Unsupervised Learning Method for Plant and Leaf Segmentation

Noor M. AL-Shakarji†‡∗, Yasmin M. Kassim○†*, Kannappan Palaniappan†
† Department of Computer Science, University of Missouri-Columbia, MO 65211, USA
‡ Department of Computer Engineering, Al-Nahrain University, Baghdad, Iraq
∗ Department of Computer Science, University of Technology, Baghdad, Iraq
Email: {nmahyd, ymkgz8}@mail.missouri.edu, palaniappank@missouri.edu

Abstract—Plant phenotyping is a recent application of computer vision in agriculture and food security. To automatically recognize plants species, we need first to extract the plant and associated substructures. Manual segmentation of plant structures is tedious, error prone and expensive. Automatic plant segmentation is useful for leaf extraction, identification, and counting. We have developed a robust and fast unsupervised approach for plant extraction and leaf detection. K-means based mask (of the pot) followed by Expectation Maximization (EM) algorithm is adapted to estimate a mixture model for identifying the foreground area for the plant. We utilized the EM with 3 RGB channels to identify the foreground versus background for plant localization. K-means has been used to extract the circular plant can as one of the intermediate result to fuse it with EM results for noise removal since the images suffered from contrast and illumination variations. For leaf segmentation, we utilized distance transform and watershed segmentation to localize the leaves individually followed by stem link algorithm to connect the stem with corresponding leaves. The results have been evaluated by the same algorithms that have been used in the contest of plant phenotyping [1]. In our work, we used A1 and A2 datasets† to test our algorithm. We achieved promising score in some evaluation metrics and comparable in the others.

Index Terms—Plant segmentation, unsupervised learning, k-mean, expectation maximization, leaf segmentation

I. INTRODUCTION

Studying plants phenotype is important for Biologists. The appearance and performance describe plant functions. Identifying and evaluating phenotypes will help to increase production and prevent plant diseases. Previously, biologists extract such information manually from the plants, however, image-based approaches help the experts to reduce time consuming by monitoring the digital plant images automatically. Many monitoring systems that have been developed commercially need a constant environment for image acquiring to prevent any fault identification. However, the situations with respect to varying plant environment will cause problems for image analysis, i.e. challenging may be illumination changes, camera resolution and plants with different shapes and sizes. Improvement in monitoring systems, would be helpful to identify and track plant growth. Several of computer vision approaches for monitoring system focus on plant and leaf segmentation, counting, and growth tracking, which remain challenging for plant imaging area. For segmentation, there is a rich history of papers that either use supervised, unsupervised, or semi-supervised methods [2] [3] [4] [5]. Although tracking is widely used on surveillance system [6] [7], there are also interesting papers in plant tracking, e.g [8] [9] segment and track leaves growing over time. In this paper we used a dataset provided through the Leaf Segmentation Challenge (LSC) of the Computer Vision Problems in Plant Phenotyping (CVPPP 2014) workshop (ECCV) [1]. This data set consists of two different species of plant (Arabidopsis and Tobacco). Many authors work on this data set with different techniques. Pape et al. used unsupervised clustering and distance maps to segment leaves [10]. Nottingham leaf segmentation process [11] uses SLIC [12] superpixels-based method that does not require any training with distance map and watershed transform to extract leaves. Yin, et al. extends a multi-leaf alignment and tracking framework [13] that used in different plant data set by using advanced segmentation [11].

The rest of this paper is organized as follows: Section II Describes our proposed method. Section III describes the evaluation metric that we used to evaluate our pipeline, while section IV shows the experimental results followed by conclusion.

II. OUR PROPOSED METHOD

There are two main steps to accomplish our goal, step 1 is to segment the region as plant and non plant regions and step 2 is to extract the individual leaves from the segmented region. For the first step, there are many methods for segmentation [14]– [20], however, we utilized the Expectation Maximization (EM) and K-means algorithms to provide fast automated system. Extracting the plant by the EM algorithm leads to noisy result due to image noisy acquisition, e.g. intensity and illumination variations, and leaves sizes. Figure 2 shows some images with

* These authors contributed equally.
†http://www.plant-phenotyping.org/CVPPP2014-dataset
these challenges that can affect the segmentation. In addition, utilizing k-means alone is not practical since it is very easy to stick in local optimal. Thus, EM and K-means are both utilized by fusing their results to get a pure and clear mask for foreground and background. For the second step, leaf segmentation, we followed some determined morphological stages with distance transform of local maxima and watershed segmentation approach to get the leaves isolated.

**A. Step1: Foreground background (FB) segmentation**

In this section, we describe how to fuse the two methods; EM and K-means, to get our final FB segmentation masks.

1) **Method1: Expectation-Maximization** : We utilized EM for the main part of the extraction since it tends to get stuck less than K-means in local optimal. The idea is to assign data points partially to different clusters instead of assigning to only one cluster. To do this partial assignment, we model each cluster using a probabilistic distribution, so there is a certain probability assigned to each data point, the final clusters assignment will depends upon the highest probability score. Mixture of Gaussian is used to represent our model which is a weighted sum of K Gaussian distributions, the summation of all the weights is 1. let’s assume \( \theta_j \) as the parameter of jth distribution with \( w_j \) is the corresponding weight as described in equation (1)

\[
p(x_i \mid \Theta) = \sum_{j=1}^{K} w_j p_j(x_i \mid \theta_j)
\]

The goal is to find parameter of the model \( \Theta \) that maximize the log likelihood over our distribution. The EM work iteratively with two main steps, the E-step which creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters, and a maximization (M) step, which computes parameters that maximizing the expected log-likelihood found on the E step. These parameter-estimates are then used to determine the distribution of latent variables in the next E step until convergence as described in equation (3). In our work, \( X \) is our feature vector that corresponds to RGB color. We train only one image to get the parameters, then test the rest of images depending upon those estimated parameters.

\[
\log(P(X \mid \Theta)) = \log(\prod_{i=1}^{N} p(x_i \mid \Theta)) = \sum_{i=1}^{N} \log(\sum_{j=1}^{K} w_j p((x_i \mid \theta_j))
\]

2) **Method2: K-means segmentation** : K-means is one of the unsupervised algorithms that usually used in many applications, the idea is to cluster similar points into isolated clusters depending upon two parameters centroid and number of clusters. In each iteration, it assigns each object to the cluster that has the closest centroid, then recalculate the new centroids from the assigned data clusters. The algorithm converges by minimizing the total distortion in our objective function as described in equation (3)

\[
J = \sum_{j=1}^{K} \sum_{i=1}^{x} \|x_i^{(j)} - c_j\|^2
\]

K-means is used to get the circular mask for the plant’s can to get rid of the false positive noise and undesired segments form background caused because of noisy image acquisition. Our input to the algorithm is a and b vector from the LAB color space since it perceptually uniform and closer to human vision, the output is three images represent the three clusters, we will use just one set of the clusters as a mask to EM results. All the intermediate results are shown in figure 3

**B. Step2: Leaf segmentation**

After extracting the plant form background, we need to identify the leaves of the plant individually. We use distance map calculation [21] to extract the centroid of each leaf from the local maximum of distance transform map. Watershed segmentation [22] is applied on distance map with optimal initial marker (centroid) over the image space. To connect the stems of the leaves, we develop the stem link algorithm (SLA) which dilate the stems to be connected to the surrounding regions, then assign the stem to the largest connected component sticked to it. The idea behind that is the fact that the stem is always oriented to the corresponding leaf, which is usually pick the larger connected leaf, so dilating them leads to know for which leaf or connected component that stem belongs to as shown in figure 4.

**III. Evaluation measures**

For our pipeline we adopted evaluation metrics that used in the LSC challenge [1]. LSC depended on four important metrics that also used in multi object segmentation. These metrics
based on Dice score for binary segmentation that measures how much overlap between ground truth and algorithm result of the binary segmentation masks. The four metrics used in our pipeline are:

- Symmetric Best Dice (SBD): used to calculate the segmentation accuracy (average Dice) among all objects (leaves), Best Dice (BD) is defined as (4)
  \[
  BD(L^a, L^b) = \frac{1}{M} \sum_{i=1}^{M} \max_{1 \leq j \leq N} \frac{2|L^a_i \cap L^b_j|}{|L^a_i| + |L^b_j|}
  \]
- Foreground - Background Dice (FBD): represents the Dice score of the foreground mask with respect to the ground truth.
- Difference in Count (DiC): the difference between the leaves in pipeline result and the ground truth (See equation (6))
  \[
  Dic = \#L^{ar} - \#L^{gt}
  \]
- $|DiC|$: which is the absolute value of DiC

IV. EXPERIMENTAL RESULTS

In this section, we discuss the evaluation of our method according to the same metric functions that have been used in the contest. We compare our results with the methods participated in the challenge as described on [11]. Our method has promising results on all the metrics, we got better values in term of SBD and comparable with the other state-of-art results as shown in table I for set A1 data set, and table II that shows our results for A2 dataset. See figure 5 which shows some of our final algorithmic results. Although A1 is very different from A2, the same parameters are used. One image for training is enough to train the parameters. In testing phase, the learned parameters are used to test all the other images in the datasets.

![Fig. 4. Leaf Detection (LD) steps](image)

### TABLE I

| Method | SBD | FBD | $|DiC|$ | DiC |
|--------|-----|-----|--------|-----|
| IPK    | 74.4 | 97.3 | 2.2    | -1.8|
| Nottingham | 68.3 | 95.3 | 3.8    | -3.5|
| MSU    | 66.7 | 94.0 | 2.5    | -2.5|
| Wageningen | 71.1 | 94.7 | 2.2    | 1.3 |
| Ours   | 74.6 | 94.9 | 2.9    | -1.4|

### TABLE II

| Method | SBD | FBD | $|DiC|$ | DiC |
|--------|-----|-----|--------|-----|
| IPK    | 76.9 | 96.3 | 1.2    | -1.0|
| Nottingham | 71.3 | 93.0 | 1.9    | -1.9|
| MSU    | 66.6 | 87.7 | 2.0    | -2.0|
| Wageningen | 75.7 | 95.1 | 0.4    | -0.2|
| Ours   | 72.1 | 94.8 | 1.3    | -0.8|

V. CONCLUSIONS

Plant and leaves segmentation algorithms have been developed for plant and leaves phenotyping segmentation. Our method has two steps, one for foreground-background segmentation that is unsupervised based, the second step isolate and segment the leaves individually based on watershed and morphological operations followed by stem linking algorithm. We have tested our work on the A1 and A2 datasets provided for the participants as described in [11]. Plant and leaves segmentation has comparable results compare with the other state-of-art results. We are currently working on achieving further improvement beside testing on dataset 3 provided in the contest.

### REFERENCES


Fig. 5. (a) Input images (selected from A1 group), (b) Ground-truth, (c) Foreground background segmentation using EM and K-means algorithms, (d) Leaf detection using watershed and morphological operations.


