Root Identification in Minirhizotron Imagery with Multiple Instance Learning

Guohao Yu, Alina Zare, Senior Member, IEEE, Hudanyun Sheng, Roser Matamala, Joel Reyes-Cabrera, Felix B. Frischi and Thomas E. Juenger

Abstract—In this paper, multiple instance learning (MIL) algorithms to automatically perform root detection and segmentation in minirhizotron imagery using only image-level labels are proposed. Root and soil characteristics vary from location to location, thus, supervised machine learning approaches that are trained with local data provide the best ability to identify and segment roots in minirhizotron imagery. However, labeling roots for training data (or otherwise) is an extremely tedious and time-consuming task. This paper aims to address this problem by labeling data at the image level (rather than the individual root or root pixel level) and train algorithms to perform individual root pixel level segmentation using MIL strategies. Three MIL methods (MI-ACE, miSVM, MIForests) were applied to root detection and compared to non-MIL approaches. The results show that MIL methods improve root segmentation in challenging minirhizotron imagery and reduce the labeling burden. In our results, miSVM outperformed other methods. The MI-ACE algorithm was a close second with an added advantage that it learned an interpretable root signature which identified the traits used to distinguish roots from soil and did not require parameter selection.

Note to Practitioners—Minirhizotrons provide an efficient and non-destructive way to collect plant roots for studying root system dynamics. However, existing software used to extract roots from minirhizotron images require significant, tedious manual marking of roots and soil in the collected imagery. Due to this slow manual process, the ability to collect useful information from a large number of minirhizotron images is bottlenecked. In this paper, we propose an automated approach to segment roots from minirhizotron images. The proposed methods not only automatically identifies and segments root pixels in imagery, but also allow for an efficient approach to label training data. This allows one to be able to re-train the models for adaptation to new environments and soil conditions. The methods we proposed in the paper only require one to label each training image as having roots or not (as opposed to labeling individual pixels).

Index Terms—segmentation, multiple instance learning, minirhizotron image, image processing, plant roots, imprecise labels

I. INTRODUCTION

THE lack of good sensors, instruments and techniques for field soil measurements is limiting decisive and rapid advances in soil, ecosystem and agronomic science. Currently, only a few techniques are available for visualization of roots in field conditions, and those that exist are labor intensive and/or limited in what they can see and measure. Because the root system of plants is involved in the absorption of nutrients and water, understanding their traits, dynamics, and behavior is important to a broad number of disciplines, including soil and plant science, agronomists, hydrologist, earth system modelers and others. For example, the study of root systems has become vital as we look for factors that could contribute to the improvement of crop yields to guarantee the food supply. Furthermore, it has been shown that the distribution of roots at depth and their response to fertilizer application substantially affected yield of rice plants [1], and changes in root system architecture, that resulted in improvements in water capture in soil, were directly associated to biomass accumulation and historical yield trends in maize [2]. Similarly, as we look for a sustainable way to manage ecosystem services, there is a critical need to study the root system, as roots are major contributors to the buildup of organic matter in soil, and the role they can play in greenhouse gas mitigation through enhancing the sequestration of carbon in soil.

Common sampling methods used to examine root system dynamics are destructive and provide limited ability to draw inferences on the plant response to stresses experienced during the growing season. For example, soil coring, excavation, trenches, or ingrowth cores extract a unique dataset in time that destroys the integrity of the root system and thus limits repeat measurements from the same plant without the introduction of confounding effects [3]. In contrast to these destructive methods, once inserted into the soil, transparent minirhizotron access tubes allow non-destructive assessment of roots over extended periods of time without repeatedly altering critical soil conditions or root processes [4, 5, 6, 7]. Roots that grow adjacent to the tube are imaged by inserting a camera (with a light source) into the tube and acquiring images along the length of the tube, providing information about roots present at different soil depths as illustrated in Fig. [1] and [1b]. Since minirhizotron tubes allow for the collection of root images over time, this enables monitoring of complex root growth and turnover dynamics [4]. After collection, minirhizotron images are used to measure and characterize root traits, a process that currently is time consuming and tedious since standard analysis approaches involve manually outlining and labeling roots using commercially available software such as WinRhizo.
Although minirhizotrons significantly advance our ability to study plant root systems, data extraction from the acquired images commonly limits the extent of their use as the majority of software available to process images collected in the minirhizotron requires human differentiation of roots and soil. For example, high frequency data collection can produce large numbers of images to manually segment and track. Therefore, the use of machine learning algorithms capable of detecting and quantifying important root traits can greatly accelerate data extraction. Although efforts to automate the labeling and segmentation of roots in minirhizotron imagery have been undertaken, success to extract root traits in heterogeneous soil textures and depths has been limited. For instance, Zeng et al. [8] used a local entropy thresholding (LET) and AdaBoost approach to extract roots from the green band of minirhizotron imagery after preprocessing with linear stretching and matched filters. The AdaBoost algorithm was used to combine five measures of connected components from the binarized LET results. This method works well for detecting young bright roots. However, the matched filter portion of this algorithm imposes strict constraints on root shape, which may result in the missed detection of small, slim, or jagged roots. Furthermore, Zeng et al. [9] also proposed a method in which a trained classifier was used to identify seed points along root centerlines. Then, using a rule base, seed points were connected along root centerlines and extended to find all root regions. The challenge in this approach is the ability to identify root seed points reliably with a simple classifier. Shojaedini and Heidari [10] defined a level set method to segment roots in minirhizotron imagery. This method was then extended by incorporating preprocessing steps that included thresholding and application of the curvelet transform to enhance weak root edges [11]. A significant challenge for root identification that has received limited consideration in these reported methods is the complex background in minirhizotron imagery that varies in both texture and intensity. Due to this natural complexity, applications of unsupervised learning methods for root identification have mixed performance results, as they tend to be very sensitive to variation in soil conditions, lighting and root color. However, labeling individual root pixels over a large data set to train a supervised learning algorithm is extremely time consuming, tedious, and prone to error.

In this paper we hypothesized that the ability to learn from imprecisely labeled data of multiple instance learning methods will significantly reduce the effort associated to label training data for the application of supervised learning algorithms used in the interpretation (or processing) of minirhizotron root images. In this study we compared the ability of three MIL methods to detect roots of switchgrass plants over a range of soil textures and colors in minirhizotron images and demonstrated that application of these methods could substantially increase the speed of minirhizotron image processing.

II. METHODOLOGY

The proposed method consists of three steps: (1) image pre-processing and feature extraction; (2) root detection using an MIL algorithm (MI-ACE, miSVM, or MIForests); and (3) post-processing to help reduce false detections.

A. Pre-processing and Feature Extraction

The minirhizotron imagery collected for this study has vertical striping artifacts which is a common occurrence that happens when the numerous sensors in the scan head are slightly differently calibrated resulting in background stripes, or slightly darker or lighter hue bands as shown in Fig. 2a. Pre-processing of the imagery was conducted to remove this striping noise by subtracting the column mean from each column in the image and adding the global image mean, $\bar{I}$, back to the image as shown in (1):

$$I_p(m,n,b) = I(m,n,b) - \sum_m I(m,n,b) + \bar{I}(b)$$

where $I(m,n,b)$ is the pixel in the $m^{th}$ row, $n^{th}$ column and $b^{th}$ band of image $I$, $\bar{I}(b)$ is the global mean of $I$ in band $b$, and $I_p$ is the resulting pre-processed image. An example of the pre-processed image is shown in Fig. 2b.

After pre-processing, the images are oversegmented into superpixels using the Simple Linear Iterative Clustering (SLIC) algorithm [12]. Superpixels are groups of pixels which are spatially-contiguous and perceptually similar in color as shown in Fig. 2c. This step helps to reduce the overall computational load since processing is done on the reduced superpixel level as opposed to the extremely large number of original pixels.

(a) Original image (b) Pre-processed (c) Superpixel segmentation of a sub-image

Fig. 2: An example of the image pre-processing sequence: (a) shows the original minirhizotron image, (b) shows the pre-processed image after removing vertical striping noise, and (c) is an example cropped sub-image illustrating the superpixel-segmentation.
After superpixel segmentation, the mean, variance, and entropy of each band in the RGB and LAB color spaces (mean-R, mean-G, mean-B, mean-L, mean-a, mean-b, var-R, var-G, var-B, var-L, var-a, var-b, H-R, H-G, H-B, H-L, H-a, H-b) are computed to make an 18-dimensional feature vector for each superpixel. After feature extraction, the superpixel feature vectors of each image are scaled by their order of magnitude so that they are normalized for equal weight across the features,

\[ x(i) = \frac{F_i(I_p(m,n))}{s(i)} \]  

where \( F_i \) is the \( i^{th} \) feature generation function, \( s(i) \) is the scaling value for the \( i^{th} \) feature, and \( x(i) \) is the scaled result for the \( i^{th} \) feature. These scaled feature vectors are the features used within the MIL algorithms to differentiate between root superpixels and non-root superpixels.

B. Multiple Instance Learning for Root Segmentation

MIL algorithms [13], [14], [15], [16], [17], [18], [19], [20] only require data to be labeled at the bag level. Bags are a multi-set of instances (superpixels) and each bag is labeled as either “positive” or “negative.” A bag is labeled as a positive bag if at least one of the instances in the bag is an instance of the positive target class (which is, in this case, roots). If none of the instances in a bag is of the target class, the bag is labeled as a negative bag. The advantage of this framework is that the bags can be constructed in such a way to ease the labeling burden. For example, in the case of root segmentation, each image can correspond to a bag and, thus, images can be labeled as either containing roots (i.e., a positive bag) or not containing roots (i.e., a negative bag). It is much easier and more efficient to produce these image-level binary labels as opposed to tracing out individual roots in each image.

To describe this more precisely, let \( B = \{B_1, \ldots, B_K\} \) be the set \( K \) bags with label \( L = \{L_1, \ldots, L_K\} \), where 0 represents a negative bag and 1 represents a positive bag. Each bag, \( B_i \) contains \( N_i \) superpixels, \( B_i = [x_1^i, x_2^i, \ldots, x_{N_i}^i] \). Each feature vector, \( x_k^i \in \mathbb{R}^{18} \) is computed on a superpixel as described in Section II-A. A bag, \( B_i \), is labeled as positive, \( L_i = 1 \), if there exists at least one feature vector in the bag corresponding to a root superpixel.

Three MIL algorithms are investigated for root detection and segmentation, the multiple instance adaptive cosine/coherence estimator (MI-ACE) [20], the multiple instance support vector machine (miSVM) [21], and the multiple instance random forest (MIForests) algorithms. MI-ACE is a supervised multiple instance learning algorithm that estimates a discriminative target signature (in this case, root signatures) during training and, then, uses this discriminative signature within the Adaptive Cosine Estimator detector [22], [23], [24] to detect root in unlabeled test imagery. The advantage of MI-ACE (and other MIL concept learning methods [16], [17], [25], [26], [18], [19]) is that the estimated discriminative target signature can be examined to obtain insight into what characterizes the target (i.e., root) class and can be easily interpreted to understand what root characteristics are being used to detect the roots and distinguish them from the background (i.e., soil). The miSVM algorithm iteratively trains a support vector machine [27] classifier while enforcing the multiple instance learning constraints that each positive bag contains a target and each negative bag does not contain any target. MIForests uses a deterministic annealing approach to optimize an objective function that enforces the MIL constraints on the data while training a random forest classifier. MIForests initially trains a random forest classifier using all data instances and each
instance is labeled the same as its bag label. Then, iteratively refines the instance labels while optimizing the MILs objective function.

We trained the three MIL algorithms with pre-processed minirhizotron images with image-level labels. The trained models were then used to predict the roots in other images and to determine the degree of confidence of those predictions as illustrated in Fig. 3 [4] [5] (see Experimental Setup section for details).

C. Post-processing

After computing the confidence map, some post-processing is applied to reduce the number of false detections. First, the confidence map is binarized by thresholding as shown in Fig. 3 [4]-[6]. After thresholding, root shape and size characteristics are used to filter out likely false detections. The size and eccentricity are computed on each connected component made by pixels identified as root. Connected components with a small size and a small eccentricity are removed from the binarized image. The results after post-processed by size are shown in Fig. 3 [7]-[9]. The results after post-processing by eccentricity are shown in Fig. 3 [10]-[12].

III. EXPERIMENTS

A. Data Description

Switchgrass (Panicum virgatum L.) is recognized as the most promising perennial grass for bioenergy production in the U.S. Nonetheless, despite being the subject of abundant and multidisciplinary research, there is a lack of understanding of the contributions of the switchgrass root system to the adaptation to different environments, biomass productivity, and substantial ecosystem services [28]. The extensive root system of switchgrass is increasingly recognized to contribute to long-term carbon sequestration, high resource use efficiency, and high biomass productivity [29], [30]. However, despite the essential biological function, information about switchgrass root system dynamics in space and time is limited, at least in part because of the challenges associated with access to the soil-root interface without disturbing the soil environment, which can impair subsequent measurements of the same plant [31].

Switchgrass root images were collected from minirhizotron access tubes that were installed in a 2-y old switchgrass field at the U.S. Department of Energy National Environmental Research Park at Fermilab in Batavia, IL, USA. Minirhizotron access tubes measuring 1.82 m in length were installed at 60 off the horizontal axis using an angled, guided hydraulic soil core sampler. The tubes were inserted to reach a maximum vertical depth of approximately 1.2 m, and foam caps were installed on the end protruding from the soil to insulate the tubes, block light, and protect them from UV damage. Images were collected by inserting a CI-602 in-situ root imager (CID BioScience, Camas, WA, USA) in to the minirhizotron tubes. This scanner was set to collect 360-degree images at 300 dpi in 28 cm depth increments.

B. Experimental Setup

The pixel level segmentation was estimated with MIL algorithms (MI-ACE, miSVM, MIForests) using training imagery with only image level labels. The training data for this study contained 34 positive bags and 34 negative bags. Each bag contains instances from an image, and the bag label is the same as the image label.

To help reduce computational complexity, a subset of the instances in each bag were selected and used within the MIL algorithms. These instances were selected in such a way to ensure a wide distribution of root and non-root instances with varying color properties. Specifically, a 200-bin histogram of the green channel values was constructed using the feature vectors across all superpixels. Then, one superpixel corresponding to each non-empty histogram bin was selected at random (uniform random selection) and kept for processing. If all bins are non-empty, this results in 200 instances per bag. Three multiple instance learning models: MI-ACE, miSVM, and MIForests were trained multiple times. Each time, the instances in each bag were randomly selected again using the approach described above and the same training data was used for all three models.

For each method, parameters were tuned to maximize performance. Performance was measured using the F-score [52] on the validation data set. For the miSVM algorithm, a radial basis function (RBF) kernel was used and the parameters were set to be $C = 10$ and $\gamma = 1$. These parameters were determined by varying the RBF kernel width, $\gamma$, from $2^{-15}$ to $2^3$ by a factor of 4. Similarly, the parameter $C$ was varied from $2^{-5}$ to $2^{15}$ by a factor of 4. The parameter settings with the best performance over this range were selected as the parameters to be used for comparison. For MIForests, the parameters were set to be $t = 100$ and $x_d = 4$. These parameters were selected by varying the number of trees $t$ from $2^4$ to $2^{13}$ by a factor of 2 and the dimension of the subset of input variables $x_d$ was varied from 2 to 16 by a factor of 2. MI-ACE did not have parameters that required tuning. After training and validation, the models were tested by a set of images with pixel level labels.

C. Root Detection Results

MIL methods: MIL methods (MI-ACE, MIForests, and miSVM) were run 30 times and the root detection results were evaluated by Receiver Operating Characteristic (ROC) curves, shown in Fig. 4 [4]-[6]. The mean and variance of the true positive rate (TPR) of different methods at false positive rate (FPR) ranging from 0.01 to 0.06 are presented in Table I. It was found that miSVM outperformed the other approaches with a higher average TPR at the same FPR while MI-ACE was a close second. The performance of miSVM was also more consistent than that of the others. Namely, miSVM had the least variance in performance. In contrast, the MIForests variance was one order of magnitude larger than the others, indicating that the MIForests approach was more sensitive to samples selected in each bag than the others.

Comparison with non-MIL methods: The results of MI-ACE, MIForests, and miSVM were compared with non-MIL
The ROC curves show that non-MIL methods can not compete with MIL methods when dealing with imprecisely labeled root data (i.e., labeled at the image level). The ROC curves of non-MIL methods were below the ROC curves of the MIL methods which means that at the same FPR value, the average TPR of non-MIL methods was less than the TPR value of MIL methods. For the case of FPR = 0.03 in Table 1, the average TPR of miSVM was 0.78 which detected 26% more roots than the SVM method and 20% more roots than the RF method. The average TPR of MI-ACE was 0.75 which detected 23% more roots than the SVM method and 17% more roots than the RF method. However, the average TPR of MIForests was similar to the average TPR of RF because the variance of MIForests TPR was one order of magnitude greater than that of non-MIL methods, and some of the ROC curves of the MIForests method overlapped with the RF method ROC curves. Therefore, the average performance of MIForests was similar to the RF.

**Post-processing results:** Results were also evaluated after post-processing. The confidence maps from each image were thresholded at FPR values ranging from 0.01 to 0.05. These binarized maps were constructed by filtering with an eccentricity or size value lower than a fixed threshold to remove connected components. The threshold was varied to produce the ROC curves shown in Fig. 5. In the post-processing step, it was important to decrease the FPR without decreasing TPR too much so that false roots were removed from binarized maps and root objects were not affected. The eccentricity can effectively decrease FPR and keep TPR changing slowly such that at the same FPR, the TPR on post-processing ROC curves using eccentricity was larger than the TPR on Org ROC curves as shown in Fig. 5 (a)-(c). This indicated that the root objects had larger eccentricity than non-root objects. Size can also be used to remove false roots. The ROC curves using size to remove likely false roots show that root objects had larger size than non-root objects in Fig. 5 (d)-(f). When the threshold on size was large, the post-processing ROC curves dropped steeply indicating that too many root objects were removed because the size of these root objects were smaller than the corresponding threshold. This experiment proved that eccentricity and size were useful attributes to separate roots from non-root objects.

**Evaluating training features via root signature:** An advantage of the MI-ACE algorithm (in addition the lack of parameters that need to be determined) was that a discriminative root signature was estimated. This signature helped

### Table I: Mean and variance of TPR of comparison algorithms at FPR varying from 0.01 to 0.06.

| FPR  | MI-ACE TPR | MIForests TPR | miSVM TPR | SVM TPR | RF TPR |
|------|------------|---------------|-----------|---------|--------|
|      | mean | var | mean | var | mean | var | mean | var | mean | var |
| 0.01 | 0.66 | 0.38e-03 | 0.49 | 3.7e-03 | 0.70 | 0.21e-03 | 0.40 | 0.87e-03 | 0.51 | 0.47e-03 |
| 0.02 | 0.77 | 0.33e-03 | 0.54 | 4.1e-03 | 0.75 | 0.12e-03 | 0.46 | 0.74e-03 | 0.55 | 0.33e-03 |
| 0.03 | 0.52 | 0.32e-03 | 0.58 | 4.1e-03 | 0.78 | 0.10e-03 | 0.52 | 0.67e-03 | 0.58 | 0.26e-03 |
| 0.04 | 0.77 | 0.29e-03 | 0.61 | 4.0e-03 | 0.80 | 0.08e-03 | 0.55 | 0.56e-03 | 0.60 | 0.23e-03 |
| 0.05 | 0.79 | 0.25e-03 | 0.63 | 3.8e-03 | 0.81 | 0.07e-03 | 0.57 | 0.55e-03 | 0.61 | 0.20e-03 |
| 0.06 | 0.80 | 0.23e-03 | 0.66 | 3.6e-03 | 0.82 | 0.06e-03 | 0.59 | 0.52e-03 | 0.63 | 0.22e-03 |

![Fig. 4: Root detection results with different methods across 30 runs without post-processing.](image)
to identify the unique root attributes that distinguish root pixels from soil. Fig. 6 plots the estimated discriminative root target signatures. The signature bands of mean-R, mean-G, mean-B, mean-L, var-R, var-G, and var-B had positive values with small variances, indicating that these features were generally larger in value for roots than for soil and background. The signature bands of H-a, and H-b had negative values, indicating that roots generally had smaller values in these features than background. The larger the absolute value was, the more informative that feature was in distinguishing roots from soil. The signature bands of mean-a, H-R, H-G, H-B, H-L, H-a, H-b, var-R, var-G, var-B, var-L, var-a, var-b were close to zeros which meant the difference between roots and soil at those feature bands were small. However, in some feature bands, the root signature values were widely spread. At a widely spread band, the signature value depended on the initial value and the differences between training results were large, for example the value of signature at mean-b in Fig. 6.

Evaluating sensitivity of MIL methods to features: The features in these signatures in Fig. 6 can be divided into three groups by their corresponding signature value. The first group was the signature features with large average absolute signature value and small variance, the second group was the features that had small average absolute signature value with small variance, and the third group was features whose signature value had large variance. The comparison experiments with different selected features were done on the same training samples to examine the effects of features with different type of signature value. The experiments using all 18 features were compared with experiments using 9 and 17 features. The

| FPR | MI-ACE-9 mean | MI-ACE-9 var | MI-ACE-17 mean | MI-ACE-17 var | MI-ACE-18 mean | MI-ACE-18 var |
|-----|---------------|--------------|---------------|--------------|---------------|--------------|
| 0.01 | 0.68          | 0.1e-4       | 0.66          | 0.54e-3      | 0.66          | 0.38e-3      |
| 0.02 | 0.74          | 0.03e-4      | 0.72          | 0.34e-3      | 0.72          | 0.33e-3      |
| 0.03 | 0.77          | 0.04e-4      | 0.76          | 0.26e-3      | 0.75          | 0.32e-3      |
| 0.04 | 0.79          | 0.03e-4      | 0.78          | 0.21e-3      | 0.77          | 0.29e-3      |
| 0.05 | 0.80          | 0.04e-4      | 0.80          | 0.19e-3      | 0.79          | 0.25e-3      |
| 0.06 | 0.81          | 0.04e-4      | 0.81          | 0.16e-3      | 0.80          | 0.23e-3      |

TABLE II: Mean and variance of TPR of MI-ACE at FPR ranging from 0.01 to 0.06. The number of features of training dataset were varied to be 9,17, or 18.
features which had large average absolute signature values with small variances were used in experiments with 9 features. Those feature were mean-R, mean-G, mean-B, mean-L, var-R, var-B, var-B, H-a, and H-b. In the experiment with 17 features, the mean-b feature with very large variance was removed from the feature set and all other features were kept.

For MI-ACE and miSVM methods, the features with large average absolute value were important for detection results as seen in Table III and IV. It was also noted that the variances of detection results of MI-ACE and miSVM were significant improved by removing features which had large signature variance and small average absolute signature value. As seen in Table III and IV the variance of TPR with 9 features was one order of magnitude smaller than the variance of TPR with 18 features. Therefore, large average absolute signature values dominated the detection results and the larger the value was, the more consistent the results were for the MI-ACE and miSVM methods.

In contrast, we found that the features with small variance were important for the MIForests method. The sensitivity analysis showed an improvement in the variance when the features with the largest variance were removed as shown in Table III. Although the variance of detection using only the features with the small variance and large mean further improved the variance, it also reduced the TPR mean. Thus, the use of all features with small variance regardless of their mean values was important for the MIForests method.

**Qualitative results:** Root detection was qualitatively compared at FPR = 0.03 by setting the thresholds of area and eccentricity to 300 and 0.95 respectively in Fig. 8. In general, the MIL methods outperformed the non-MIL methods. The miSVM method detected the most roots in images and the MIL methods outperformed the non-MIL methods.

### D. Experiment: Effect of Label Accuracy

1) **Sample versus Image Level Label Training:** In this experiment, we compared the detection accuracy of the different methods when using training data paired with image level label versus training data paired with instance level label. The experimental runs using sample level label were trained using 10 images that were randomly selected from a total of 30 images that had been manually labeled. About 1000 samples labeled as root and 1000 samples labeled as soil were randomly selected in each image. In order to run MIL methods with instance level labeled data, we put only one instance per bag and the bag label was assigned to the label of the only instance in the bag.

ROC curves corresponding to instance and image level label accuracy can be seen in Fig. 7a-7c. Among all three MIL methods, MI-ACE was the most robust algorithm to noisy labeled training data. This was indicated by the overlapping of the ROC curves regardless of label accuracy (Fig. 7a). MIForests and miSVM were more sensitive to label accuracy than MI-ACE (Fig. 7b and 7c). The ROC curves of MIForests and miSVM with image level label were lower than the ROC curves of MIForests and miSVM with instance level label. The results of MIForests and miSVM improved as the label became more accurate. The results also showed that, the variance of ROC curves training with instance level data was larger than that training with image level data, especially for MI-ACE and miSVM. The results of training with image level data was more consistent because it was easy to get a large number of images with image level label, but it was expensive to label every instance and hard to get a large amount of training data with instance level label. A large amount of training data helped the generalization of the learning model so that the test results were more consistent.

2) **Small Bag versus Image Level Label:** We also compared results using training data paired with image level label to results using training data paired with small bag level label. Small bags were generated by SLIC algorithm which were large enough to group about 10 adjacent instances into a bag. It was a positive bag if there were root samples in the bag. Otherwise, it was a negative bag. Similarly to training data with sample level label, the experiment runs using small bag level label were trained using 10 images that were randomly selected from a total of 30 images. About 100 positive bags and 100 negative bags were randomly selected in each image.

ROC curves corresponding to small bag level label and image level label can be seen in Fig. 7d-7f. The results of MI-ACE and miSVM using image level label training data were comparable with results using small bag level label because the ROC curves of MI-ACE and miSVM with different label accuracy overlapped with each other. But the results using image level labeled data were more consistent than the results using small bag level labeled data because there were more labeled data with image level label that improved the generalization of the learning model, indicating the importance of a large amount of data to the learning results. For MIForests, the detection results with small bag level training data were much better than that with image level
training data, proved again the importance of precise label of training data to the learning results of MIForests.

IV. SUMMARY

We propose a new approach to detect roots in minirhizotron images. Our approach uses image level label data to segment roots using MIL methods significantly reducing the effort needed to label training data for the application of supervised learning algorithm. We tested the capability of three MIL methods to achieve competitive root detection over a range of soil textures and colors. Using this approach, we found 1. MI-ACE was useful in identifying the unique root attributes that distinguished root pixels form soil, and 2. the miSVM method performed the best in detecting roots with image-level labels. Overall, our results suggest application of MIL methods can substantially improve the imaging analysis bottleneck currently occurring in the study of roots by the minirhizotron technique.

REFERENCES

[1] K. Liu, A. He, C. Ye, S. Liu, J. Lu, M. Gao, Y. Fan, B. Lu, X. Tian, and Y. Zhang, “Root morphological traits and spatial distribution under different nitrogen treatments and their relationship with grain yield in super hybrid rice,” Scientific reports, vol. 8, no. 1, p. 131, 2018.

[2] G. L. Hammer, Z. Dong, G. McLean, A. Doherty, C. Messina, J. Schussler, C. Zinselmieier, S. Paszkiewicz, and M. Cooper, “Can changes in canopy and/or root system architecture explain historical maize yield trends in the us corn belt?” Crop Science, vol. 49, no. 1, pp. 299–312, 2009.

[3] H. Taylor, D. Upchurch, J. Brown, and H. Rogers, “Some methods of root investigations,” in Developments in Agricultural and Managed Forest Ecology. Elsevier, 1991, vol. 24, pp. 553–564.

[4] M. Johnson, D. Tingey, D. Phillips, and M. Storm, “Advancing fine root research with minirhizotrons,” Environmental and Experimental Botany, vol. 45, no. 3, pp. 263–289, 2001.

[5] G. Bates, “A device for the observation of root growth in the soil,” Nature, vol. 139, no. 3527, pp. 966–967, 1937.

[6] J. Waddington, “Observation of plant roots in situ,” Canadian Journal of Botany, vol. 49, no. 10, pp. 1850–1852, 1971.

[7] B. Rewald and J. E. Ephrath, “Minirhizotron techniques,” Plant roots: The hidden half, pp. 1–15, 2013.

[8] G. Zeng, S. T. Birchfield, and C. E. Wells, “Detecting and measuring fine roots in minirhizotron images using matched filtering and local entropy thresholding,” Machine Vision and Applications, vol. 17, no. 4, pp. 265–278, 2006.

[9] ——, “Rapid automated detection of roots in minirhizotron images,” Machine Vision and Applications, vol. 21, no. 3, pp. 309–317, 2010.

[10] S. Shojaeipour and M. Heidari, “A new method for root detection in minirhizotron images: Hypothesis testing based on entropy-based geometric level set decision,” International Journal of Engineering-Transactions A: Basics, vol. 27, no. 1, p. 91, 2013.

[11] B. H. RAHMANZADEH and S. Shojaedini, “Novel automated method for minirhizotron image analysis: Root detection using curvelet transform,” 2016.

[12] R. Achanta, A. Shaji, K. Smith, A. Luncchi, P. Fua, and S. Süsstrunk, “Slic superpixels,” Tech. Rep., 2010.

[13] T. G. Dietterich, R. H. Lathrop, and T. Lozano-Pérez, “Solving the multiple instance problem with axis-parallel rectangles,” Artificial intelligence, vol. 89, no. 1-2, pp. 31–71, 1997.

[14] J. Bolton, P. Gader, H. Frigui, and P. Torrione, “Random set framework for multiple instance learning,” Information Sciences, vol. 181, no. 11, pp. 2061–2070, 2011.

[15] Y. Chen, J. Bi, and J. Z. Wang, “Miles: Multiple-instance learning via embedded instance selection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 1931–1947, 2006.

[16] O. Maron and T. Lozano-Pérez, “A framework for multiple-instance learning,” in Advances in neural information processing systems, 1998, pp. 570–576.

[17] Q. Zhang and S. A. Goldman, “Em-dd: An improved multiple-instance learning technique,” in Advances in neural information processing systems, 2002, pp. 1073–1080.

[18] A. Zare and P. Gader, “Pattern recognition using functions of multiple instances,” in Pattern Recognition (ICPR), 2010 20th International Conference on. IEEE, 2010, pp. 1092–1095.

[19] C. Jiao and A. Zare, “Functions of multiple instances for learning target signatures,” IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 8, pp. 4670–4686, 2015.

[20] A. Zare, C. Jiao, and T. Glenn, “Discriminative multiple instance hyperspectral target characterization,” IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 10, pp. 2342–2354, 2018.

[21] S. Andrews, I. Tsochantaridis, and T. Hofmann, “Support vector machines for multiple-instance learning,” in Advances in neural information processing systems, 2003, pp. 577–584.

[22] J. Broadwater and R. Chellappa, “Hybrid detectors for subpixel targets,” IEEE transactions on pattern analysis and machine intelligence, vol. 29, no. 11, 2007.

[23] S. Kraut and L. L. Scharf, “The cfar adaptive subspace detector is a scale-invariant glrt,” IEEE Transactions on Signal Processing, vol. 49, no. 1, pp. 2538–2541, 1999.

[24] S. Kraut, L. L. Scharf, and L. T. McWhorter, “Adaptive subspace detectors,” IEEE Transactions on Signal Processing, vol. 49, no. 1, pp. 1–16, 2001.

[25] A. Shrivastava, J. K. Pillai, V. M. Patel, and R. Chellappa, “Dictionary-based multiple instance learning,” in Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014, pp. 160–164.
A. Shrivastava, V. M. Patel, J. K. Pillai, and R. Chellappa, “Generalized dictionaries for multiple instance learning,” *International Journal of Computer Vision*, vol. 114, no. 2-3, pp. 288–305, 2015.

C. Cortes and V. Vapnik, “Support-vector networks,” *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

M. R. Schmer, K. P. Vogel, R. B. Mitchell, and R. K. Perrin, “Net energy of cellulosic ethanol from switchgrass,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 2, pp. 464–469, 2008.

L. Comas, S. Becker, V. M. V. Cruz, P. F. Byrne, and D. A. Dierig, “Root traits contributing to plant productivity under drought,” *Frontiers in plant science*, vol. 4, p. 442, 2013.

M.-A. de Graaff, J. Six, J. D. Jastrow, C. W. Schadt, and S. D. Wulfsbeger, “Variation in root architecture among switchgrass cultivars impacts root decomposition rates,” *Soil Biology and Biochemistry*, vol. 58, pp. 198–206, 2013.

G. Berhongaray, I. Janssens, J. King, and R. Ceulemans, “Fine root biomass and turnover of two fast-growing poplar genotypes in a short-rotation coppice culture,” *Plant and soil*, vol. 373, no. 1-2, pp. 269–283, 2013.

D. M. Powers, “Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation,” 2011.

**Guohao Yu** received the B.S. degree in electrical engineering from the Harbin Engineering University, Harbin, China, and the M.S. degree in optical engineering from the Chinese Academy of Sciences, Shanghai, China in 2008 and 2011, respectively. He is currently a graduate research assistant working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Florida. His research interests include multiple instance learning, semantic segmentation and minirhizotron image analysis.

**Alina Zare** received the Ph.D. degree from the University of Florida, Gainesville, in 2008. She is currently an associate professor in the Department of Electrical and Computer Engineering, University of Florida. Her research interests include machine learning and pattern recognition, sparsity promotion, target detection, hyperspectral image analysis, and remote sensing. She is a recipient of the 2014 National Science Foundation CAREER award and the 2014 National Geospatial-Intelligence Agency New Investigator Program Award. She is a senior member of the IEEE.

**Hudanyun Sheng** received the B.Sc. degree in physics from Tongji University, Shanghai, China, in 2015, the M.Sc. degree in Industrial and Systems Engineering from University of Florida, Gainesville, Florida. She is now working towards the M.Sc. degree in Electrical and Computer Engineering in University of Florida, and at the same time working as a graduate research assistant in the Machine Learning and Sensing Lab.

**Roser Matamala** received her M.Sc. and Ph.D. degree in biological sciences from the University of Barcelona, Spain, in 1993 and 1997 respectively. She was a Research Associate at Duke University in 1998 and then at Argonne National Laboratory in 2000. Currently, she is a Scientist at Argonne and a Fellow at the Northwestern-Argonne Institute of Science and Engineering and at the UChicago Consortium for Advanced Science and Engineering. She is an ecologist and biogeochemist who studies the carbon, water, and energy cycles at the ecosystem level and the changes occurring due to human factors, such as land-use and global climate changes. She is an investigator for many Department of Energy, Office of Science, Biological and Environmental Research projects. In these projects, she has been mapping the stocks and characterizing the composition of soil organic carbon in the Arctic region. She is also studying climate adaptation and sustainability of bioenergy crops across continental scale environmental gradients. In addition, she is developing new sensors for detecting changes in roots and soil moisture funded by ARPA.E and NSF. She has published over 50 articles in scientific journals, book chapters and scientific reports.

**Joel Reyes-Cabrera** was born in Leon, Nicaragua, where he lived 17 years and then moved to Costa Rica. He received his B.Sc. degree in agricultural sciences from EARTH University at Costa Rica in 2009. He entered graduate school at the University of Florida and received his M.Sc. and Ph.D. degrees in 2013 and 2017, respectively. His main areas of research interests are agricultural water management, plant physiology, bioenergy, and carbon fluxes in the agroecosystem. Dr. Reyes-Cabrera is currently a postdoctoral fellow in the Department of Plant Sciences at the University of Missouri where he conducts research on the root biology of bioenergy crops.

**Felix B. Fritschi** is a native of Peru and received an Ing. HTL degree from the Swiss College of Agriculture, an M.S. in Agronomy from the University of Florida, and a Ph.D. in Plant Biology from the University of California, Davis. He joined the University of Missouri in 2007 and currently is the C. Alice Donaldson Professor of Bioenergy Crop Physiology and Genetics in the Division of Plant Sciences. Dr. Fritschi is a member of the Interdisciplinary Plant Group at the University of Missouri and a Fellow of the Crop Science Society of America. Dr. Fritschi has a long-standing interest in sustainable crop production and an active research program aimed at enhancing crop resource use efficiency, particularly as related to production under drought and heat stress conditions. Dr. Fritschi’s program leverages physiological mechanism-based strategies and genetics to better understand and ultimately reduce the impacts of abiotic stress on crop plants. He is also engaged in the development of innovative tools and techniques to better characterize plant growth under field conditions, exploit natural genetic variation, and accelerate crop improvement.
Thomas E. Juenger is a Professor in the Department of Integrative Biology at the University of Texas at Austin. He received his Ph.D. in Ecology and Evolutionary Biology at the University of Chicago (1999) and completed postdoctoral research at the University of California Berkeley as a Miller Fellow (1999-2002). Tom’s research focuses on the interface of ecological and evolutionary quantitative genetics in natural plant populations. Much of his research has focused on studies of the genetics of physiological traits, abiotic stress tolerance, and local adaptation and gene-by-environment interaction. More recently, he has led several projects exploring local adaptation in C4 perennial grasses including studies of plant root systems.
Fig. 8: Qualitative examples of root segmentation results with different methods. The 1st column is the original images. The 2nd column is the groundtruth (GT).