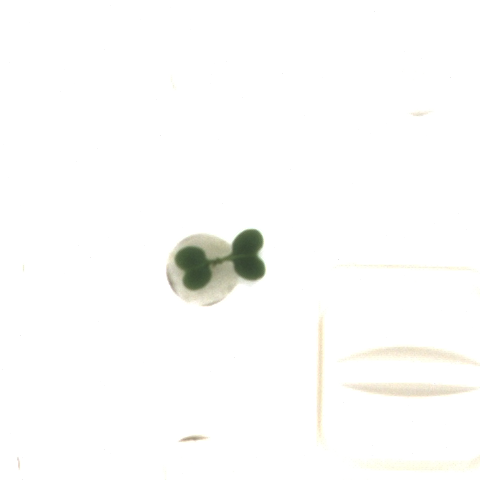


Figure 1 - Example 2

Figure 2 - Example 1 of input image

**2) Part 1**

This part, we could call the initial one, since we are supposed to separate everything that is plant from everything that is not a plant. Plant’s most prominent characteristics when compared to it’s background is it’s color. We will discuss that more in depth in chapter 2. Above this point we can see some examples of images that we need to process.

**3) Part 2**

Here however the task is more complicated, since we not only have to have our plants separated from the background, we also need to separate each individual leaf. For this we can use beforementioned object location and recognition of separate objects even though they do not appear to be separated on the initial image. In part 2 of this project, our input will be the output of part 1, since it’s pointless to perform the same task twice.

**2. Part 1 – plant segmentation**

**1) Short introduction**

As mentioned before, the goal of this part of the project is to, for each input image in our set of 900 images, produce output image which will be a binary mask for the plant present on input image, so in short the output in this part is binary mask containing plant for an input image. We also add some output used for comparison (Dice coefficient as well as Jaccard index more about which a little later).

**2) Ground truth for part 1**

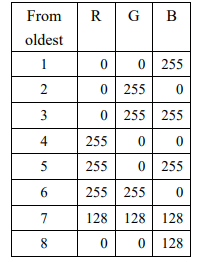
In addition to plant images (Figures 1 and 2), we are given our ground truth set containing images of plants with leaves labeled from youngest to oldest according to this color table:

Figure 3 - Ground truth color scheme

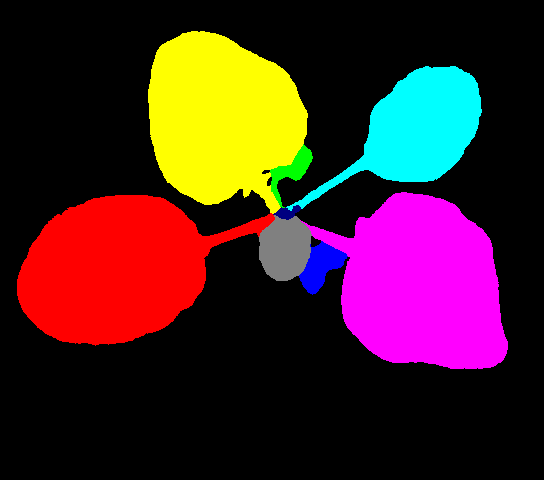
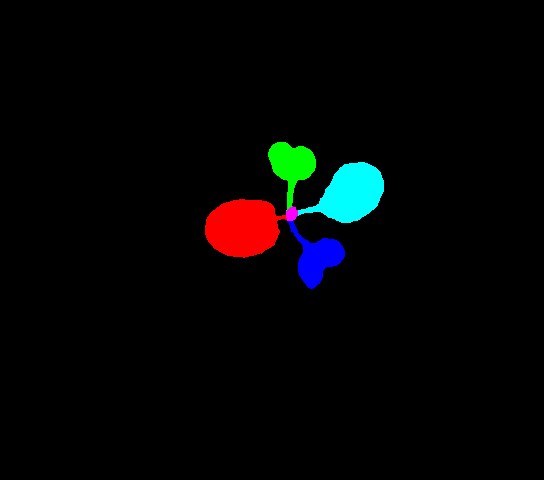
Below we can see some examples of said ground truth images we are given.

Figure 4 - Ground truth example 1

Figure 5 - Ground truth example 2

We can see that on the left picture we theoretically have 5 different leaves whereas on the picture on the right we have 8 different leaves, which is the maximum number of them in our set.

Now those ground truth images are virtually useless to us at this stage since to make a comparison and to calculate both Dice coefficient as well as Jaccard index we need to turn them into binary images. Since plants on them are separated flawlessly, we can simply separate what is colorful from what Is not colorful and it’ll produce us binary masks which we can use later on for comparison purposes.

The algorithm used for this is rather simple, we only convert the image to HSV scale and then apply thresholding:

Figure 6 - Code snipper for ground truth to binary mask transformation (full code in ground\_truth\_init.py)

mask = cv.inRange(hsv, (0, 0, 0), (255, 255, 255))

imask = mask > 0

green = np.zeros\_like(img, np.uint8)

green[imask] = img[imask]

grey = cv.cvtColor(green, cv.COLOR\_BGR2GRAY)

ret, thresh = cv.threshold(grey, 0, 255, cv.THRESH\_BINARY)

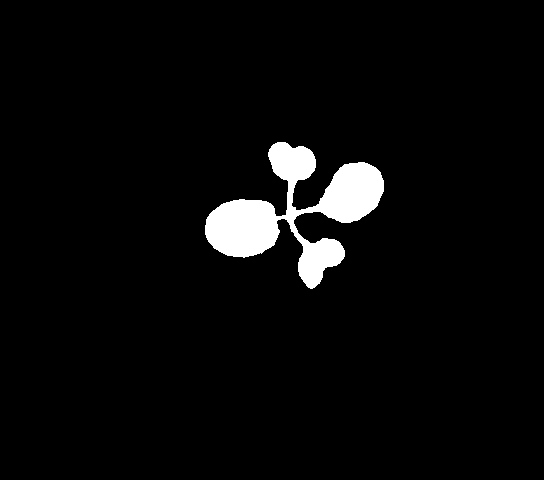
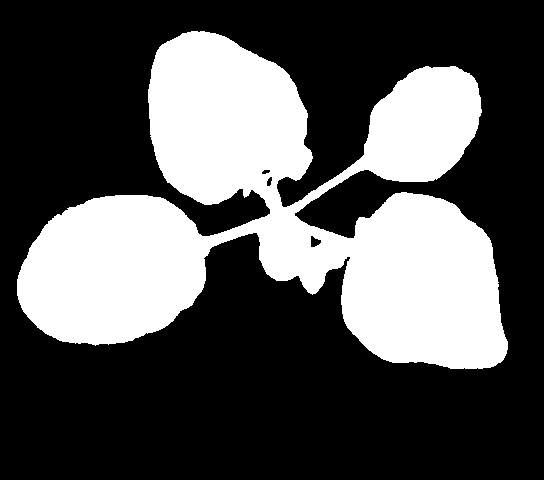
After those operation we get binary masks which we later on use to make comparisons to calculate how good was our algorithm of separating plants from not plants.

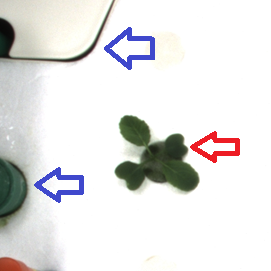
Figure 7 - Binary mask for example 2

Figure 8 - Binary mask for ground truth example 1

Figures 8 and 9 are the resulting masks after converting ground truth images to binary masks.

**3) Description of algorithm used**

Having those binary masks we can proceed to our actual algorithm. First part of this algorithm is very similar to the one used to obtain ground truth binary masks, however this case is not so straight forward. If we take a look again at some example images:

In case of actual images of plants we need to take into consideration that plants may not be as prominent as it was the case in ground truth images. As marked on figure 10, there may appear some regions which will have color similar to that of plant (blue arrows) and slightly more dense regions, where the lightning will make plant appear very similar in color to it’s surroundings, again making it hard to separate. Therefore we first apply thresholding for some specifically picked range in HSV (in my case the range was picked using mostly trial and error)

LOWER\_BOUND = np.array([40, 40, 37], np.uint8)

UPPER\_BOUND = np.array([70, 200, 150], np.uint8)

Figure 10 - Example plant, regions of interest may cause problems

Figure 9 - Example plant, clear image

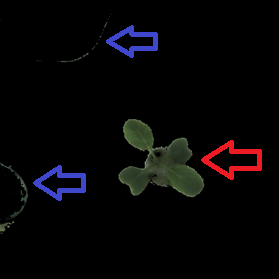
Which gave me the following results for example from figures 9 and 10;

Figure 9.2 - Green mask of example from figure 9

Figure 10.2 - Green mask for fig. 10

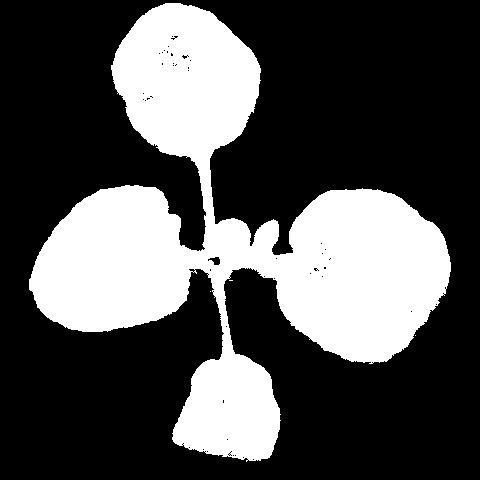
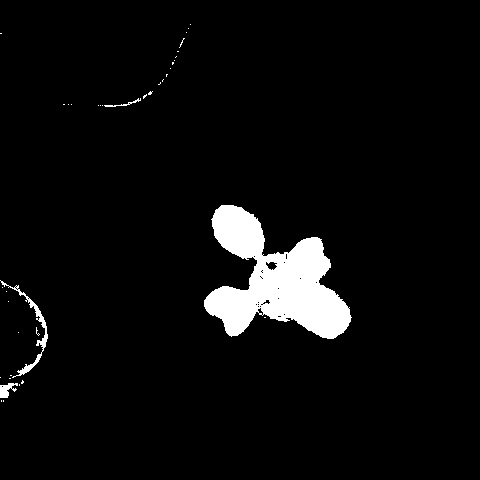
We can see that just as we predicted, fig. 9 example gave us rather clear plant with little to no noise and clear image whereas fig. 10 example produced some noise in regions marked by blue arrows, as well as unclear image in region marked via red arrow (part of plant’s pot got also marked as plant). Now we can transform those green images to binary masks:

Figure 10.3 - Binary mask for fig. 10.2

Figure 9.3 - Binary mask for fig. 9.2

Here it’s even more evident that there’s a vast difference in quality between those 2 images and their masks respectively. We can attempt to slightly correct them using 2 operations: opening and closing. Opening serves as a basic workhorse of morphological noise removal. Opening removes small objects from the foreground of an image, placing them in the background, whereas closing removes small holes in the foreground, changing small islands of background into foreground. First let’s see how those operation affect the first plant. After opening we obtain the following result:

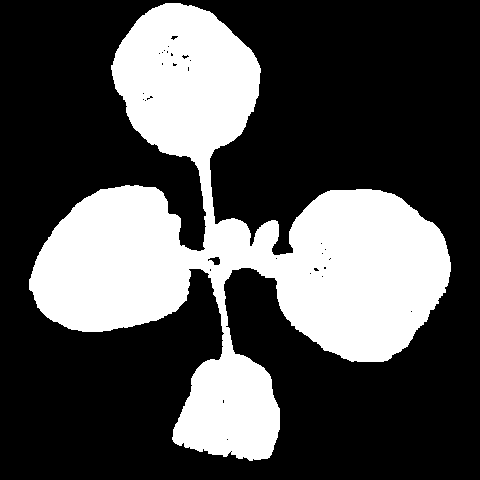
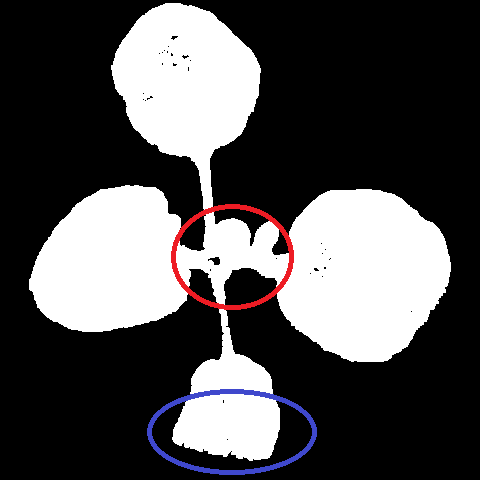
On the first look it might appear that not much has changed, therefore we will look at two specific regions to see what the operation of opening did to the binary mask.

Figure 9.4.2 - 9.4 with regions marked

Figure 9.4 - 9.3 after opening

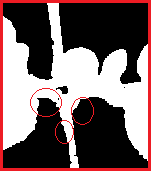
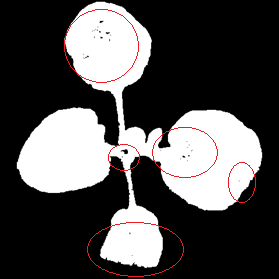
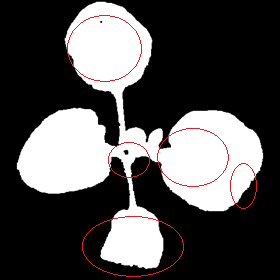
As we can see the process of opening very slightly reduced the noise present in the picture. Now let’s see what the process of closing will change.

Figure 9.5.6 - after closing

Figure 9.5.5 - before closing

Figure 9.5.4 - after opening

Figure 9.5.3 - before opening

Figure 9.5.1 - before opening

Figure 9.5.2 - after opening

In this case we can clearly see regions of interest at the first glance. We immediately notice that most of the holes present in gaps before closing operating were, in fact, closed after it and the image looks much smoother thanks to this.

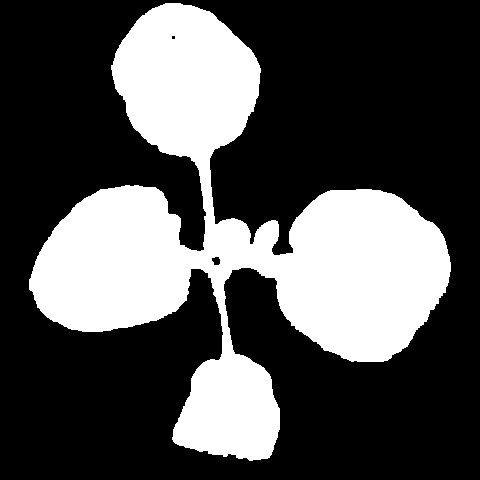
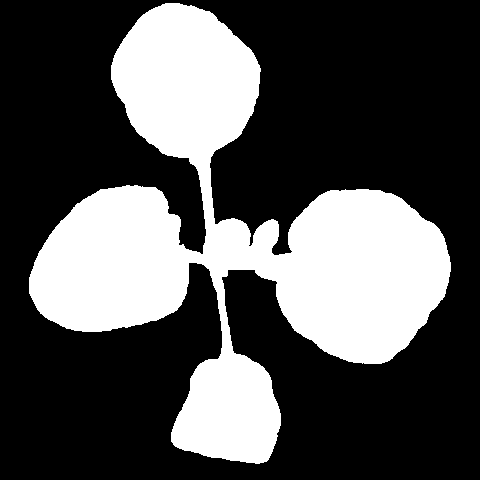
 As we can see, for this particular image our result was very satisfactory.

Figure 9.6.2 - obtained binary mask

Figure 9.6.1 - original image

Actually we can also compare our result to the ground truth binary mask we obtained earlier. Clearly we see that our result is nearly perfect in this case. Although we need to remember that this example was pick specifically to showcase example with good conditions, where the plant can be easily separated from the rest of the image. Now, knowing what to look out for we can move on to the second example, quickly going through each stage of the process.

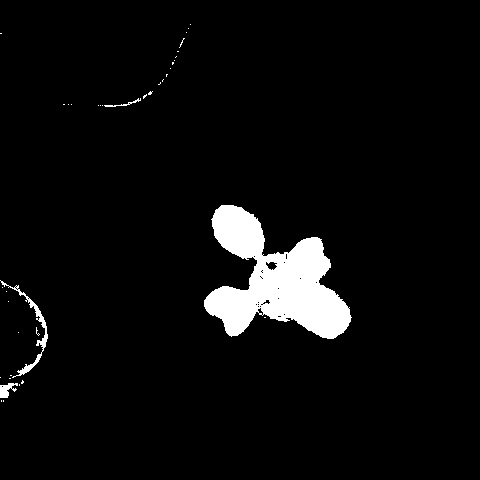
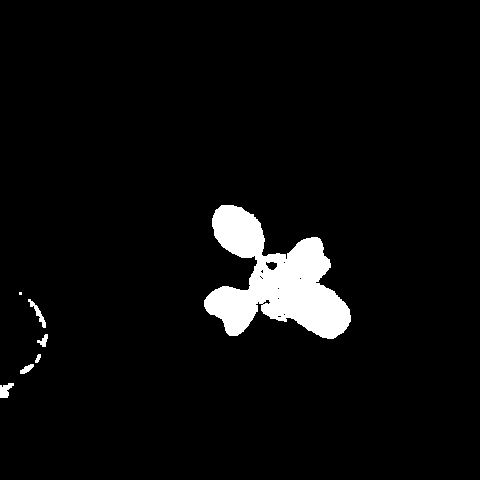


Figure 9.6.3 - ground truth binary mask

Figure 10.4.2 - after opening

Figure 10.4.1 - before opening

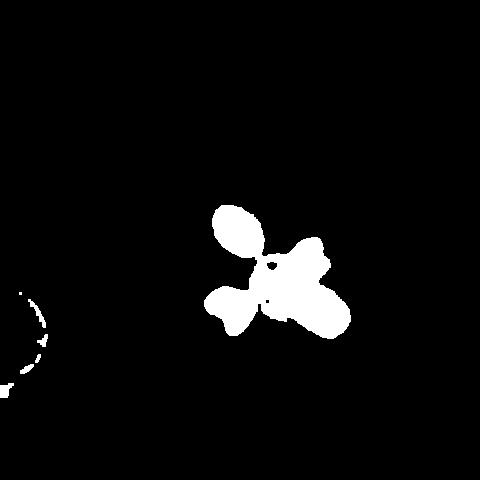
We can see that opening managed to get rid of the noise in the upper part of the image, however it wasn’t able to do anything useful about the mess in the middle of the plant.

Figure 10.5.2 - after closing - final result

Figure 10.5.1 - original image

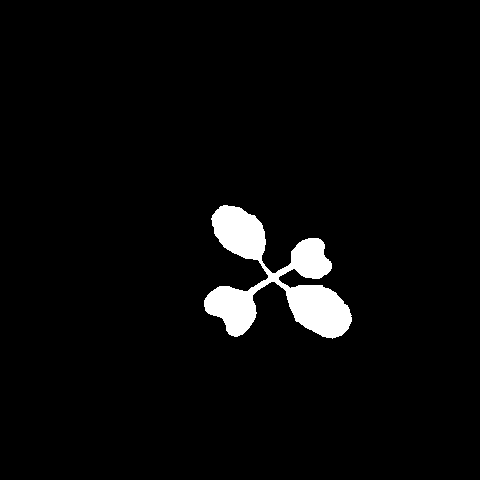
In contrast with the previous example, in this case our result could be labeled as rather terrible. The bottom right part of our plant looks like one big mess, whereas on the ground truth image It has clear stems going to each individual leaf. The lack of quality of the mask we obtained comes from the fact that on original image the pot in the near proximity of bottom right leaves has nearly the same color as the leaves, making it hard to segment the leaves without the pot. Additionally we were only able to get rid of the noise from the top of the image, however we the one on the lower left is still there, making the result worse. Those 2 example showed pretty well that even though we use the same algorithm for both images, results for one can be almost perfect, even when results for the other one are terrible.

Figure 10.5.3 - ground truth binary mask

**4) Results**

There are 2 types of assessment method we use to determine the quality of our prediction.

* Intersection over Union metric (Jaccard index) – statistic used in understanding the similarities between sample sets. It emphasizes the similarity between finite sample sets.
* Dice coefficient – statistic also used to gauge the similarity of two samples. Where A is target and B is prediction. asdasda

The results obtained for example used in algorithm description are as follows:

For the first example, where our result was very good, we obtained:

$ python comp\_good.py

Dice: 0.8440679771988706 Jaccard: 0.730205548898516

Where for the 2nd example we got:

$ python comp\_bad.py

Dice: 0.9873771824658591 Jaccard: 0.97506906359294

To calculate Dice coefficient and Jaccard’s index I used simple script and the fact that , where J is Jaccard’s index and D is Dice coefficient.

def calculateDiceC(im1, im2):

    im1 = np.asarray(im1).astype(np.bool)

    im2 = np.asarray(im2).astype(np.bool)

    intersection = np.logical\_and(im1, im2)

    return 2 \* intersection.sum() / (im1.sum() + im2.sum())

The results for each individual plant are as follows:

|  |  |  |
| --- | --- | --- |
| Plant Id | Dice coefficient | Jaccard’s index |
| 0 | 0.9253934259647113 | 0.8611462541958379 |
| 1 | 0.9348765191709363 | 0.8777165615044495 |
| 2 | 0.9311788492832148 | 0.871220455039402 |
| 3 | 0.9178763088482359 | 0.848217552534396 |
| 4 | 0.922676941918169 | 0.8564533498066876 |

Whereas for the whole set, the result is:

|  |  |
| --- | --- |
| Dice coefficient | Jaccard’s index |
| 0.9264004090370536 | 0.8628919169074337 |

As we can see the algorithm to separate plants from given image works rather well for this set, however it’s definitely not perfect. It takes between 7 and 7,5 seconds to execute python script calculating this for all 900 photos (not taking into account writing results to files and calculating dice coeffs and jaccard’s indexes).

**2. Part 2 – leaf labeling**

**1) Short introduction**

Having obtained the binary masks for each plant in the first part of this project, now we have to find and label each individual leaf for each plant. This task is much more sophisticated since we have to first find leaves, then mark them and then use that information to color each leaf in a given color.

**2) Description of algorithm**

The algorithm takes binary masks obtained in the 1st part of the project and then performs few operations on them in order to obtained desired result. First we erode the image, with iterations parameter bigger for later days (so that once the plant gets bigger, we erode a little more than for smaller plants). After erosion, ideally we have all leaves split so we can begin to actually find each leaf individually hence we find contours using cv2.findContours() function. Then, if the contour area isn’t too small, we find center of each contour, which is in theory the center of the leaf and then we use that information to color every pixel in the binary mask according to to which leaf center it’s closest.