

Programming Project 03

Francel Lamprecht, Christine Robinson, Regina Wehler

Winter Semester 2018/19

Contents

1	Introduction	2
2	Methodology	2
2.1	Segmentation	2
2.2	Objects Recognition	2
2.3	Object Analysis	4
2.4	Explorative Data Analysis	4
3	Results	4
4	Discussion & Outlook	4
5	References	4

1 Introduction

****REMOVE LATER!**This is a dummy sentence that shows how citations work (Adams et al., 2018).******

Plant phenotyping is a research area in the plant sciences that focuses on the quantitative measurement of both functional and structural properties of plants. During the last decade plant phenotyping evolved to non-destructive, image-analysis-based phenotyping. This computer driven approach allows the characterization of plants in a high-throughput manner (Walter et al., 2015).

The focus of this plant phenotyping research project was on the processing, segmentation and classification of Arabidopsis, to gain valuable insight on how the image-analysis-based phenotyping in the life sciences works.

Arabidopsis is often used as a model plant, as the plant is small in size, with a rapid life cycle of about six weeks (Koornneef and Meinke, 2010). Arabidopsis is a rosette plant, with the individual leaves being on separate, distinguishable stems. This characteristic of Arabidopsis makes leaf separation from each other less complex. Given that we were only exploring viable methods in this research area, Arabidopsis was our chosen plant, as it gave us the opportunity to explore with different algorithms in an easier way than what the other given data set, Tobacco, would have allowed us to do.

2 Methodology

The image analysis approach consists of three main steps (figure 1). First, the input RGB input image is segmented into two class, plant and background by a trainable Weka segmentation classifier. Second, the binary class output is used to identify single leaf objects by watershed segmentation. The resulting binary mask is utilized to analyze properties of the leaf objects in the single red, green and blue channel, respectively.

2.1 Segmentation

The plant is separated from the background with the help of a trainable Weka segmentation classifier, which is available as a plugin for Fiji. Weka provides a GUI to train machine learning algorithms to produce pixel-based segmentations. The user can add traces to classes and train the classifier with those. Afterwards, traces/regions of interest can be adjusted and the classifier can be re-trained to improve classification. Six representative pictures are selected from the 2017 data set for training (plant029, plant145 and plant159 from A1; plant032, plant034 and plant037 from A2). Pictures are chosen to cover the whole range of plant green shades and the range of background characteristics of the given data sets. The Weka Experimenter was used to assess the performance of different machine learning algorithms. Based on these results (table 1), FastRandomForest was used as a classifier. By applying the trained classifier to the RGB input images, for each input image a binary classification image is obtained (figure 1-b).

2.2 Objects Recognition

A plant consists of leaves that are attached to each other. To be able to analyze leaves individually, they need to be separated. Here, the watershed algorithm was used to separate touching objects. The algorithm first calculates an Euclidean distance map and determines center points as points which are, from a topological view, the ultimate eroded points, As the algorithm's name indicates, this topological map is "flooded" with water and at each

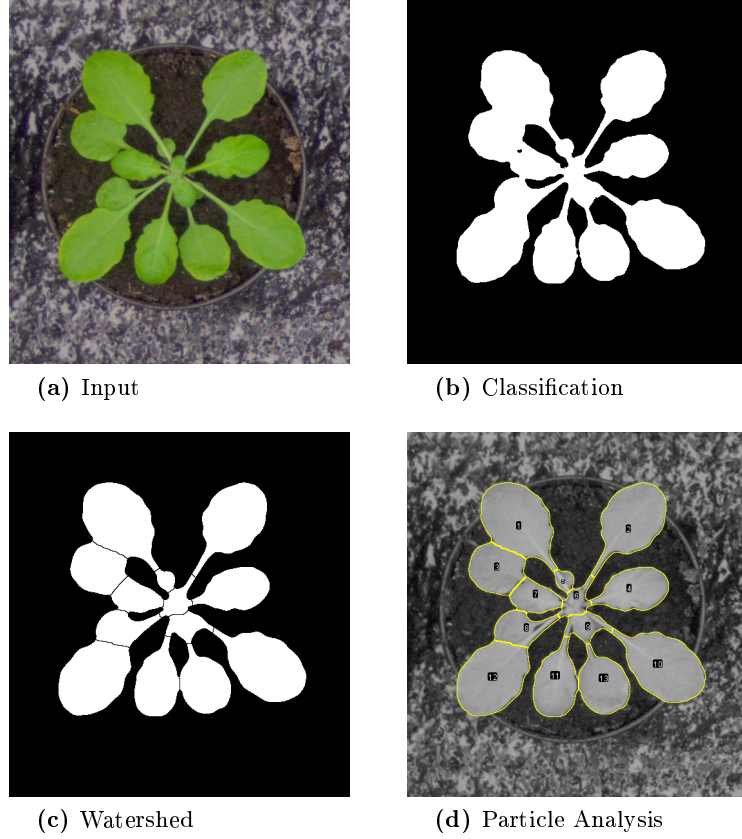


Figure 1 – The analysis of plant leaves is conducted in three steps. The RGB input image (a) is classified into plant and background by a trainable Weka segmentation classifier (b). Single leaf objects are identified by watershed segmentation (c). The binary mask is used to analyze properties in the split red, green and blue channel (d).

Table 1 – Comparison of classifier performance by Weka Experimenter. The default algorithm parameters were kept.

Algorithm	Percent correct	Precision	Recall	F score	Matthews correlation	AUC
FastRandomForest	99.93	1.00	1.00	1.00	1.00	1.00
SMO	92.52	0.89	0.93	0.91	0.85	0.93
<i>k</i> -nearest Neighbors	95.38	0.95	0.93	0.94	0.90	0.95
RandomSubSpace	99.85	1.00	1.00	1.00	1.00	1.00
Bagging	99.77	1.00	1.00	1.00	1.00	1.00
DesicisonTable	99.05	0.99	0.9	0.99	0.98	1.00

collision of two "watersheds", a line separating two objects is drawn. The output of the algorithm is the binary image with added watershed lines (figure 1-c).

2.3 Object Analysis

Particle Analyzer and additional stuff until csv export

2.4 Explorative Data Analysis

Everything in Python

3 Results

4 Discussion & Outlook

- watershed is working best for circular objects, that's why it sometimes fails on leaves with long stems

5 References

- Adams, T., Dörpinghaus, J., Jacobs, M., and Steinhage, V. (2018). Automated lung tumor detection and diagnosis in CT Scans using texture feature analysis and SVM. In *Communication Papers of the 2018 Federated Conference on Computer Science and Information Systems*, Annals of Computer Science and Information Systems, pages 13–20. PTI.
- Koornneef, M. and Meinke, D. (2010). The development of arabidopsis as a model plant. *The Plant Journal*, 61(6):909–921.
- Walter, A., Liebisch, F., and Hund, A. (2015). Plant phenotyping: from bean weighing to image analysis. *Plant Methods*, 11(1):14.