

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/258518075>

# Image-Based plant phenotyping with incremental learning and active contours

Article in Ecological Informatics · September 2014

DOI: 10.1016/j.ecoinf.2013.07.004

---

CITATIONS  
91

READS  
624

---

3 authors:



Massimo Minervini  
IMT School for Advanced Studies Lucca

19 PUBLICATIONS 658 CITATIONS

[SEE PROFILE](#)



Mohammed Abdelsamea  
Birmingham City University

31 PUBLICATIONS 330 CITATIONS

[SEE PROFILE](#)



Sotirios Tsaftaris  
The University of Edinburgh

183 PUBLICATIONS 2,028 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Tumour Parcellation and Quantification for Immunohistochemistry Images of Colorectal Cancer [View project](#)



Non-rigid Histological Image Registration using Convolutional Neural Network [View project](#)

# Image-based plant phenotyping with incremental learning and active contours

Massimo Minervini<sup>a</sup>, Mohammed M. Abdelsamea<sup>a</sup>, Sotirios A. Tsaftaris<sup>a,\*</sup>

<sup>a</sup>*IMT Institute for Advanced Studies, Piazza S. Ponziano, 6, Lucca, Italy*

---

## Abstract

Plant phenotyping investigates how a plant's genome, interacting with the environment, affects the observable traits of a plant (phenome). It is becoming increasingly important in our quest towards efficient and sustainable agriculture. While sequencing the genome is becoming increasingly efficient, acquiring phenotype information has remained largely of low throughput. Current solutions for automated image-based plant phenotyping, rely either on semi-automated or manual analysis of the imaging data, or on expensive and proprietary software which accompanies costly hardware infrastructure. While some attempts have been made to create software applications that enable the analysis of such images in an automated fashion, most solutions are tailored to particular acquisition scenarios and restrictions on experimental design. In this paper we propose and test, a method for the segmentation and the automated analysis of time-lapse plant images from phenotyping experiments in a general laboratory setting, that can adapt to scene variability. The method involves minimal user interaction, necessary to establish the statistical experiments that may follow. At every time instance (i.e., a digital photograph), it segments the plants in images that contain many specimens of the same species. For accurate plant segmentation we propose a vector valued level set formulation that incorporates features of color intensity, local texture, and prior knowledge. Prior knowledge is incorporated using a plant appearance model implemented with Gaussian Mixture Models, which utilizes incrementally information from previously segmented instances. The proposed approach is tested on *Arabidopsis* plant images acquired with a static camera capturing many subjects at the same time. Our validation with ground truth segmentations and comparisons with state-of-the-art methods in the literature show that the proposed method is able to handle images with complicated and changing background in an automated fashion. An accuracy of 96.7% (Dice Similarity Coefficient) was observed, which was higher than other methods used for comparison. While here it was tested on a single plant species, the fact that we do not employ shape driven models and we do not rely on fully supervised classification (trained on a large dataset) increases the ease of deployment of the proposed solution for the study of different plant species in a variety of laboratory settings. Our solution will be accompanied by an easy to use graphical user interface and, to facilitate adoption, we will make the software available to the scientific community.

**Keywords:** Agriculture, Sustainability, Phenotyping, Plant segmentation, Machine learning, Gaussian mixture model, Active contour model

---

## 1. Introduction

Plants have always been a crucial source of food, feed, fiber, and fuel. Thus, striving for a more sustained agriculture [1], together with the breeding industry, researchers try to identify, improve, and breed key traits to satisfy growing demand, increase resistance to parasites and diseases, and minimize environmental impact (less water, less fertilizer).

Understanding biological function and the complex processes involved in the development of plants relies on understanding the interaction between genetic information and the environment, and how they affect the phenotype (the appearance or behavior) of the organism and consequently desirable traits. Model plant systems, such as *Arabidopsis*

---

\*Principal corresponding author

Email addresses: [massimo.minervini@imtlucca.it](mailto:massimo.minervini@imtlucca.it) (Massimo Minervini), [mohammed.abdelsamea@imtlucca.it](mailto:mohammed.abdelsamea@imtlucca.it) (Mohammed M. Abdelsamea), [s.tsafaris@imtlucca.it](mailto:s.tsafaris@imtlucca.it) (Sotirios A. Tsaftaris)

thaliana, have been used extensively for this purpose [2]. Unlike humans and most other mammals, they are considered ideal models to study natural variation and decipher the “genotype to phenotype” link [3] – arguably a problem that concerns all life science fields. It is expected, that the integration of these findings with bioinformatics and systems biology will lead to the construction of a “virtual plant analog”, which will permit researchers to investigate gene activity at every stage of plant development [4, 5].

The rate of throughput for acquiring genetic information (with sequencing and microarrays) has achieved game-changing levels [6]. However, as of today, inexpensive and automated phenotyping (phenomics) remains a bottleneck [7, 8]. Until recently most phenotypes (e.g., related to plant growth) were acquired in destructive ways (e.g., weigh the plant, or cut out and measure a leaf) or involved human survey (e.g., measuring leaf size or plant radius) *in situ* without destructing the plant. Naturally these methods are faced with low throughput and high productivity cost. Consequently, there has been a growing interest towards developing solutions for the automated analysis of visually observable traits of the plants. Several consortia such as the International and European Plant Phenotyping Networks (IPPN and EPPN), and the iPlant Collaborative Project [9] have been established to promote and accelerate phenotype discovery and analysis, and increase our understanding of biology [5, 10, 11]. Several approaches based on imaging techniques and computer vision have been proposed to increase the throughput of non-destructive phenotyping (without penalizing accuracy) with solutions that can be affordable and easy to deploy [8, 12]; however, such systems usually require sophisticated analysis algorithms to segment the plant from the background. As of now the majority of solutions pose strict experimental conditions to ease the complexity of the analysis task that follows.

In this paper we propose and test, an algorithm and a software system for the automated segmentation and analysis of time-lapse top-view plant images from phenotyping experiments of *Arabidopsis* rosettes. Example images of *Arabidopsis* rosettes (referring to the circular and radial cluster of leaves they form when growing) are shown in Figure 1. We use data which we acquired in a general laboratory setting with a static camera that captures many plants at the same time [13], with the purpose of collecting test data (which are lacking in the public domain) and demonstrating the challenging aspects of the problem of plant segmentation. While our paper focuses on the algorithm and the software solution, as an example that relates it to phenotyping experiments we measure plant growth, estimated through projected rosette area of 19 *Arabidopsis Columbia* (Col-0) wild-type plants. The system involves minimal user interaction (necessary to establish the statistical experiments that follow) and at every time instance (in our context a digital photograph) segments the plants in images that contain many instances of the same species (an example input image is shown in the bottom of Figure 1). We rely on a combination of level set and learning based segmentation to incrementally incorporate information from previous time instances, allowing us to adapt to changes in the scene. We learn an appearance model of the plant relying on Gaussian Mixture Models (GMM) of color and texture features collected from previously segmented instances. For each unseen instance we classify each pixel in the image to obtain a probabilistic map of pixels most likely belonging to a plant.

The probabilistic output assists the localization of a plant within an image (many plants exist in an image). Each individual plant is then segmented using a new active contour model that incorporates probabilistically weighted (using the model output) features of pixel intensity and texture. Once each plant has been segmented, several measurements relevant to the study of plant growth are extracted. Finally, the plant appearance model is updated by re-training the GMM to include the newly processed data in an online and incremental fashion.

Our inclusion of the iterative learning aspect allows us to handle images with complicated and changing background in an automated fashion, which challenge currently available solutions. Overall the proposed approach obtains an accuracy in segmentation higher than 96.4% (Dice Similarity Coefficient) and is not affected by challenges in the scene. While here it was tested on one mutant of *Arabidopsis*, the fact that we do not employ shape driven models and we do not rely on fully supervised classification (trained on a large dataset) favors the deployment of the proposed solution for the study of different plant species in a variety of laboratory settings. The automated phenotyping solution proposed in this paper improves upon the accuracy results obtained by the state of the art in plant phenotyping, even in an environment that is not strictly controlled, thus accommodating a broader range of experimental scenarios – handling multiple plants simultaneously without an explicit scene arrangement and with minimal user intervention. Our solution will be accompanied by an easy to use graphical user interface and, to facilitate adoption, we will make the software available to the scientific community.

The remainder of the paper is organized as follows. In Section 2, we review several approaches for image based phenotyping with regard to acquisition and analysis, while Section 3 presents the proposed approach, discussing the requirements imposed by the application (i.e., the segmentation of rosette plants) and the methodology we adopted

based on an incremental learning framework that combines active contour models and GMMs. In Section 4 we present experimental results on the use of our approach for the study of *Arabidopsis* plant growth, with respect to a reference method and a specialised software proposed in [14]. We also compare with recent innovation in image segmentation. We describe our testbed setup for the collection of time-lapse imaging data used here, and detail all parameters involved in the operation of the systems being compared. Subsequently, we present and discuss quantitative results showing the ability of our solution to automatically and accurately segment plants in images, in order to calculate projected area and estimate growth. Finally, Section 5 offers conclusions and directions for future work.

## 2. Related work

Plant phenotyping can occur on small scale (controlled laboratory settings of a growth chamber), in the green house, or in the field [10, 15, 16]. While each setup aims to address different experimental questions, the majority of early stage experiments occur in growth chambers that offer controlled laboratory settings in a small scale using plant models. In the following paragraphs we quickly overview growth chamber phenotype acquisition strategies and we subsequently review image processing approaches towards extracting phenotypes.

### 2.1. Phenotype acquisition

The introduction of automation and digital imaging allowed the rapid collection of time-lapse images of plants in a non-destructive fashion [17]. These images were later analyzed offline mostly using expert analysis by manipulating photograph analysis software or semi-automated methods [17]. This had a particular impact on small scale experiments of plant models in growth chambers and several approaches have been proposed throughout the years. Most of the solutions are customized [13, 17, 18, 19, 20, 21] imposing strict experimental setups (e.g., uniform black background). Their cost and implementation complexity could range from a few dollars (e.g., [13, 22, 23]) to a few thousands (e.g., [18, 19, 21, 24]), or to hundreds of thousands of dollars (e.g., the solutions provided by commercial entities such as LemnaTec and CropDesign).

High resolution images can also be captured in the visible and/or infrared spectrum. Optionally, also 3D depth information is acquired using high grade [24, 25] or lower grade (e.g., using Microsoft’s Xbox Kinect) hardware [26]. However, their limited field of view requires that images are acquired with each plant in isolation and usually involve some robotic (or manual) placement of the plant (e.g., [23]) or the camera (e.g., [18, 21]), which lowers throughput or increases equipment cost.

On the other hand, some approaches assume static cameras with a fixed field of view [13, 14, 22, 23] and are largely used in growth-chamber experiments of small plant models such as *Arabidopsis*. Such systems are simpler to deploy, are more affordable (since they do not use automation), and have higher throughput since they can image many plants at a time. However, due to the fixed camera, the per plant imaging resolution can be lower, and only few imaging angles are obtained (usually only a top-view is used); thus, requiring sophisticated image analysis algorithms.

### 2.2. Image analysis for phenotyping

In the following paragraphs we briefly describe approaches aiming at segmenting the plant from the background, which in turn is used to measure plant growth. To appreciate the segmentation problem at hand, Figure 1 shows the variability of appearance and shape (i) among several mutants of the *Arabidopsis* family, (ii) within individuals of the same mutant, and (iii) through the life-cycle of the same plant. Although several other methods have been developed for the segmentation (and classification) of individual leaves [27, 28, 29] (e.g., for plant identification purposes [30, 31]), the analysis of plant vein structure [32, 33], and the study of root or hypocotyl growth [34, 35], in the following we will focus on those segmenting the whole plant.

The majority of articles in plant phenotyping, since they assume strict conditions as to the composition of the scene, follow simplified approaches to the subsequent intensity segmentation problem. They usually employ thresholding based approaches in one channel (e.g., the native green channel of an RGB acquisition or on one channel of other color spaces) or multiple channels. Those with strict restrictions on the scene rely either on calibrated thresholds (e.g., [23]), or on histogram based methods such as Otsu’s thresholding or Gaussian mixture modeling of intensities to identify data driven thresholds [14]. Some methods rely on combining more than one channel indirectly (e.g., in [14] individual segmentations from each channel obtained via thresholding are combined) or directly via unsupervised

clustering methods [13, 36]. However, changes in the scene can pose a challenge. Such changes occur frequently as Figure 2 illustrates. For example the soil could appear more or less dry, moss (which is green in appearance) can grow on the soil, reflections due to water and illumination distortions can be also present. (While some users will not remove such data from analysis, this lowers the number of instances used for analysis and consequently the statistical power of the experiment. Furthermore, approaches such as disturbing the soil to remove out moss may change a plant’s position, breaking the location correspondence from previous images.) From an imaging perspective, few systems utilize additional information from other imaging sources (such as depth or infrared information) to facilitate the segmentation approach [24, 25, 26]; however, few laboratories have the capacity to deploy such sensing technology due to its higher cost and lower (in some cases) throughput.

On the other hand, one approach towards dealing with complex backgrounds and changes in the scene is to introduce prior knowledge into the segmentation approach itself. Although naturally one would consider to introduce shape priors in deformable models and contour formulations [37, 38], it would require defining and learning shape priors of all possible plant shapes. As Figure 1 illustrates, the complexity of shape even among mutants of the same plant or even among samples of the same plant species can be significant. Thus, building appropriate plant models is rather complex and learning such shape models would require a large amount of previously labeled information; as a result, no such method exists in the literature. An alternative approach would be to rely instead on leaf models, and try to segment individual leaves within the scene. Along these lines, Felzenszwalb [39] represents shapes by deformable triangulated polygons to detect precisely described objects, including maple leaves. De Vylder et al. [29] proposed a probabilistic parametric active contour model that optimizes an energy function by maximizing the probability that the contour is on the edge of a leaf. In simplifying the shape descriptor, the authors in [28] proposed a parametric active polygon model to take advantage of prior knowledge on leaf shapes to design a flexible time-efficient model, in a classification and leaf recognition problem. Although these methods are capable of segmenting individual leaves in isolation, they require prior training and large labeled datasets, usually they cannot handle occlusions, and they would require significant post processing to compose the final complete plant segmentation.

### 3. Proposed approach

As outlined in the previous section, the current state of the art in the automated analysis of images from phenotyping experiments is limited to a couple of commercial solutions and few freely available software. There is a great interest in developing an approach that:

- can accommodate most laboratory settings;
- does not rely on explicit scene arrangement;
- can handle multiple plants in the same image;
- can tolerate low spatial resolution and out of focus discrepancies, attributed to a fixed broad field of view (thus accommodating affordable sensors);
- is robust to changes in the scene;
- requires minimal user interaction (for experimental setup and training);
- is scalable to large population sizes or high sampling rates;
- offers high accuracy and consistent performance; and
- is automated and high throughput.

Such an approach, when combined with affordable sensing hardware (e.g., commercial cameras) would provide a truly affordable and easy to deploy phenotyping collection and analysis system that can satisfy the needs of most laboratories. Here we propose a software solution that satisfies the above requirements using several innovations, which have not been previously considered in the context of plant segmentation. First of all we rely not only on color features, but also on texture information extracted from the images to effectively discern foreground (plant) from

background. To increase robustness and offer consistent performance we incorporate a learning component in our solution. To favor scalability, we learn the color and texture appearance of plants using a multi-dimensional Gaussian Mixture Model (GMM), which minimizes additional computational and storage requirements. The appearance model is learned incrementally after each instance has been automatically segmented, thus reducing the need for prior labeling significantly. Finally, we propose a new vector valued level set formulation to segment each plant in the image by incorporating all features and prior information, in a joint energy optimization framework.

Figure 3 provides a general view of the proposed system. Once a new image is acquired, the first step consists in localizing plant objects in the image and obtaining an approximate plant segmentation (which is used as an initialization for the next step). Subsequently, each region containing a plant is segmented with an active contour model to accurately delineate the plant from the background. Both of these steps can take advantage of the plant appearance model employed here. Each segmented plant is labeled with an identifier coherently with previous images. This permits us to follow each plant individually across time and link all new measurements and analysis output to the corresponding plant. Finally, the processed data, the segmentations, and several indices of interest to the plant community are added to a repository. In addition, the features of foreground (the plant) are stored and used to update the plant appearance model.

Overall, the proposed solution has been designed to work automatically and minimize user interaction without compromising accuracy and generalization. The user at the beginning of the experiment informs the application about the number of plants present in the scene and their grouping if any (e.g., which mutants are present). Our only assumption regarding the composition of the scene, is that plants should not be touching (this requirement is reasonable and common among all phenotyping platforms to facilitate object separation). To use the plant appearance model, some initial training is necessary. This initial training can be provided by the user offline or alternatively the user can operate aspects of the proposed solution without the model. This can take advantage of the fact that some scenes can be less complex (e.g., plants are smaller, or moss has not appeared yet). The user is prompted to make corrections to the segmentations (if necessary) on those (early-training) images, and subsequently the model is trained based on this output. The user could also adjust some parameters regarding the model during this training part. Afterwards the plant model is updated at each iteration and is used at each step, without any user supervision.

In the following paragraphs we present each aspect of the proposed system in detail, providing insights into both mathematical formulation and design choices.

### 3.1. Image Features

From a computer vision perspective, a laboratory setup for plant phenotyping experiments presents several challenges such as neon light illumination, water reflection, shadows, and moss (some examples are shown in Figure 2), contributing to noise and scene complexity. To eliminate issues of non uniform illumination (due to lighting distortion from neon lights and shadowing), when utilizing color information we convert the RGB color space to the 1976 CIE  $L^*a^*b^*$  color space [40]. We use the  $a^*$  component which determines the color position between green and red, and the  $b^*$  component (colors between blue and yellow) as color features.

While relying only on the intensity of pixels may appear adequate, there are several conditions that would challenge this assumption. For example, the color intensity of moss can be very close to that of plant leaves. This motivates the utilization of texture features along with intensity to describe foreground and background. Several texture representations can characterize texture content at each pixel location, taking into account its neighborhood [41, 42]. Some of these approaches (e.g., co-occurrence matrices, Haralick features, and Gabor filters) result in a multi-resolution, multi-scale representation of local texture characteristics. This in turn introduces additional feature dimensions.

In this work we use a texture detection filter obtained via grayscale morphological operations as a local texture descriptor. In order to detect high texture regions in an image  $I$ , the response of a pillbox filter is linearly combined with a Difference of Gaussians (DoG) filter. A pillbox filter is a circular uniform kernel  $H_\rho$  of radius  $\rho$ . The DoG filter operates by subtracting a blurred version of an intensity image from another blurred version of the same image, where the different blurring is usually obtained by convolving with a pair of Gaussian kernels ( $K_{\sigma_H}, K_{\sigma_L}$ ) of different standard deviation ( $\sigma_H, \sigma_L$  respectively). The filtered output  $f$  is defined as:

$$f(I; \rho, \sigma_H, \sigma_L) = H_\rho * I_i + (K_{\sigma_H} * I_j - K_{\sigma_L} * I_j), \quad (1)$$

where  $I_i$  and  $I_j$  are channels of the same multi-channel image  $I$ , and  $*$  denotes the discrete 2D convolution operator. The response of the “Texture From Blurring” (TFB) filter finally is:

$$f_{\text{TFB}}(I; \alpha, \rho, \sigma_H, \sigma_L) = \exp(-\alpha |f(I; \rho, H, L)|), \quad (2)$$

where  $\alpha$  sets the decrease rate. The pillbox filter responds to smooth regions versus high texture regions, while the DoG filter responds to edges. In the context of plant images, the combination of these two operations permits the separation of high texture regions (e.g., belonging to moss or earth) from smooth regions (e.g., belonging to leaves and stems).

Our system thus relies on color information ( $a^*$ ,  $b^*$ ) and texture features (TFB) as a feature space to perform the image analysis tasks described in the following sections. Figure 4 shows examples of the features employed as extracted from an image. (For these examples TFB was found using  $I_i = a^*$ ,  $I_j = L^*$ ,  $\sigma_H = 4$ ,  $\sigma_L = 1$ ,  $\rho = 3$ , and  $\alpha = 1/50$ ). Observe, their effectiveness in discerning between the object of interest (plant) and other background regions.

### 3.2. Plant Appearance Model

To safeguard our method from scene changes and increase its accuracy, we designed the algorithm such that it can utilize the segmentation outcomes of previous instances. Although, we will present in subsequent sections the details of our segmentation strategy, here we discuss that we can learn a plant model given the features of an image and its available segmentation into foreground and background regions. This model classifies pixels in an unseen instance according to their likelihood of belonging to a plant. Thus, with this process, we learn incrementally how a plant appears, and we feed back to the system this information.

We must consider several aspects to allow us to achieve this benefit without increasing complexity of design. Given a set of segmented images, there are several reasons to avoid training a sophisticated supervised classification algorithm that assigns pixels to either the foreground or the background class. First it has to allow for fast training to permit fast updates, second it has to be robust towards (possibly) mislabelled data which are automatically produced by previous time instances, and third it should have low data storage requirements. Classical supervised classification algorithms retain both foreground and background features and thus require additional storage and also their online (incremental learning) implementation is not straightforward. The solution we adopt here uses only foreground information and features in an appearance model.

Our solution is given an input image (and the features that can be extracted from it) and a segmentation mask identifying plants to learn the multi-dimensional distribution of the feature space using a multivariate Gaussian Mixture Model formulation. Accordingly, the density function for an observation (a pixel location with its features) with the  $d$ -dimensional feature vector  $x \in \mathbb{R}^d$  is given by:

$$p(x|\Theta) = \sum_{j=1}^M \pi_j p(x|\Theta_j), \quad (3)$$

where  $M > 1$  is the number of components of the mixture,  $\Theta_j = (\mu_j, \Sigma_j)$  is the set of parameters defining the  $j$ -th component of the mixture (i.e., mean  $\mu_j$  and covariance matrix  $\Sigma_j$ ), and  $\pi_j$  is the prior probability of pattern  $x$  belonging to the  $j$ -th component, such that  $0 \leq \pi_j \leq 1$ , for  $j = 1, \dots, M$ , and  $\sum_{j=1}^M \pi_j = 1$ . Each component of the mixture is characterized by a multivariate Gaussian distribution:

$$p(x|\Theta_j) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_j|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}(x - \mu_j)^\top \Sigma_j^{-1} (x - \mu_j)\right). \quad (4)$$

Given the density function, the log-likelihood function is defined as:

$$\ell(\Theta) = \sum_{i=1}^N \log \left( \sum_{j=1}^M \pi_j p(x_i|\Theta_j) \right), \quad (5)$$

where  $N$  is the number of available data points.

We maximize the log-likelihood function and estimate the unknown parameters of the distribution and the pixel's prior  $\Theta = (\pi_j, \mu_j, \Sigma_j)$ , for  $j = 1, \dots, M$ , using the Expectation-Maximization algorithm [43], to fit the GMM model to the available plant data (foreground). To eliminate the need to store the collection of all feature vectors  $x$  for all seen images, we update the GMM after each segmented image in an online fashion as in [44]. Thus, our storage requirements are limited to storing original images, their segmentation outputs, and relevant metadata (see Section 3.4).

Once the model is learned, for an unseen image we extract its features and obtain the probability of each pixel belonging to the plant model, by evaluating Eq. (3). The output of the plant model applied in a new image is a probability map  $P$  (with values in the interval  $[0, 1]$ ), which contains an estimate of the probability of any pixel in the image belonging to a plant.

### 3.3. Plant Localization

While some methods assume certain grid arrangement of pots [13, 14], one of our design criteria is not to impose conditions on the scene and the arrangement of the rosettes. This necessitates a process to localize plant objects in the scene. This step isolates rectangular regions of interest (of reduced size) containing plants, and estimates approximate plant segmentations which are then subsequently used as input to higher-complexity steps (e.g., the segmentation presented in the next section).

Although several approaches can be employed, for simplicity we use a  $K$ -means (with  $K = 2$ ) clustering algorithm operating in the feature space discussed in Section 3.1, to cluster a pixel as plant or background in the original image. To initialize the  $K$ -means we adopt the following schemes to get a good choice of initial cluster centroids. In the absence of a plant model or prior information we calculate initial centroids automatically from histogram thresholding in the Excess Green color space [45]. We transform the RGB to Excess Green ( $ExG$ ) domain, using  $ExG = 2G - R - B$ , where  $R$ ,  $G$  and  $B$  are the three components of the RGB color space. Subsequently, we use Otsu's thresholding [46] to identify a single threshold. Alternatively, in the presence of a plant appearance model we take advantage of it to get a good choice of initial cluster centroids. In both cases, we threshold the  $ExG$  or the probability map  $P$ , respectively, using a fixed threshold and consider all pixels above this threshold as foreground, while pixels below this threshold are considered as background. The initial cluster centroids are found by averaging foreground and background pixels respectively. We should note that we also allow the user to optionally calibrate and provide initial cluster centroids offline.

Subsequently, after the  $K$ -means has converged, we find a square bounding box that contains each plant object and output this region. A two cluster  $K$ -means performed on a complex image (i.e., plants surrounded by several other distinct objects) exhibits a bias towards over-segmentation; however, this behavior provides sufficient guarantees that all parts of a plant are included in its bounding box.

We also output a binary mask that serves as an approximate plant segmentation to be used as input to the active contour segmentation. Depending on the mode of operation, this binary mask can either be obtained from the  $K$ -means clustering or from the thresholded probability map.

The end result of this process is a collection of rectangular regions of interest, whose union reconstructs the original scene, and a collection of rough plant segmentations.

#### 3.3.1. A Probabilistic Vector Valued Active Contour Model

While the previous step provides a collection of rectangular regions of interest and an initial rough segmentation, the goal of this step is to obtain a highly accurate segmentation of each plant. Operating on a smaller portion of the image allows us to use more complex algorithms, which likely would not have been efficient and effective in the full image. Our motivation for employing an active contour method for object segmentation is its ability to model arbitrarily complex shapes and handle implicitly topological changes such as merging and splitting. Thus, the level set based segmentation method can be effectively used for extracting a foreground layer with fragmented appearance, such as leaves of the same plant.

Since we want to take advantage of the existence of multiple image features (in the following we refer to them as channels) we build upon the Chan-Sandberg-Vese active contour model for vector valued images [47]. The innovative aspects of our approach, compared to [47] and other similar works, are that: (a) we introduce the median of the foreground distribution in addition to its mean in the overall energy functional, and (b) we incorporate a probabilistic prior in our formulation.

Recently it was shown that using the median in region based level sets can adapt better to images where the object of interest maybe composed from different intensity classes [48]. Here we use the median as descriptor of non-symmetry in the distribution of the foreground for each channel thus, increasing the discriminative power between the foreground and background distributions with minimal computational overhead.

Without prior knowledge it is known that active contours may also erroneously segment regions that appear to have high statistical similarity with the object of interest. Chen et al. [49] used a non-parametric technique (namely Kernel Density Estimation) to model the shape variation, previously proposed by Cremers et al. [50] to incorporate both shape and intensity prior information. Leventon et al. [51] proposed to incorporate prior information about intensity and curvature profile of the structure using a training set of images and boundaries. They model the intensity distribution as a function of signed distances from the object boundary, rather than modeling only the intensity of the object as a whole. On the other hand, Lee et al. [52] proposed a supervised active contour model, which estimates a multivariate mixture density function from training samples using either parametric or non-parametric density estimation methods. This density is used to measure how likely each image pixel is to be an element of each subset in a new probabilistic active contour formulation. We rely on a new mechanism to incorporate prior information. We use the information provided by the appearance model to weigh pixels according to the probability of belonging to the foreground or background. Thus, we do not rely on shape (which has its own complications as we discussed previously) and we decouple the prior knowledge from the active contour model increasing the flexibility of our approach.

In this paper, the overall energy functional for the proposed model consists of two parts: an “Image-based Feature” term  $E^{\text{lbF}}$ , which utilizes the intensity information of each channel, and a “Prior” term  $E^P$  to incorporate the prior knowledge obtained based on the plant appearance model described previously. In the following we use  $\vec{\cdot}$  to denote vectors.

The  $E^{\text{lbF}}$  term following the formulation in [47] is defined as:

$$E^{\text{lbF}}(C, \vec{c}^+, \vec{m}^+, \vec{c}^-) = \int_{in(C)} \frac{1}{N} \sum_{i=1}^N \lambda_i^+ e_i^+(z) dz + \int_{out(C)} \frac{2}{N} \sum_{i=1}^N \lambda_i^- e_i^-(z) dz, \quad (6)$$

$$e_i^+(z) = |I_i - c_i^+|^2 + |I_i - m_i^+|^2, \quad (7)$$

$$e_i^-(z) = |I_i - c_i^-|^2, \quad (8)$$

where  $z$  denotes a pixel location in an image channel  $I_i$ ,  $i = 1, \dots, N$ ,  $\lambda_i^+$  and  $\lambda_i^-$  define the weight of each term (inside and outside the contour),  $\vec{c}^-$  is the vector valued representation of the mean for each channel outside the contour, and  $\vec{c}^+$  and  $\vec{m}^+$  are the vector-valued representations of the mean and median respectively for each channel inside the contour. Our key difference compared to [47] is the introduction of a term corresponding to the median. The way we estimate these statistical quantities will be described shortly.

Following standard level set formulations [47] we replace the contour curve  $C$  with the level set function  $\phi$  [53]:

$$E^{\text{lbF}}(\phi, \vec{c}^+, \vec{m}^+, \vec{c}^-) = \int_{\phi \geq 0} \frac{1}{N} \sum_{i=1}^N \lambda_i^+ e_i^+(z) dz + \int_{\phi < 0} \frac{2}{N} \sum_{i=1}^N \lambda_i^- e_i^-(z) dz. \quad (9)$$

The vectors  $\vec{c}^+$ ,  $\vec{m}^+$ , and  $\vec{c}^-$  are defined in similar fashion to other intensity driven active contour models as statistical averages and medians:

$$\begin{cases} \vec{c}^+(\phi) = \text{average}(I_i \in \phi(z) \geq 0), \\ \vec{m}^+(\phi) = \text{median}(I_i \in \phi(z) \geq 0), \\ \vec{c}^-(\phi) = \text{average}(I_i \in \phi(z) < 0), \end{cases} \quad (10)$$

for each channel  $I_i$ ,  $i = 1, \dots, N$ , inside or outside the contour.

Using the level set function  $\phi$  to represent the contour  $C$  in the image domain  $\Omega$ , the energy functional can be written as follows:

$$E^{\text{lbF}}(\phi, \vec{c}^+, \vec{m}^+, \vec{c}^-) = \int_{\Omega} \frac{1}{N} \sum_{i=1}^N \lambda_i^+ e_i^+(z) H(\phi(z)) dz + \int_{\Omega} \frac{2}{N} \sum_{i=1}^N \lambda_i^- e_i^-(z) (1 - H(\phi(z))) dz, \quad (11)$$

where  $H$  is the Heaviside function.

By keeping  $\vec{c}^+$ ,  $\vec{m}^+$ , and  $\vec{c}^-$  fixed, we minimize the energy function  $E^{\text{IbF}}(\phi, \vec{c}^+, \vec{m}^+, \vec{c}^-)$  with respect to  $\phi$  to obtain the gradient descent flow as:

$$\frac{\partial \phi}{\partial t} = \zeta^{\text{IbF}} = \delta(\phi) \left[ -\frac{1}{N} \sum_{i=1}^N \lambda_i^+ e_i^+(z) + \frac{2}{N} \sum_{i=1}^N \lambda_i^- e_i^-(z) \right], \quad (12)$$

where  $\delta$  is the Dirac delta function.

After Eq. (12) converges, the evolving curve  $C$  will separate the object from the background based on the non-symmetric property of foreground distribution. However, when the background and foreground are not easily separable without prior knowledge the level set may converge to a wrong boundary.

To introduce prior knowledge we require matrices  $P_{\text{in}}$ ,  $P_{\text{out}}$  (the size of the image) where  $P_{\text{in}}(z) \equiv p(z \in \Omega_1)$ , i.e., the probability of pixel at location  $z$  belonging to the foreground class  $\Omega_1$ , and naturally  $P_{\text{out}}(z) \equiv p(z \in \Omega_2)$ , i.e., the probability of pixel at location  $z$  belonging to the background class  $\Omega_2$ . Notice that  $P_{\text{in}}(z) + P_{\text{out}}(z) = 1$ .

In the proposed active contour formulation we utilize both  $P_{\text{in}}$  and  $P_{\text{out}}$  to weigh each channel individually, hence the prior energy term can be described as follows:

$$E^P(\phi, \vec{c}^\oplus, \vec{m}^\oplus, \vec{c}^\ominus) = \int_{\Omega} \frac{1}{N} \sum_{i=1}^N \lambda_i^\oplus e_i^\oplus(z) dz + \int_{\Omega} \frac{2}{N} \sum_{i=1}^N \lambda_i^\ominus e_i^\ominus(z) dz, \quad (13)$$

$$e_i^\oplus(z) = |I_i \cdot P_{\text{in}} - c_i^\oplus|^2 + |I_i \cdot P_{\text{in}} - m_i^\oplus|^2, \quad (14)$$

$$e_i^\ominus(z) = |I_i \cdot P_{\text{out}} - c_i^\ominus|^2, \quad (15)$$

where now  $\lambda_i^\oplus$  and  $\lambda_i^\ominus$  define the weights of each term,  $\cdot$  denotes point-wise multiplication,  $\vec{c}^\oplus$ ,  $\vec{m}^\oplus$ , and  $\vec{c}^\ominus$  are defined as follows:

$$\begin{cases} \vec{c}^\oplus = \text{average}(I_i \cdot P_{\text{in}} \in \phi(z) \geq 0), \\ \vec{m}^\oplus = \text{median}(I_i \cdot P_{\text{in}} \in \phi(z) \geq 0), \\ \vec{c}^\ominus = \text{average}(I_i \cdot P_{\text{out}} \in \phi(z) < 0), \end{cases} \quad (16)$$

for  $i = 1, \dots, N$ .

The level set formula based on the prior term is defined as follows:

$$\frac{\partial \phi}{\partial t} = \zeta^P = \delta(\phi) \left[ -\frac{1}{N} \sum_{i=1}^N \lambda_i^\oplus e_i^\oplus(z) + \frac{2}{N} \sum_{i=1}^N \lambda_i^\ominus e_i^\ominus(z) \right]. \quad (17)$$

To derive our final joint level set functional form, we follow the approach in [47] and replace  $\delta(\phi)$  in Eqs. (12) and (17) by  $|\nabla \phi|$ . The former has a small effective range while the latter has the effective range of the whole image. Furthermore, in order to efficiently regularize the level set evolution, we convolve the level set function with a Gaussian kernel [54]. Finally, the level set formulation of our model combining Eqs. (12) and (17) becomes:

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| [(1 - \lambda) \zeta^{\text{IbF}} + \lambda \zeta^P], \quad (18)$$

where  $\lambda$  balances the influence of  $E^{\text{IbF}}$  and  $E^P$ . A larger value of  $\lambda$  emphasizes the effect of the  $E^P$  term versus the  $E^{\text{IbF}}$  term, whereas a smaller  $\lambda$  reduces the effect of the prior energy term. In particular,  $\lambda = 0$  implies that the model is utilizing image-based features only, without relying on prior knowledge.

### 3.3.2. Implementation

The main steps of our segmentation model can be summarized as follows. For each region of interest obtained from the localization step:

1. initialize the level set function  $\phi$  to be binary;
2. evolve the level set function according to Eq. (18);

3. smooth the level set function with a Gaussian kernel;
4. if the curve evolution has converged, output a binary mask of the segmentation, otherwise return to step 1.

While the previous steps are general, we outline below details of our implementation for this particular application. We use as  $I_1, I_2$  respectively the  $a^*$  and  $b^*$  components of the  $L^*a^*b^*$  color space, which is regarded as representing global information. Texture features are encoded in the channels  $I_3, \dots, I_N$  as local information since they describe local neighborhood structure [55]. Furthermore, since the prior knowledge used here is learned only from foreground objects we use as  $P_{\text{in}}$  the output of the plant appearance model, and as  $P_{\text{out}} = 1 - P_{\text{in}}$ . To accommodate that the prior knowledge of the background is less reliable (since we did not train on the background) we use a smaller weight  $\lambda_i^\ominus$ . Alternatively, if we did not want to provide any prior knowledge for the background we could model  $P_{\text{out}}$  to follow a uniform distribution as suggested in [52]. Finally, to initialize the level set we use the output of plant localization (see Section 3.3).

### 3.4. Plant Labeling and Analysis

After we have obtained all segmentations of foreground from the active contour model, we recompose the original scene (containing all plants), and we obtain a binary representation of all plant objects in the original scene. In phenotyping experiments, plants usually belong to distinct groups, e.g., mutants of the same species or specimens undergoing different treatments. Therefore, plants need to be labeled and followed individually across time in order to maintain correspondence of individual measurements for each plant. The goal of this step is to assign a unique label to pixels of the same plant (intra-frame accuracy) and to keep assigning that label to the same plant across time (inter-frame accuracy).

The binary mask representing foreground objects is composed by a number of connected components. Due to possible errors in the segmentation (e.g., under-segmentation of the stems due to lack of resolution), portions of the same plant may result in disconnected objects (e.g., a leaf may not be connected to its originating plant).

To address this problem we take advantage of the radial shape of *Arabidopsis* rosettes, and label disconnected objects in groups that minimize the Euclidean distance from a centroid. If  $N_{\text{subj}}$  is the number of plants in the scene (recall from Section 3 that this parameter is fixed by the user), we perform the task of intra-frame labeling by finding the centroids of the  $N_{\text{subj}}$  largest connected components. These are then used to initialize a  $K$ -means clustering on the pixel coordinates, with  $K = N_{\text{subj}}$ . Thus, each pixel in the foreground is labeled according to which centroid is closer in the Euclidean distance sense. To obtain the final label for all pixels in a given component, a connected component having different labels (e.g., a leaf of one plant close to another plant) is assigned a single label with a majority vote strategy.

Having available plant centroids and labels from the previous image, coherency in inter-frame labeling is maintained by assigning a plant in the current image the label of the closest plant (in the Euclidean sense) in the previous image. This approach tolerates small shifts of pots (which can occur when staff are handling the experiment or when plants are watered), but not significant movements or shuffling. In our setting we assume that plants do not touch and it is the user's responsibility to arrange pots in the scene with enough distance between each other (an assumption common among many phenotyping analysis platforms). To accommodate touching plants a plant shape model is necessary, which as we outlined in Section 2.2 can be rather complex and specific to a particular plant species.

After a successful segmentation, several visual phenotypes can be extracted. For example, plant growth is estimated with several indices [14] that reflect the area of the plant, its roundness, and overall color intensity. The measurements, the indices, along with the plant identifier are written in a tabulated format that can be imported by several statistical analysis and plotting software.

## 4. Results and Discussion

### 4.1. Experimental Setup

We implemented our system in Matlab (release 2011b), on a machine equipped with Intel Core 2 Duo CPU E8200 2.66 GHz and 4 GB memory, running 64-bit Linux. Due to lack of publicly available datasets we devised a test bed to image *Arabidopsis thaliana* in our laboratory. Since our goal was to produce a good range of test images several challenging situations, such as water and moss growth were allowed to occur. The scene consists of top-view images of  $N_{\text{subj}} = 19$  *Arabidopsis thaliana* Columbia (Col-0) wild-type rosettes, acquired over a period of 12 days. The plants

were imaged with a 7 megapixel commercial camera (Canon PowerShot SD1000) following the setup discussed in [13]. The images were stored and processed in raw format to avoid any distortion introduced by compression. Figure 5 shows an example image from the dataset, illustrating the arrangement of the plants and the complexity of the scene.

We segmented manually all plants to obtain ground-truth segmentations. To quantify the accuracy of the segmentation algorithms, we adopt the following metrics:

$$\text{Precision (\%)} = \frac{TP}{TP + FP} \quad (19)$$

$$\text{Recall (\%)} = \frac{TP}{TP + FN} \quad (20)$$

$$\text{Jaccard (\%)} = \frac{TP}{FP + TP + FN} \quad (21)$$

$$\text{Dice (\%)} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (22)$$

where  $TP$ ,  $FP$ , and  $FN$  represent the number of true positive, false positive, and false negative pixels, respectively, calculated by comparing algorithmic result and ground-truth segmentation masks. Precision is the fraction of pixels in the segmentation mask that matches the ground truth, whereas recall is the fraction of ground-truth pixels contained in the segmentation mask. The Jaccard and Dice similarity coefficients are used to measure the spatial overlap between algorithmic result and ground truth. All of these metrics are expressed in percentages, with larger values representing higher agreement between ground truth and algorithmic result.

We compare the proposed approach with state-of-the-art methods in plant phenotyping and recent image segmentation approaches. The first plant phenotyping method (referred thereafter as Reference) is an approach based on  $K$ -means segmentation, due to its widespread adoption in intensity-based plant segmentation. For fair comparison, the  $K$ -means was applied on the same intensity and texture features as the proposed approach, and we carefully selected initial centroids for each image in the dataset. We used the same plant labeling and analysis procedure as the one proposed here (Section 3.4). We also adopted the Rosette Tracker software proposed in [14], which is made available as an ImageJ plugin by the authors. It was operated by performing proper color calibration and by enabling the options for removing moss and clutter, according to the characteristics of our dataset. We gave as input to Rosette Tracker the number of plants in the scene (as required by the software). However since Rosette Tracker assumes that plants are arranged in rows and the imaging axis is parallel to these rows, plants were often mislabeled. This required manual post-processing to correctly assign labels to plant parts, in order to include Rosette Tracker in the evaluation.

To demonstrate that even state-of-the-art methods in color image segmentation are challenged, we chose four algorithms covering a span of recent innovations in computer vision and used their reference software implementations. In particular, we considered the following methods: *gPb-owt-ucm* [56], a segmentation method that relies on contour detection and spectral clustering, providing a hierarchical representation of the image and a final segmentation after user annotation; *MSRM* [57], an interactive segmentation approach that relies on a new region merging framework to fuse a super-pixel segmentation obtained using [58]; the *CoSand* [59] method which relies on the temperature maximization of the anisotropic heat diffusion formulation on a graph representation of the image (an initial super-pixel segmentation is obtained using [58]); and finally, although it is not a true segmentation method, we also used *SDSP* [60], a saliency detection algorithm, because it uses similar features as the proposed approach: it combines priors related to frequency (implemented using log-Gabor features), color (CIE  $L^*a^*b^*$  color space), and location of the object. For all methods standard parameters as recommended by their respective authors were used; and for the methods relying on interaction, unbiased and automated foreground and background annotations were obtained by skeletonizing and dilating the ground truth mask, as previously done in [61].

Since *Arabidopsis* plants do not have very pronounced veins and given the imaging resolution of our setup, for the results that follow we chose the following parameters (which were kept constant throughout the experiment): we used  $M = 2$  mixtures for the GMM implementing the plant appearance model, and only one TFB texture image, obtained using  $\sigma_H = 4$ ,  $\sigma_L = 1$ , and  $\rho = 3$ . Thus, we used 3 image features including the  $a^*$  and  $b^*$  color components and the TFB texture feature. To initialize the contour we used the probability map, by applying a fixed threshold (0.5). The parameters of the active contour formulation were set as follows:  $\lambda = 0.6$  (i.e., we rely almost equally on each of the

two terms of Eq. (18)),  $\lambda_i^+ = \lambda_i^- = 1, \forall i$  (in other words each channel  $I_i$  in Eq. (6) has the same weight), and since we learn only on the foreground, by choosing  $\lambda_i^\oplus = 1$  (i.e., the channel weights of the foreground), and  $\lambda_i^\ominus = 0.01$  (i.e., the weights of the background channels), we rely mostly on the foreground prior to drive the curve evolution. We recall that parameters  $\lambda_i^+, \lambda_i^-$  refer to the image-based feature term, while  $\lambda_i^\oplus, \lambda_i^\ominus$  refer to the prior term in the energy functional. We should note that no morphological operations have been performed on the segmentation output of any method, thus the results presented in the following reflect the true output of the algorithms used.

#### 4.2. Results

In this section we present first the results obtained by evaluating the proposed system on the dataset previously described, in terms of both segmentation accuracy and validity as a tool for plant phenotyping applications. The performance of the proposed system is compared with the Reference method and Rosette Tracker. Then, we also compare the proposed approach with more recent segmentation methods [56, 57, 59, 60]. Finally, to illustrate the novelty of our active contour formulation and the importance of using prior knowledge and texture features, we describe and compare variants of the proposed system.

To illustrate the output of our algorithm Figure 5 shows an example image from the dataset. For visual comparison we also include the ground truth segmentation and labeling and include outputs from the Reference and Rosette Tracker approaches. Overall, visually a significant agreement can be observed between the Proposed method and ground truth, whereas Reference appears noisy (e.g., moss and earth in the pots are also included in the foreground). On the other hand, Rosette Tracker reports plant objects in regions belonging to the background, likely attributed to the fact that it does not include a plant localization step. We should note that in our quantitative analysis that follows we manually post-process the results of Rosette Tracker to assign labels correctly to foreground objects and to eliminate wrongly detected regions.

To better appreciate visually the differences in segmentation accuracy and illustrate the complexity introduced as plants grow or as other objects appear in the background, Figure 6 shows segmentation outcomes for a single plant observed across time. It is evident that Reference has a clear bias towards over-segmentation (in other words it cannot separate moss from plant), while Rosette Tracker deals with moss by being conservative, thus under-segmenting, but without completely eliminating moss from the segmentation. Overall, Rosette Tracker shows substantial loss of leaf and stem portions (and thus it will affect growth trend measurements and other phenotyping markers). On the other hand, the Proposed is unaffected by moss or any other noise in the background, while still preserving the entirety of the plant structure with smooth boundaries.

While the previous examples demonstrated visually the accuracy of our method, Figure 7 shows segmentation accuracy over time, measured quantitatively using the Dice Similarity Coefficient. It is readily seen, that the performance of Reference starts decreasing very soon as plant (and moss) grow and the scene's complexity increases. After the seventh day its performance seems to improve, but this outcome is misleading. It occurs only because as plants grow they cover most of the pot and consequently most of the moss. Rosette Tracker as well is not robust to changes in the scene. It exhibits an oscillating level of accuracy, depending on events that occur in the scene, e.g., some plants experienced drought during the first days, moss incrementally appeared below some plants, the trays got shifted around the sixth day, and water was present in the trays while acquiring the last image of the dataset (a cut out of which is shown in the bottom of Figure 2). Such challenging images were included in the dataset to investigate the behavior of the algorithms under different conditions than the ideal case, that are still likely to occur in practice. The proposed system adapts to these changes and accurately segments the plant. Thanks to the plant appearance model that is learned over time and is integrated in the active contour segmentation, the proposed system responds to the aforementioned challenges appearing in the scene, significantly better than the other methods; it maintains a very high accuracy (above 90%) throughout the whole cycle.

From a phenotyping perspective, segmentation accuracy is important, because it reflects the ability of the system to test phenotyping hypotheses. Figure 8 shows the growth pattern of a plant, comparing its estimate obtained by automatic segmentations with the ground truth. Growth was estimated from the projected rosette area and the reported results were normalized by plant area at the beginning of the experiment. The actual growth of the plant appears steady over time and it follows a linear trend. The growth pattern is characterized by periodic saddles, which reflect the natural circadian rhythm of the plant, which is regulated by the light / dark cycles (ambient temperature did not deviate much in our climate controlled office). During the first two days all of the systems provided satisfactory results (i.e., they follow the ground truth close); however, as the conditions in the scene changed, different outcomes can be

Table 1: Segmentation accuracy reported for Proposed, Rosette Tracker, and Reference methods, shown as mean (standard deviation).

System	Accuracy (%)			
	Precision	Recall	Jaccard	Dice
<b>Proposed</b>	97.08 (1.83)	95.86 (2.96)	93.17 (3.22)	96.44 (1.76)
Rosette Tracker	88.86 (6.49)	78.83 (24.37)	71.20 (22.29)	80.37 (22.57)
Reference	60.82 (14.55)	99.87 (0.21)	60.74 (14.43)	74.65 (10.62)

observed that are not natural and would jeopardize subsequent statistical analysis. Reference reported an unnatural exponential growth of the plant after the second day, due to moss in the pots that was erroneously segmented along with plants.

On the other hand, while Rosette Tracker is less susceptible to moss compared to Reference (see Figure 6), still there is a tendency to overestimate plant area. After the sixth day Rosette Tracker exhibits an oscillating behavior that over-amplifies any changes in the scene. In particular, the last day, a layer of water in the tray caused Rosette Tracker to severely under-segment the plants.

Most of the approaches for image-based plant phenotyping (as it is the case for Reference and Rosette Tracker) can operate appropriately only under strict assumptions on the composition and stability of the scene. When such assumptions are violated, one must discard these images (or the results) which hurts the validity of the phenotyping experiment (with consequences on the statistical power of the experiment). As it is shown in Figure 8, our proposed system closely approximates the actual growth pattern of the plant, proving robust to any changing and challenging conditions in the dataset (e.g., moss, tray shifts, water). The fidelity in recovering the growth pattern is brought to such an extent that the circadian rhythm of the plant can be readily observed.

This level of accuracy is observed across all our dataset and using additional accuracy metrics. Table 1 reports averaged results of segmentation accuracy over the whole dataset. The Reference method shows poor accuracy in terms of precision, Jaccard and Dice, and a very high recall value due to the constant over-segmentation (i.e., the plant is fully contained in the segmentation mask, along with large portions of earth and moss from the background). Rosette Tracker appears more balanced in the overall results, although it leaves substantial room for improvement. Our proposed system achieves very high accuracy values (above 90%) for all of the employed metrics.

Such capability of accurately delineating plant objects in images enables researchers to test phenotyping hypothesis in experiments with subjects from different species or undergoing different treatments, using automated phenotype collection solutions. Primarily, our approach is a tool to study phenotypes related to size, shape, and growth pattern; however, also quantification of any other visual phenotypes observable through digital images (e.g., color variations, timing of flowering) assumes a segmentation of the plant from the background, which our software provides.

While previously we compared our approach with state-of-the-art methods in plant phenotyping, it is critical to showcase the challenge posed by the problem compared to recent innovations in the field of image segmentation (as outlined in the experimental setup) [56, 57, 59, 60]. Figure 9 shows the results of this comparison. Observe that, while on the first image all algorithms produce comparable results, as the images become more challenging (moss or water presence) the segmentation results deteriorate, while our method demonstrates greater accuracy and closely estimates the ground truth. This behavior was observed across all images in the dataset. Using the same performance metrics (shown as average and standard deviation in parenthesis) as previously, the best performing method, *gPbowt-ucm* [56], had Precision 89.04(6.40)%, Recall 96.20(2.78)%, Jaccard 85.86(5.49)%, and Dice 92.29(3.36)%. This method requires human interaction and our testing showed that its final result depends largely on the detail and precision of such user annotation. Thus, to eliminate any user bias all interactive methods were initialized using morphological operations from the ground truth masks. This provides highly accurate annotations and even with this ideal annotations the interactive methods are unable to match the accuracy of the proposed approach (e.g., the Dice similarity coefficient for our Proposed approach was 96.44%, according to Table 1), illustrating the challenges of segmenting plants in complicated background. Finally, most of these methods are computationally demanding both in terms of system memory and time, largely due to the construction of a graph representation.

The proposed approach shows superiority when compared to other methods but is important from an image processing viewpoint to identify which component of the approach contributes to this gain in accuracy. We designed

Table 2: Variants of the proposed system introduced to show the contribution of each component.

System	Learning	Features	Median	Comments
<b>Proposed</b>	2-GMM	$a^*, b^*, TFB$	Yes	$\lambda = 0.6, \lambda_i^+ = \lambda_i^- = \lambda_i^\oplus = 1, \text{ and } \lambda_i^\ominus = 10^{-2} \forall i$
Proposed-a	No	$a^*, b^*, TFB$	Yes	$\lambda = 0, \text{ others same as Proposed}$
Proposed-b	No	$a^*, b^*, TFB$	No	$\vec{m}^+ = \vec{c}^+, \text{ and } \vec{m}^\oplus = \vec{c}^\oplus, \text{ others same as Proposed-a}$
Proposed-c	No	$a^*, b^*$	No	$\lambda_3^+ = 0, \text{ and } \lambda_3^- = 0, \text{ others same as Proposed-b}$

Table 3: Effects of different components (detailed in Table 2) on segmentation accuracy shown as mean (standard deviation).

System	Accuracy (%)			
	Precision	Recall	Jaccard	Dice
<b>Proposed</b>	97.08 (1.83)	95.86 (2.96)	93.17 (3.22)	96.44 (1.76)
Proposed-a	85.10 (9.37)	98.83 (1.19)	84.21 (9.08)	91.17 (5.54)
Proposed-b	83.25 (10.29)	98.99 (1.10)	82.50 (9.98)	90.08 (6.21)
Proposed-c	73.72 (17.82)	99.39 (0.92)	73.21 (17.28)	83.39 (11.92)

a series of experiments to highlight the importance of using texture, prior knowledge and the median (in modeling the distribution of the foreground inside the contour). These scenarios can be easily tested by altering weights in the level set formulation. To this end, we use different values for the controlling parameter  $\lambda$  of  $E^{lbF}$  and  $E^P$  terms and the internal weight parameters  $\lambda_i^+, \lambda_i^-, \lambda_i^\oplus$  and  $\lambda_i^\ominus$ . Table 2 summarizes conceptually the different versions used for comparison, while the last column includes the change in parameters. As shown in Table 2, the Proposed model refers to our proposed approach, Proposed-a is a version without considering the prior knowledge, while Proposed-b refers to a version without prior knowledge and without the median for the foreground distribution. Finally, Proposed-c is a model version where the prior knowledge, median, and texture components are disabled, essentially resembling the model in [47]. All versions shared the same initialization of the active contour model, in order to highlight the individual contribution of each component.

To demonstrate the effect of each component qualitatively, Figure 10 shows the segmentation result of a plant image with the different versions considered here (all of them share the same contour initialization). Evidently the proposed method has the best agreement with the ground truth. Clearly the median and the texture contribute to the accuracy, incrementally but never reach the agreement of the Proposed that includes learning. Proposed-c on the other hand which relies only on color information completely over-segments.

The same quantitative conclusions can be reached also when we compare the accuracy of these modified systems using accuracy metrics standard in image segmentation. As Table 3 shows the Proposed model gives a substantial performance improvement across the whole dataset. Compared to the other versions, this one is not affected by the complexity of the background distribution when prior knowledge is included. The difference in performance between Proposed-b and Proposed-c highlights the importance of considering the texture component within the level set framework as local region-based information; it can differentiate between foreground and background distributions better. The effect of adding the median descriptor in the level set energy functional for the minimization is also important because it increases the accuracy (e.g., Dice is higher than Proposed-b) but also reduces the measurement deviation across the dataset. This is in agreement with the findings of [48] related to the robustness to the skewness of the distribution which we extended here to vector valued formulation.

## 5. Conclusions

We propose a novel approach for the segmentation of plants in image based phenotyping experiments. We propose a new vector valued active contour model which incorporates prior knowledge reflecting the likelihood of a pixel to belong to a plant. We build a plant appearance model based on Gaussian Mixture Models and train this model based on the output of the segmentations. Thus, we use prior instances in an incremental learning fashion to accommodate

changes in scenes and complexity of the background. We rely on color and texture features, but we aim to balance complexity of the approach (i.e., storage and processing requirements) with scalability to larger experiments.

We tested our proposed system with several top-view images of *Arabidopsis* collected using a time-lapse digital camera in our laboratory over a span of a few days. For comparison we implemented a reference unsupervised segmentation method, commonly employed in image based plant phenotyping, and used another method publicly available [14]. We also compared with several recent innovations in image segmentation [56, 57, 59, 60]. For validation we manually labeled each image in our dataset. The proposed approach achieves an overall accuracy of 96.44% (Dice Similarity Coefficient), which is significantly higher than the other two methods. Our experiments also show that the proposed approach can accurately estimate plant growth and is not susceptible to changes in the scene. Furthermore, we show that it is the integration of prior knowledge, texture features, and a new level set formulation that achieves this increase in accuracy.

Most of the currently available solutions for analysis of images for plant phenotyping are tailored to specific acquisition scenarios. As a result the proposed segmentation approaches cannot be generalized to any laboratory environment. Our proposed solution involves minimal interaction and employs simple yet effective machine learning techniques to learn from the output (and possible feedback of the user). We use an appearance model that can accommodate several plant species since it does not require shape information.

While some approaches are free and open source (e.g., [14, 23]) the current state of the art in phenotyping analysis are commercial software that accompany professional (and costly) phenotyping infrastructures. Our proposed solution is supplemented by a graphical user interface to facilitate interaction with the user. This will be soon publicly available (a release of our implementation in Matlab is expected by late 2013).

Currently we tested our software using only *Arabidopsis* images from a single mutant family. Testing on additional mutants and other rosette plant models is left for future work. While our approach assumes that little motion occurs between images (i.e., no shuffling) in the future we will consider to incorporate several safety checks to alert the user of possible errors. While here we used a single top-view image it is possible that acquiring images at different depth and focus, and their later fusion to a single fully focused image, could further increase the accuracy of our approach. To increase adoption (and relieve the need for Matlab licenses) in the future we will convert our algorithms to use open source alternatives such as Python. Finally, for users that prefer hosted solutions, we intend to incorporate our solution within the PhytoBisque service of the iPlant Collaborative framework [9].

## Acknowledgments

This work was supported in part by a Marie Curie Action: “Reintegration Grant” (ref# 256534) of the EU’s Seventh Framework Programme. The authors would like to thank Prof. Pierdomenico Perata and his group from Scuola Superiore Sant’Anna, for providing us with plant samples and instructions on growth conditions of *Arabidopsis*. They also thank Dr. Cristian Rusu, for valuable feedback on the manuscript. Finally, they thank the reviewers for providing detailed comments and constructive critique on the manuscript.

## References

- [1] J. D. Peleman, J. R. R. van der Voort, Breeding by Design, *Trends in Plant Science* 8 (7) (2003) 330–334.
- [2] F. Licausi, M. Kosmacz, D. A. Weits, B. Giuntoli, F. M. Giorgi, L. A. C. J. Voesenek, P. Perata, J. T. van Dongen, Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization, *Nature* 479 (7373) (2011) 419–422.
- [3] R. C. O’Malley, J. R. Ecker, Linking genotype to phenotype using the *Arabidopsis* unimutant collection, *The Plant Journal* 61 (6) (2010) 928–940.
- [4] H. Holtorf, M.-C. C. Guittion, R. Reski, Plant functional genomics, *Die Naturwissenschaften* 89 (6) (2002) 235–249.
- [5] P. N. Benfey, T. Mitchell-Olds, From Genotype to Phenotype: Systems Biology Meets Natural Variation, *Science* 320 (5875) (2008) 495–497.
- [6] E. R. Mardis, Next-generation DNA sequencing methods, *Annual review of genomics and human genetics* 9 (1) (2008) 387–402.
- [7] R. T. Furbank, M. Tester, Phenomics – technologies to relieve the phenotyping bottleneck, *Trends in Plant Science* 16 (12) (2011) 635–644.
- [8] E. P. Spalding, N. D. Miller, Image analysis is driving a renaissance in growth measurement, *Current Opinion in Plant Biology* 16 (1) (2013) 100–104.
- [9] S. A. Goff, M. Vaughn, S. McKay, E. Lyons, A. E. Stapleton, D. Gessler, N. Matasci, L. Wang, M. Hanlon, A. Lenards, A. Muir, N. Merchant, S. Lowry, S. Mock, M. Helmke, A. Kubach, M. Narro, N. Hopkins, D. Micklos, U. Hilgert, M. Gonzales, C. Jordan, E. Skidmore, R. Dooley, J. Cazes, R. McLay, Z. Lu, S. Pasternak, L. Koesterke, W. H. Piel, R. Grene, C. Noutsos, K. Gendler, X. Feng, C. Tang, M. Lent, S.-J. Kim, K. Kvilekval, B. S. Manjunath, V. Tannen, A. Stamatakis, M. Sanderson, S. M. Welch, K. A. Cranston, P. Soltis, D. Soltis, B. O’Meara, C. Ane, T. Brutnell, D. J. Kleibenstein, J. W. White, J. Leebens-Mack, M. J. Donoghue, E. P. Spalding, T. J. Vision, C. R. Myers, D. Lowenthal, B. J.

- Enquist, B. Boyle, A. Akoglu, G. Andrews, S. Ram, D. Ware, L. Stein, D. Stanzione, The iPlant Collaborative: Cyberinfrastructure for Plant Biology, *Frontiers in Plant Science* 2 (34).
- [10] E. Finkel, With 'Phenomics,' Plant Scientists Hope to Shift Breeding Into Overdrive, *Science* 325 (5939) (2009) 380–381.
- [11] J. L. Micol, Leaf development: time to turn over a new leaf?, *Current Opinion in Plant Biology* 12 (1) (2009) 9–16.
- [12] M. Golzarian, R. Frick, K. Rajendran, B. Berger, S. Roy, M. Tester, D. Lun, Accurate inference of shoot biomass from high-throughput images of cereal plants, *Plant Methods* 7 (1) (2011) 2+.
- [13] S. A. Tsaftaris, C. Noutsos, Plant Phenotyping with Low Cost Digital Cameras and Image Analytics, in: *Information Technologies in Environmental Engineering, Environmental Science and Engineering*, Springer Berlin Heidelberg, 2009, pp. 238–251.
- [14] J. De Vylder, F. Vandebussche, Y. Hu, W. Philips, D. Van Der Straeten, Rosette Tracker: An Open Source Image Analysis Tool for Automatic Quantification of Genotype Effects, *Plant Physiology* 160 (3) (2012) 1149–1159.
- [15] J. W. White, P. Andrade-Sanchez, M. A. Gore, K. F. Bronson, T. A. Coffelt, M. M. Conley, K. A. Feldmann, A. N. French, J. T. Heun, D. J. Hunsaker, M. A. Jenks, B. A. Kimball, R. L. Roth, R. J. Strand, K. R. Thorp, G. W. Wall, G. Wang, Field-based phenomics for plant genetics research, *Field Crops Research* 133 (2012) 101–112.
- [16] A. Walter, B. Studer, R. Kölliker, Advanced phenotyping offers opportunities for improved breeding of forage and turf species, *Annals of Botany* 110 (6) (2012) 1271–1279.
- [17] D. Leister, C. Varotto, P. Pesaresi, A. Niwergall, F. Salamini, Large-scale evaluation of plant growth in *Arabidopsis thaliana* by non-invasive image analysis, *Plant Physiology and Biochemistry* 37 (9) (1999) 671–678.
- [18] C. Granier, L. Aguirrezzabal, K. Chenu, S. J. Cookson, M. Dauzat, P. Hamard, J.-J. Thioux, G. Rolland, S. Bouchier-Combaud, A. Lebaudy, B. Muller, T. Simonneau, F. Tardieu, PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in *Arabidopsis thaliana* permitted the identification of an accession with low sensitivity to soil water deficit, *New Phytologist* 169 (3) (2006) 623–635.
- [19] G. A. Pereyra-Irujo, E. D. Gasco, L. S. Peirone, L. A. N. Aguirrezzábal, GlyPh: a low-cost platform for phenotyping plant growth and water use, *Functional Plant Biology* 39 (11) (2012) 905–913.
- [20] S. Arvidsson, P. Pérez-Rodríguez, B. Mueller-Roeber, A growth phenotyping pipeline for *Arabidopsis thaliana* integrating image analysis and rosette area modeling for robust quantification of genotype effects, *The New Phytologist* 191 (3) (2011) 895–907.
- [21] A. Walter, H. Scharr, F. Gilmer, R. Zierer, K. A. Nagel, M. Ernst, A. Wiese, O. Virnich, M. M. Christ, B. Uhlig, S. Jünger, U. Schurr, Dynamics of seedling growth acclimation towards altered light conditions can be quantified via GROWSCREEN: a setup and procedure designed for rapid optical phenotyping of different plant species, *New Phytologist* 174 (2) (2007) 447–455.
- [22] J. Kokorian, G. Polder, J. J. B. Keurentjes, D. Vreugdenhil, M. Olortegui Guzman, An ImageJ based measurement setup for automated phenotyping of plants, in: *Proceedings of the ImageJ User and Developer Conference*, 2010, pp. 178–182.
- [23] A. Hartmann, T. Czauderna, R. Hoffmann, N. Stein, F. Schreiber, HTPheno: An image analysis pipeline for high-throughput plant phenotyping, *BMC Bioinformatics* 12 (1) (2011) 148+.
- [24] G. van der Heijden, Y. Song, G. Horgan, G. Polder, A. Dieleman, M. Bink, A. Palloix, F. van Eeuwijk, C. Glasbey, SPICY: towards automated phenotyping of large pepper plants in the greenhouse, *Functional Plant Biology* 39 (11) (2012) 870–877.
- [25] G. Alenyà, B. Dellen, C. Torras, 3D modelling of leaves from color and ToF data for robotized plant measuring, in: *International Conference on Robotics and Automation (ICRA)*, IEEE, 2011, pp. 3408–3414.
- [26] Y. Chéné, D. Rousseau, P. Lucidarme, J. Bertheloot, V. Caffier, P. Morel, E. Belin, F. Chapeau-Blondeau, On the use of depth camera for 3D phenotyping of entire plants, *Computers and Electronics in Agriculture* 82 (2012) 122–127.
- [27] X. Tang, M. Liu, H. Zhao, W. Tao, Leaf Extraction from Complicated Background, in: *2nd International Congress on Image and Signal Processing (CISP)*, IEEE, 2009, pp. 1–5.
- [28] G. Cerutti, L. Tougne, A. Vacavant, D. Coquin, A Parametric Active Polygon for Leaf Segmentation and Shape Estimation, in: G. Bebis, R. Boyle, B. Parvin, D. Koracin, S. Wang, K. Kyungnam, B. Benes, K. Moreland, C. Borst, S. DiVerdi, C. Yi-Jen, J. Ming (Eds.), *Advances in Visual Computing*, Vol. 6938 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2011, pp. 202–213.
- [29] J. De Vylder, D. Ochoa, W. Philips, L. Chaerle, D. Van Der Straeten, Leaf Segmentation and Tracking Using Probabilistic Parametric Active Contours, in: *Proceedings of the 5th International Conference on Computer Vision/Computer Graphics Collaboration Techniques, MIRAGE 2011*, Springer-Verlag, Berlin, Heidelberg, 2011, pp. 75–85.
- [30] H. Goëau, P. Bonnet, A. Joly, I. Yahiaoui, D. Barthélémy, B. Nozha, J.-F. Molino, The ImageCLEF 2012 Plant identification Task, in: *CLEF 2012*, 2012.
- [31] A. Arora, A. Gupta, N. Bagmar, S. Mishra, A. Bhattacharya, A Plant Identification System using Shape and Morphological Features on Segmented Leaflets: Team IITK, CLEF 2012, in: *CLEF (Online Working Notes/Labs/Workshop)*, 2012.
- [32] C. A. Price, O. Symonova, Y. Mileyko, T. Hille, J. S. Weitz, LEAF GUI: segmenting and analyzing the structure of leaf veins and areoles, *Plant Physiology* 155 (1) (2010) 236–245.
- [33] S. Dhondt, D. Van Haerenborgh, C. Van Cauwenbergh, R. M. Merks, W. Philips, G. T. Beemster, D. Inzé, Quantitative analysis of venation patterns of *Arabidopsis* leaves by supervised image analysis, *The Plant Journal* 69 (3) (2012) 553–563.
- [34] A. French, S. Ubeda-Tomás, T. J. Holman, M. J. Bennett, T. Pridmore, High-Throughput Quantification of Root Growth Using a Novel Image-Analysis Tool, *Plant Physiology* 150 (4) (2009) 1784–1795.
- [35] L. Wang, I. V. Uilecan, A. H. Assadi, C. A. Kozmik, E. P. Spalding, HYPOTrace: Image Analysis Software for Measuring Hypocotyl Growth and Shape Demonstrated on *Arabidopsis* Seedlings Undergoing Photomorphogenesis, *Plant Physiology* 149 (4) (2009) 1632–1637.
- [36] A. Clément, B. Vigouroux, Unsupervised segmentation of scenes containing vegetation (*Forsythia*) and soil by hierarchical analysis of bi-dimensional histograms, *Pattern Recognition Letters* 24 (12) (2003) 1951–1957.
- [37] A. L. Yuille, D. S. Cohen, P. W. Hallinan, Feature extraction from faces using deformable templates, in: *Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 1989, pp. 104–109.
- [38] D. Cremers, F. Tischhäuser, J. Weickert, C. Schnörr, Diffusion Snakes: Introducing Statistical Shape Knowledge into the Mumford-Shah Functional, *International Journal of Computer Vision* 50 (3) (2002) 295–313.
- [39] P. F. Felzenszwalb, Representation and detection of deformable shapes, *IEEE Transactions on Pattern Analysis and Machine Intelligence*

- 27 (2) (2005) 208–220.
- [40] Colorimetry, Tech. Rep. 15.2, CIE Publication, Vienna: Central Bureau of the CIE (1986).
- [41] M. Tuceryan, A. K. Jain, Texture Analysis, 1993, Ch. Texture Analysis, pp. 235–276.
- [42] P. Howarth, S. Rüger, Evaluation of Texture Features for Content-Based Image Retrieval, in: Image and Video Retrieval, Vol. 3115 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2004, pp. 326–334.
- [43] D. M. Titterington, A. F. M. Smith, U. E. Makov, Statistical Analysis of Finite Mixture Distributions, Wiley, 1985.
- [44] A. Declercq, J. H. Piater, Online Learning of Gaussian Mixture Models: a Two-Level Approach, in: Proceedings of the 3rd International Conference on Computer Vision Theory and Applications (VISAPP), Vol. 1, 2008, pp. 605–611.
- [45] M. R. Golzarian, M. K. Lee, J. M. A. Desbiolles, Evaluation of Color Indices for Improved Segmentation of Plant Images, Transactions of the American Society of Agricultural and Biological Engineers 55 (1) (2012) 261–273.
- [46] N. Otsu, A threshold selection method from gray-level histograms, IEEE Transactions on Systems, Man and Cybernetics 9 (1) (1979) 62–66.
- [47] T. F. Chan, Sandberg, L. A. Vese, Active Contours without Edges for Vector-Valued Images, Journal of Visual Communication and Image Representation 11 (2) (2000) 130–141.
- [48] M. M. Abdelsamea, S. A. Tsaftaris, Active Contour Model driven by Globally Signed Region Pressure Force, in: Proceedings of the 18th International Conference on Digital Signal Processing (DSP), 2013.
- [49] S. Chen, R. J. Radke, Level Set Segmentation with Both Shape and Intensity Priors, in: 12th International Conference on Computer Vision, IEEE, 2009, pp. 763–770.
- [50] D. Cremers, S. J. Osher, S. Soatto, Kernel density estimation and intrinsic alignment for knowledge-driven segmentation: Teaching level sets to walk, in: International Journal of Computer Vision, 2004, pp. 36–44.
- [51] M. Leventon, O. Faugeras, E. Grimson, W. Wells, Level Set Based Segmentation with Intensity and Curvature Priors, in: Proceedings of IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, 2000, pp. 4–11.
- [52] C. P. Lee, W. Snyder, C. Wang, Supervised Multispectral Image Segmentation using Active Contours, in: Proceedings of the International Conference on Robotics and Automation (ICRA), IEEE, 2005, pp. 4242–4247.
- [53] H.-K. Zhao, T. Chan, B. Merriman, S. Osher, A Variational Level Set Approach to Multiphase Motion, Journal of Computational Physics 127 (1) (1996) 179–195.
- [54] K. Zhang, L. Zhang, H. Song, W. Zhou, Active contours with selective local or global segmentation: A new formulation and level set method, Image and Vision Computing 28 (4) (2010) 668–676.
- [55] X.-F. Wang, D.-S. Huang, H. Xu, An efficient local Chan-Vese model for image segmentation, Pattern Recognition 43 (3) (2010) 603–618.
- [56] P. Arbelaez, M. Maire, C. Fowlkes, J. Malik, Contour Detection and Hierarchical Image Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 33 (5) (2011) 898–916.
- [57] J. Ning, L. Zhang, D. Zhang, C. Wu, Interactive image segmentation by maximal similarity based region merging, Pattern Recognition 43 (2) (2010) 445–456.
- [58] A. Levinstein, A. Stere, K. Kutulakos, D. Fleet, S. Dickinson, K. Siddiqi, TurboPixels: Fast Superpixels Using Geometric Flows, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (12) (2009) 2290–2297.
- [59] G. Kim, E. P. Xing, L. Fei-Fei, T. Kanade, Distributed Cosegmentation via Submodular Optimization on Anisotropic Diffusion, in: 13th International Conference on Computer Vision (ICCV), 2011, pp. 169–176.
- [60] L. Zhang, Z. Gu, H. Li, SDSP: A Novel Saliency Detection Method by Combining Simple Priors, in: International Conference on Image Processing (ICIP), 2013.
- [61] L. A. C. Mansilla, P. A. V. Miranda, Image Segmentation by Oriented Image Foresting Transform: Handling Ties and Colored Images, in: Proceedings of the 18th International Conference on Digital Signal Processing (DSP), 2013.



Figure 1: Arabidopsis images showing shape variability. Sketch representations of Arabidopsis mutants adapted from [11] (top); two plants imaged at different growth stages, earlier shown left-most (middle); and a top-view example image shown many individual plants (bottom).



Figure 2: Examples of challenging images due to changing conditions during an experiment. Examples shown are: drought where the soil appears more brown (top); green moss growing on the soil (middle); and water residual in the tray causing reflections (bottom).

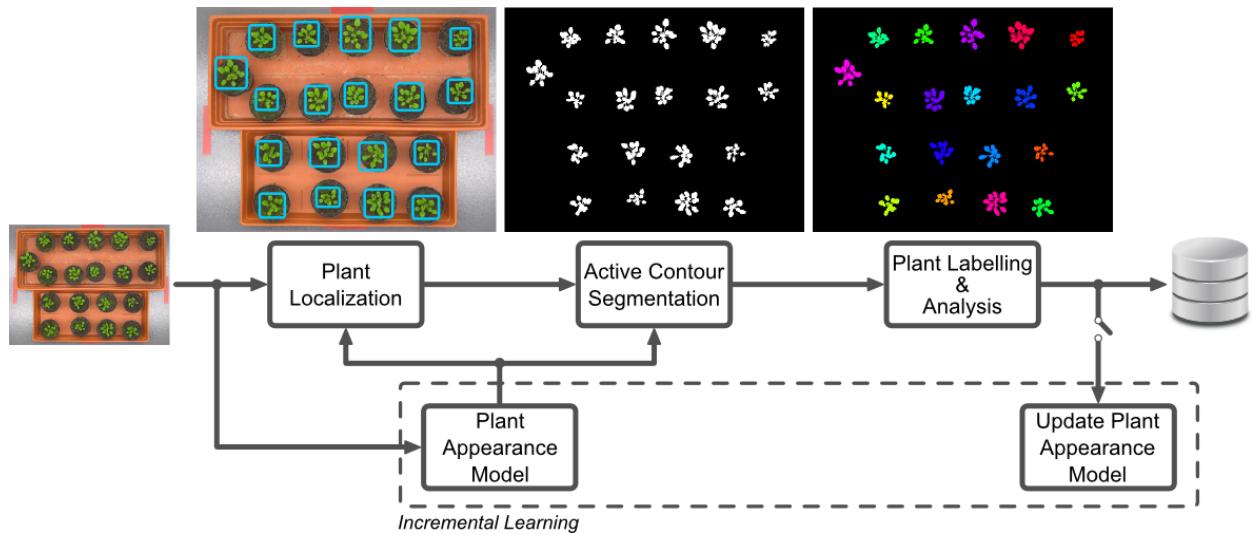


Figure 3: Graphical representation of the proposed approach.

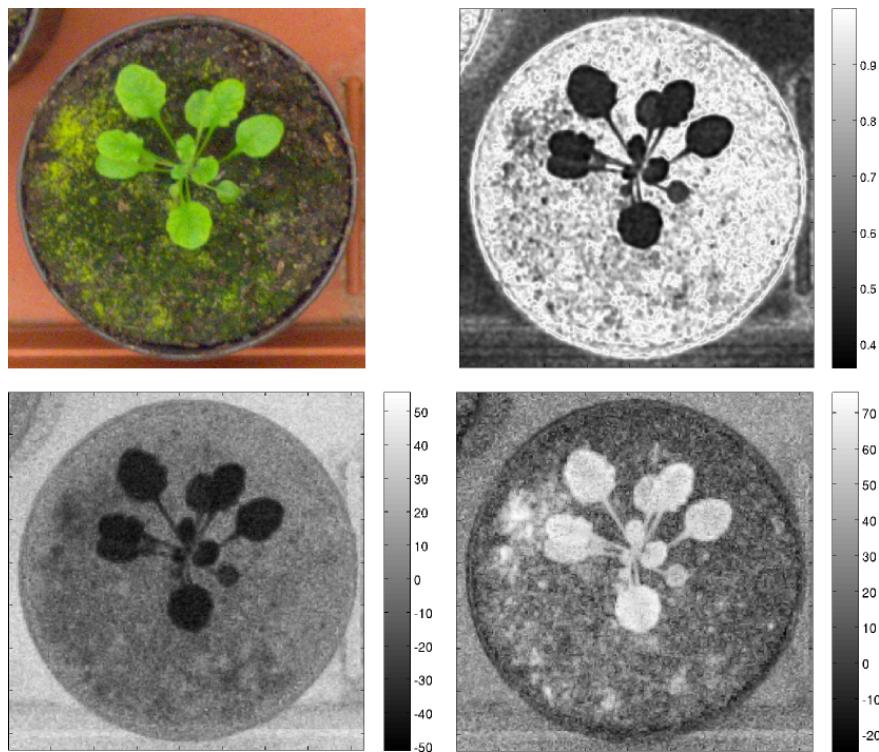


Figure 4: Examples of the image features used in our proposed system. Shown are: the original RGB image (top left); texture descriptor (top right);  $a^*$  (bottom left) and  $b^*$  (bottom right) color components.

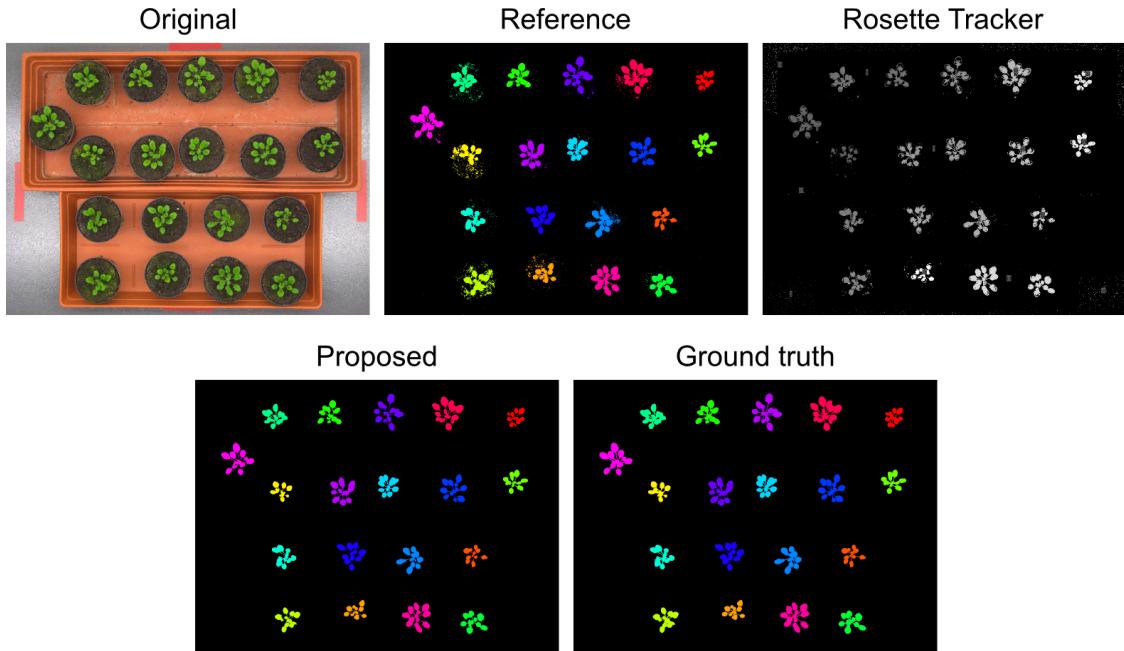


Figure 5: A sample image from the dataset, its ground truth segmentation obtained manually, and the results of the systems being compared, where plant labelling is denoted by color.

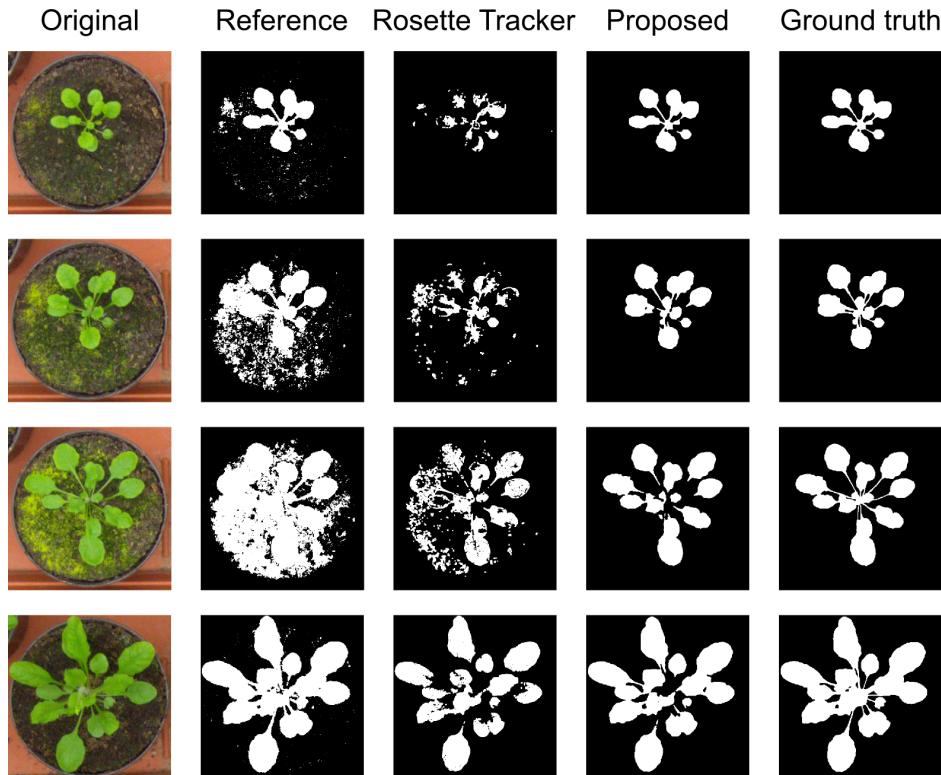


Figure 6: Segmentation result of the different systems for the same plant at different time points. Rows: the same plant in different time instances. Columns starting from left to right: the original plant, then the segmentation results of Reference, Rosette Tracker, and Proposed methods respectively, and Ground truth (right most).

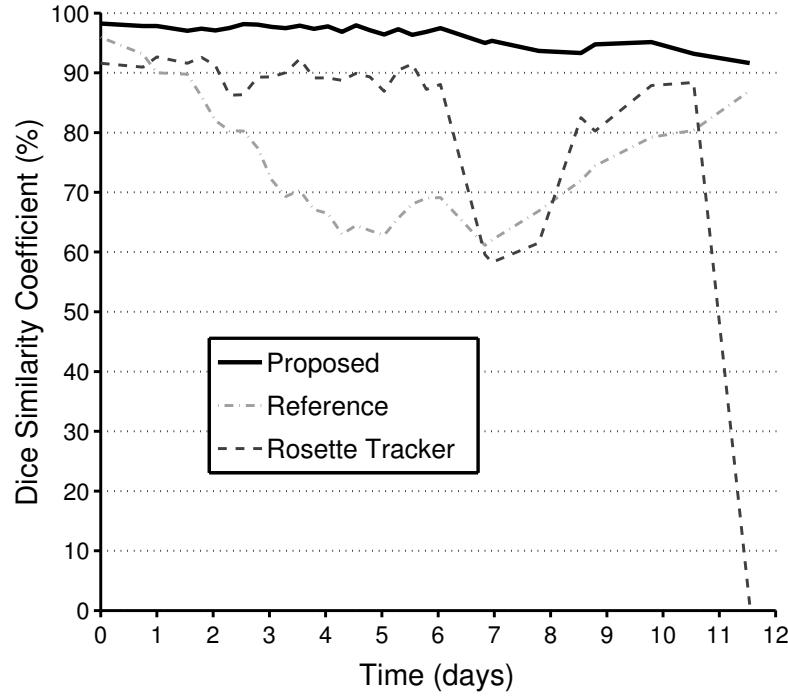


Figure 7: Segmentation accuracy over time estimated using the Dice Similarity Coefficient for the Proposed, Reference and Rosette Tracker methods.

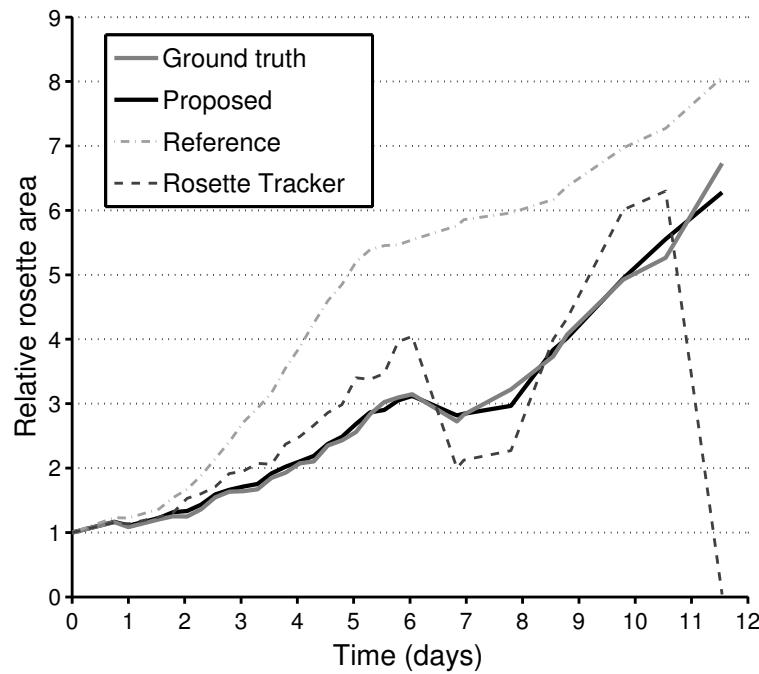


Figure 8: A plant's growth pattern obtained using the evaluated systems is compared against the ground truth. Relative area is reported as projected area normalized by the area of the plant in the first day.

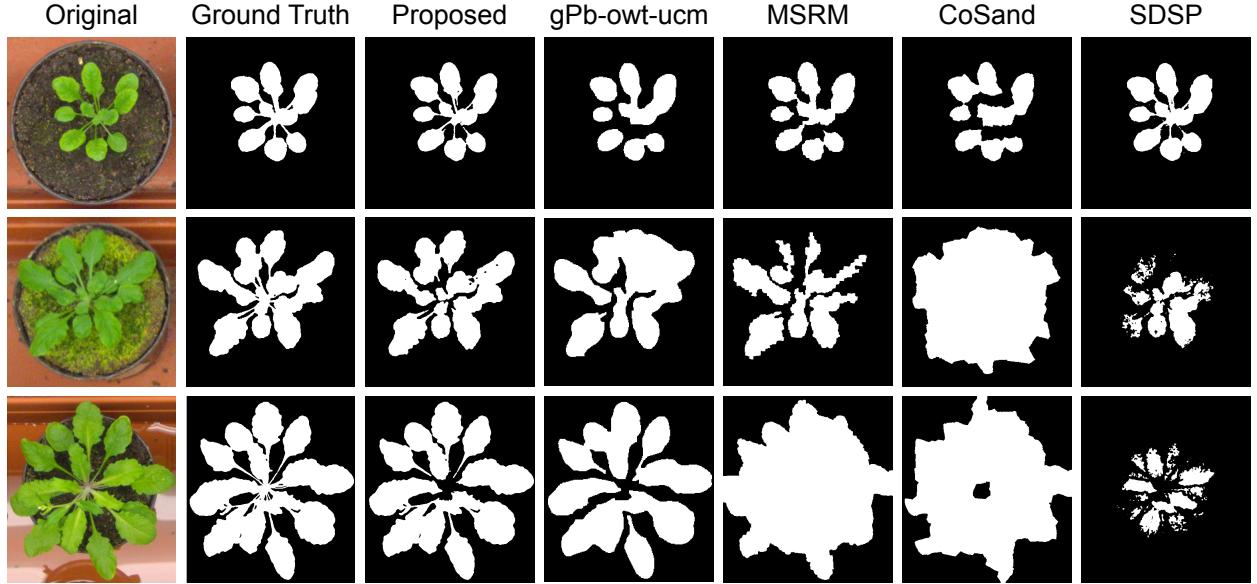


Figure 9: Example segmentation results of different state-of-the-art methods. Columns starting from left to right: the original image, the ground truth, then the segmentation results of Proposed method, *gPb-owt-ucm* [56], *MSRM* [57], *CoSand* [59], *SDSP* [60], respectively.

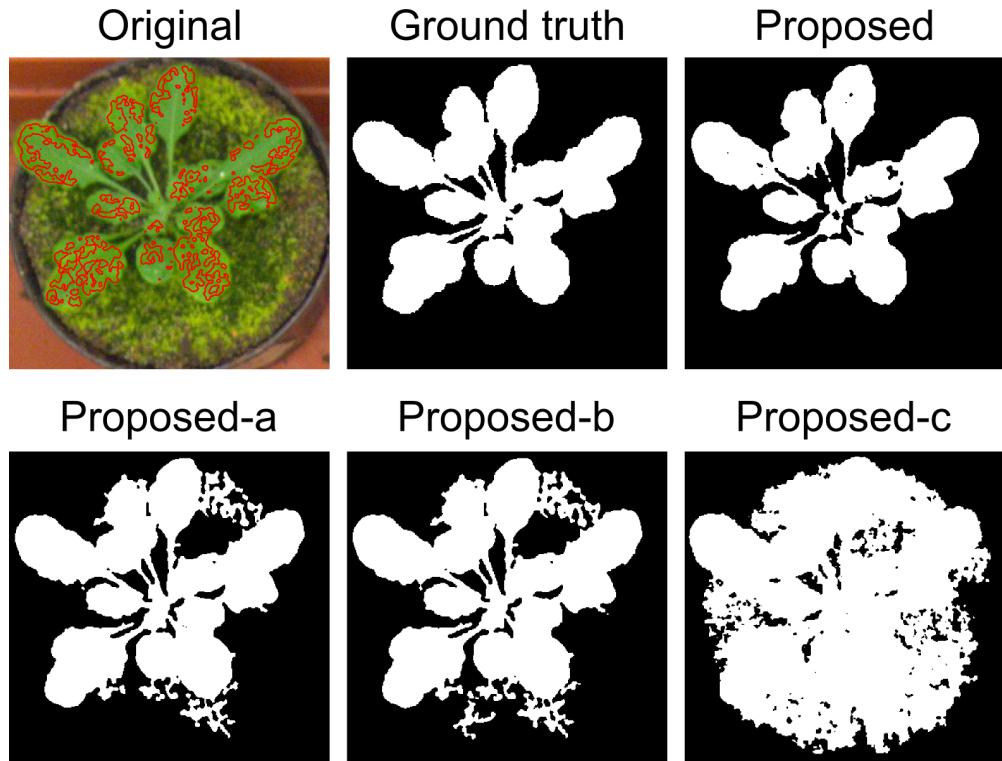


Figure 10: Segmentation outputs for a plant image (shown top left) illustrating the importance of each components of the proposed approach as described in text and summarized in Table 2. The same contour was used for initialization which is shown overlaid in red on the original image (top left).