2020 25th International Conference on Pattern Recognition (ICPR)
Milan, Italy, Jan 10-15, 2021

A Hierarchical Framework for Leaf Instance Segmentation: Application to Plant Phenotyping

Swati Bhugra, Kanish Garg, Santanu Chaudhury, Brejesh Lall
Department of Electrical Engineering, Indian Institute of Technology Delhi
{*eez138301, ee1160446, santanuc, brejesh*}@ee.iitd.ac.in

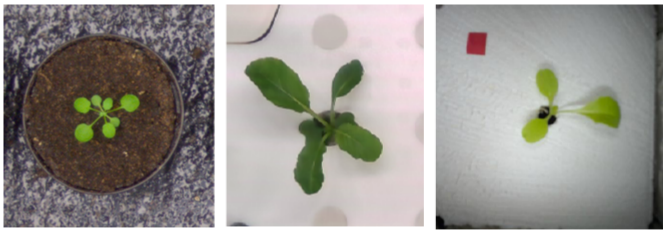
***Abstract*—Image based analysis of plants is a high-throughput and non-invasive approach to study plant traits. The quantitative estimation of many plant traits (leaf area index, biomass etc.) from plant images is primarily based on accurate segmentation of individual leaves. This is a challenging task due to the presence of overlapped leaves and lack of discernible boundaries between them. To overcome these limitations, state-of-the-art supervised deep learning algorithms have been recently employed. However, the annotations of individual leaf instances is time consuming, in addition the variability in leaf shapes and its arrangement among different plant species limits the broad utilisation of these algorithms. To relieve this bottleneck, we propose a novel framework that relies on a graph based formulation to extract leaf shape knowledge for the task of leaf instance segmentation. These shape priors are generated based on leaf shape characteristics independent of plant species. Evaluation of the proposed framework on multiple plant datasets i.e. Arabidopsis, Komatsuna and salad demonstrates its broad utility.**

2020 25th International Conference on Pattern Recognition (ICPR) | 978-1-7281-8808-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICPR48806.2021.9411981

I. Introduction

Quantification of morphological plant traits is a key component of precision agriculture [1]. The utilisation of different tools and methods for this quantitative analysis is known as plant phenotyping [2]. Traditional plant phenotyping is based on destructive sampling of plants from the fields, however this prevents comprehensive study of long term processes e.g. plant growth. In contrast, image based phenotyping is non-invasive that allows temporal plant monitoring [3]. In this context, high-throughput phenotyping platforms (HTPPs) have been established that facilitates acquisition of plant image data for growth analysis [3], [4]. Growth monitoring based on the image data is highly dependent on accurate description of leaf derived traits, such as leaf area, leaf length, leaf area index to name a few [5]. A prerequisite for computing these leaf derived traits is segmenting individual leaves from the plant image. However, leaf instance segmentation is a challenging task due to the variability in leaf pose and presence of various degrees of overlap among plant leaves. Moreover, the lack of discernible boundaries among these overlapping leaves adds to the complexity of the task even in the controlled imaging setups of HTPPs [6]. Given the images of different plant species collected in HTTPs (sample images shown in Figure 1), we address the problem of leaf instance segmentation in this paper.

Within the limited literature available for leaf instance segmentation task, two different approaches exists. One approach is based on (i) traditional computer vision based methods and the other based on (ii) supervised deep neural networks. The proposed deep neural networks (DNNs) [7]–[9] were trained on an open dataset [10] consisting of top-down views of wild-type Arabidopsis and rely on recurrent neural networks (RNNs) to sequentially segment leaf instances. Unlike common object classes (e.g. car, table), large variance with respect to leaf arrangement, shape and size exists for the plant class. Utilisation of these DNNs on different plant species will require large amounts of annotated data that capture this stochastic nature of plant class [11], [12]. However, obtaining such finely grained annotations is tedious and time-consuming [13].



(a) (b) (c)

Fig. 1. Selected images of (a) Arabidopsis, (b) Komatsuna and (c) Salad.

Therefore, with respect to the leaf instance segmentation, an approach based on traditional computer vision methods that utilises global features of leaf characteristics that are invariant across plant species, may provide a better solution without the need for such sophisticated annotations. Current computer vision based solutions proposed for this task either rely on (i) extracting information from Euclidean distance map (EDM) of the segmented plant image or (ii) extracting leaf instances based on shape matching. In reference to the former approach, authors in [14] obtained a skeleton graph image based on Euclidean distance map (EDM) and skeletonisation from the corresponding segmented plant image. This representation was then used for extracting leaf centre points and leaf split points for individual leaf segmentation (denoted as IPK method). Similarly, authors in [6] (denoted as Nottingham method) extracted leaf centres based on EDM but used them as seeds in watershed transform for individual leaf segmentation. However, the accuracy of these approaches is limited by their ability to find leaf centres and thus fails in scenarios with densely overlapped leaves. Another approach reported in [6] (denoted as Wageningen) utilised watershed transform on EDM of the segmented plant image and in the subsequent step, the basins were successively merged based

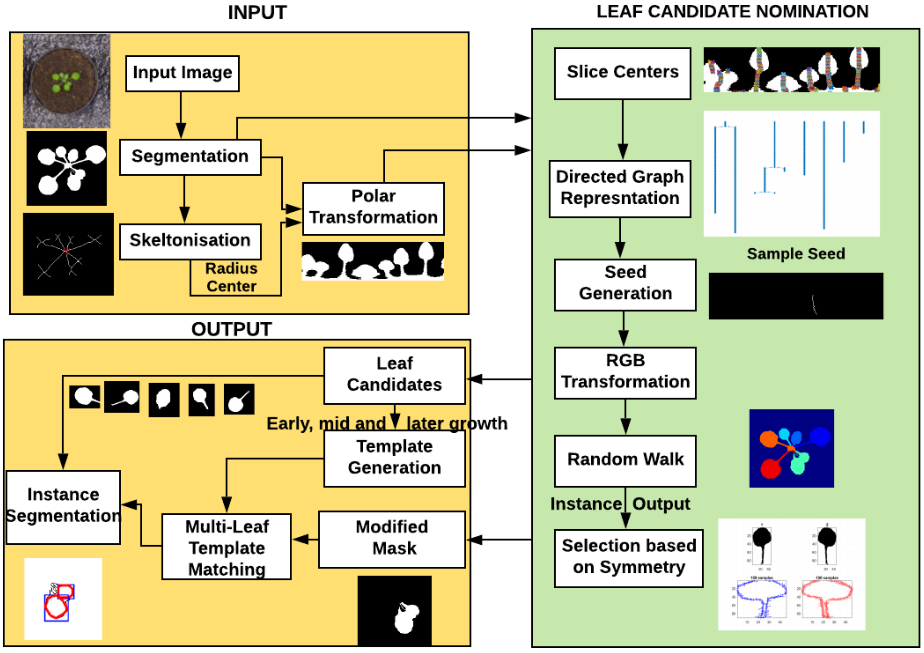


Fig. 2. An overview of the proposed framework

on a threshold. This threshold was empirically estimated from the plant image data and its value was observed to vary with plant species. In contrast to these EDM based methods, authors in [15] (denoted as MSU method) reformulated leaf instance segmentation task as a leaf alignment problem and proposed an optimisation model using the Chamfer distance [16] for aligning leaves iteratively. However, this approach required *a priori* template database which was manually generated by the authors. Similar to the deep learning based methods, the aforementioned traditional computer vision methods were also proposed for rosette plant databases namely Arabidopsis and tobacco. In contrast, [17] proposed a shape matching model for the Sorghum plant, relying on the assumption about leaf shape (i.e. triangular leaves).

As previously discussed, the accuracy of EDM based methods depends on finding the leaf centres and thus fails in overlapping scenarios. In this respect, the template based approach [15] adapts better to occlusions [6]. Our proposed framework is motivated by this work that utilises multiple leaf templates for the task of leaf instance segmentation. However, in contrast to the manually generated templates, we propose to first extract the non-occluded (i.e disconnected) leaf instances based on leaf shape characteristics that are invariably observed across plant species [10] and then utilise these instances as shape priors for template generation. In summary, the paper has the following contributions:

- Novel strategy to exploit the global feature of leaf shapes invariant across plant species.
- Automatic generation of leaf templates for the incorporation of leaf shape knowledge.

Thus, the proposed framework permits automated analysis of different plant species with rich leaf shape variation (as shown in Figure 1), thus relieving the current bottleneck of annotated data [12], [13].

II. Methodology

We present a novel leaf instance segmentation framework (shown in Figure 2) for top-view images of plants collected temporally for growth monitoring in high throughput phenotyping platforms (HTPPs). In such setups, a single plant is visible in each image (sample images shown in Figure 1) [6]. Our proposed framework uses the following definitions that takes into account the biological mechanism of plant growth [10]: Definition (i) The plant leaves spread from the centre of the plant [17], Definition (ii) The leaves are connected to the plant centre [14] and Definition (iii) The leaves are symmetric about the medial axis [10]. We outline each step of the proposed framework in the following subsections.

A. *Segmentation*

The images acquired in phenotyping platforms suffer from non-uniform illumination, moss and shadows [4]. To address the issue of scene variability, a statistical image segmentation algorithm - Mean-shift Bandwidths Searching Latent Dirichlet Allocation (MSBS-LDA) [18] is employed. The algorithm takes advantage of the Latent Dirichlet Allocation (LDA) for the design of visual words and spatial documents. In addition, Mean-shift [19], [20] is utilized for the word-document assignment that enhances the tolerance of the documents to feature differences. To find the optimum bandwidth of Mean-Shift for a stable segmentation, the authors [18] used LUV distances between the topics (clusters). However, the colour of the moss is very similar to the plant leaves thus the utilisation of LUV distance is not suitable for the investigated images. To address this, we employed distance between the texture descriptors presented in [4]. The texture detection filter is a linear combination of Difference of Gaussians (DoG) filter and a pillbox filter [4]. The pillbox filter corresponds to smooth regions in the image (i.e. regions corresponding to leaves and stems) whereas DoG responds to high texture regions (i.e. regions belonging to soil and moss). Thus, this texture descriptor permits the separability of leaves from the moss necessary for an accurate plant segmentation.

B. *Leaf Candidate Nomination*

Given the segmented image and its corresponding skeleton image [21], this step automatically selects the leaf candidates using a graph based formulation (the steps are shown in green block of Figure 2).

The plant leaves spread from the centre exhibiting different orientation (Definition (i)), thus detecting leaf candidate regions will require examining different combinations of orientations and width in the RGB image. To constrain the search of possible leaf candidate regions, firstly the segmented image is transformed into polar coordinates [17]. We compute the centre and radius required for the polar transformations based on its skeleton image [21] (the skeleton end points and skeleton branch points are denoted as nodes). The center is extracted based on the betweenness centrality [22] of the nodes in the skeleton image. The centrality measure of a node reflects how many paths go through that node. More precisely, if node k sits on a shortest path (using Dijkstra's algorithm [23]) between some nodes i and j, then this path counts towards the betweenness of node k. Based on Definition (ii) that all the leaves are connected to the center of the plant, the center corresponds to the node with the maximum betweenness centrality. Similarly, the maximum traversal path from the centre to the end nodes of the skeleton image is utilised as the radius thus preventing the truncation of plant leaves when polar transformed.

Given a polar transformed image, we use the slice representation presented in [17] to automatically generate small set of seeds. A slice is a set of connected pixels for a fixed radius in the polar transformed image. The hierarchical relationship of two vertically connected slices is defined as - the slice with a large radius is a parent of the vertically connected slices with small radius. Based on this hierarchical representation and centre of each slice as a node (as shown in Figure 3), a directed graph is generated (as shown in Figure 2). The paths between each source and sink node of the directed graph are traversed and transformed back to the Cartesian coordinates. These paths are assumed independent and utilised as multi-instance priors in the subsequent step.

The goal of leaf candidate nomination is to extract leaf representatives that can be processed by template generation step. To achieve this, random walk formulation is utilised [24]. This algorithm represent image as a graph with a fixed number of nodes (pixels) and neighboring nodes connected

Fig. 3. (a) Polar transformed sample leaves and (b) corresponding slice centers.

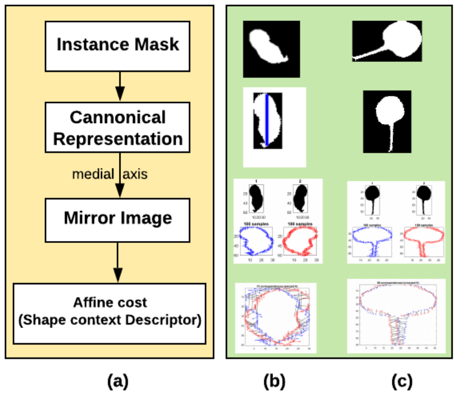


Fig. 4. (a) Workflow to quantify symmetry, (b) Non-leaf sample instance and (c) Leaf sample instance.

by weighted edges. Given the multi-label priors, generated from the previous step, it evaluates the probability with which a random walker starting at each unlabeled pixel will first reach one of the pre-labeled pixels and assigns each unlabeled pixel to the label with the highest probability. In contrast to the region growing methods and level set methods [25], [26], that solve segmentation of each leaf instance individually, this graph based approach permits the simultaneous segmentation of multiple instance priors.

The strategy of extracting seeds based on hierarchical representation of slices and its utilisation in random walk results in under-segmentation of plant image (as shown in Figure 2), this is due to the connectedness property of the vertical slices. However, this property also permits the optimal segmentation of non-occluded leaves that exists as disconnected regions in the plant image. Given the instance output of the random walk, we automatically isolate the non-occluded leaves by exploiting the symmetry property of leaves about the medial axis (Definition (iii)). Each instance obtained from random walk output is first converted to a canonical representation (as shown in step 2 of Figure 4), then the symmetry is quantified using the distance cost between the shape context descriptor [27] of the instance and its mirror image (about the medial axis) as shown in Figure 4. Since shape contexts are rich descriptors, they have also been utilised for leaf image classification [28]. The shape context descriptor quantifies the distribution of the shape relative to a given point by computing a coarse histogram of the relative coordinates of the remaining points. The histogram in shape context computation uses bins, uniform in log-polar space. To compute the dissimilarity score between the instance and its mirror image, the distance is measured as the amount of transformation necessary to align these shapes [29] (also termed as bending energy). It is observed that the utilisation of this distance between instance and its mirror image yield high separability between the two classes i.e. leaf and non-leaf candidates. For example, the shape distance of random walk instance output shown in Figure 4(b) and 4(c) was computed to be 0.3462 and 0.0311 respectively. Thus, this distance performs well even when the overlap percentage of leaves is high (as shown in Figure 4(b)).

The multiple leaf candidates are removed from the segmented image to obtain a modified mask. If the modified mask belongs to a small area of the input segmented mask (i.e. overlap ratio of modified mask and input segmented mask is less than 10%), then these leaf candidates are the final leaf instances extracted from the proposed framework. Otherwise, these candidate forms one part of the instance output and the modified mask is used as an input for multi-leaf template matching (sub-section D). The output of the leaf candidate nomination is shown as input to the output block in Figure 2.

C. Template Generation

Leaf templates have a large influence on multi-leaf template matching. Thus, for generating templates that accommodates large variations in leaf shape and size of a plant species in its growth cycle (shown in Supplementary Figure 1), we only utilise two plant images sampled at different growth stages (three in our case, early, mid and final stages).

Leaf candidates are extracted from these sampled images as discussed in the previous sub-sections. Furthermore, a post-processing pipeline to remove stems based on Hough transform [30] is utilised (as shown in Figure 5). The Hough transform generates a parameter space, where every pixel in an x–y image space is mapped into a line of an m–c parameters space and the peaks represent lines. Since the canonical form of leaves are used, the peaks with m concentrated around zero degrees corresponds to stems (shown in Figure 5). In addition, the peaks for leaf candidates with no stem were observed to have arbitrary angles (varying more than 20 degrees). These representative shapes from sampled images are then used for template generation. Since, the multi-leaf template matching algorithm is rotation and scale sensitive, thus the representative shapes are varied using geometric transformations such as scaling and rotation. This template database with different shapes, size and orientation generated through sampling of time-lapsed images is employed in multi-leaf template matching. In contrast to [15], we generate templates automatically with negligible manual annotation thus eliminating subjectivity.

D. Multi-leaf Template Matching

Given the modified segmented mask (target image) and template database, multi-leaf alignment algorithm presented in [15] is employed. For each template, the algorithm first computes the location of minimum Chamfer matching (CM) distance [16], [31] within the edge map of the target image. These individual templates are then filtered based on its overlap with the target image. Then, an optimal set of leaf candidates based on an objective function [15] with three terms: (i) average CM of the selected leaves (ii) the number of estimated leaves and (iii) the distance of the synthesised mask from the selected leaves and the mask of the target image is selected. The first term promotes the selection of leaf templates with minimum CM distance, the second term is to select an optimal number of leaf templates and the third term selects leaf templates, that when arranged generates the target mask accurately. Thus, this function leads to the joint estimation of optimal leaf instances by iteratively removing redundant leaf candidates. Based on this optimisation using gradient descent, the selected leaf templates forms another part of leaf instance output.

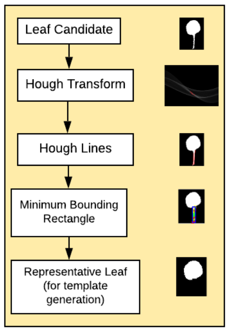


Fig. 5. Workflow to extract representative shapes from leaf candidates

In contrast to [15], we propose graph based leaf template generation method (sub-section B and C), that requires no manual selection of templates. In addition, the removal of non-occluded leaves from the segmented plant images (sub-section B) limits the analysis region for subsequent template matching (sub-section D), thus reducing its computation time.

III. Results

In this section, we show the results and evaluation of the proposed framework on the following datasets, consisting of top-down views of different plant species collected at different growth periods and the corresponding ground truth instance segmented images that has been manually annotated by experts:

- PRL dataset [32], [33].
- The plant phenotyping database [34].
- The Komatsuna dataset [35];
- The Salad dataset [36].

A prerequisite for utilisation of this framework is an accurate plant segmented image, that facilitates the extraction of reliable leaf shapes. In this respect, several approaches [4] have been proposed to segment the plant from the background. For example, the algorithm presented in IPK [6] and Wageningen [6], is based on supervised techniques i.e 3D histogram construction and neural networks respectively. However, to enable the automated analysis of the investigated images, evaluation of only unsupervised segmentation algorithm have been performed (Figure 6 visually illustrates the segmentation outcomes). Overall, we observe that the output of Gaussian mixture model [37] appears noisy whereas K-means [38],

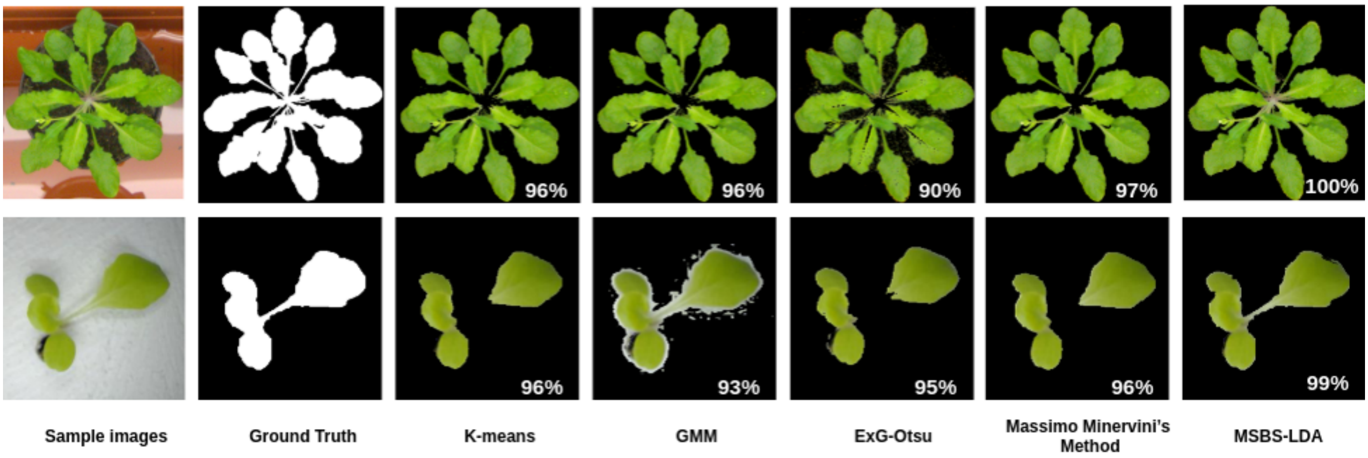


Fig. 6. Example segmentation results of different methods

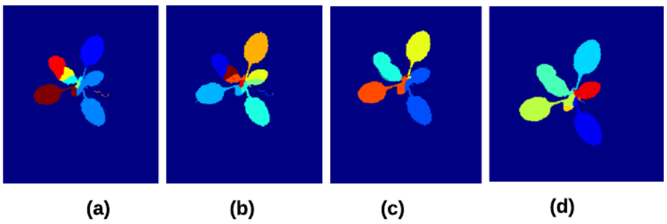


Fig. 7. Selected result with different seed placement approach using Random Walk for region growing (a) Algo 2, (b) Algo 1, (c) Algo 3 and (d) Slice based approach.

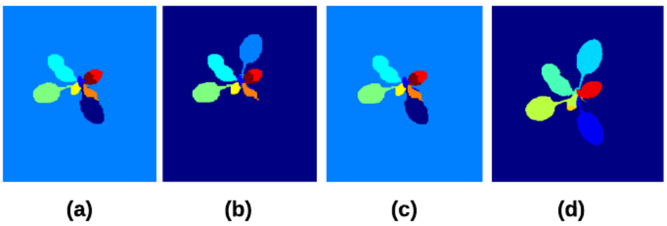


Fig. 8. Selected result with different segmentation method using the seed placement of slice approach (a) Kruskal, (b) Prim, (c) Power-watershed and (d) Random walk.

Otsu's thresholding based on excess green features [4] and Minervini's algorithm [4] over-segments the foreground resulting in disconnected plant region. On the other hand, a significant agreement with the ground truth can be observed for the modified MSBS-LDA algorithm using the texture based distance metric [4].

In the proposed framework, seed placement and subsequent segmentation approach that employ the seed priors have large influence on the leaf template generation. Thus, we conducted a comprehensive set of experiments in regards to these steps (shown in Table I). To quantitatively evaluate the performance based on the extraction of non-overlapped leaves, following criterion is utilised:

- Difference in count (DiC) [6]: mean of the differences between predicted non-overlapped leaves and ground truth non-overlapped leaves (best value when mean is zero).
- Dice (%) [6]: mean overlap between algorithmic extracted non-overlapped leaves and the corresponding ground truth non-overlapped leaves (best value when mean 100%).

TABLE I
LEAF CANDIDATE NOMINATION RESULTS ON PLANT PHENOTYPING DATABASE

Method	DiC	DICE(%)
Algo1 + Random Walk	-1.27	84.1
Algo2 + Random Walk	-1.01	82.3
Algo3 + Random Walk	-.19	85.9
Slice approach + Random walk	**-0.14**	**89.9**
Slice apporach + Prim	-2.45	80.2
Slice apporach + Kruskal	-2.3	81.2
Slice apporach + Power watershed	-2.35	83.1

With respect to seed placement, different approaches were utilised. One approach (denoted as Algo1) was based on orientation field map [39] and leaf tips of the polar transformed image. For leaf tip detection, we have chosen Harris-Stephens corner detector [40]. The seeds were generated by tracing the path from leaf tip, growing in the direction determined by the orientation map [39]. The second approach (denoted as Algo2) was based on the skeleton image of the polar transformed image. Here, the path from the end nodes in the image is traversed and at junctions (having more than one path) a minimum deviated direction is employed. These paths from the end nodes are then utilised as seeds. In contrast to the aforementioned approach, in the third approach (denoted as Algo3) all paths are traversed from the end nodes of the skeleton image and the smoothness of the path was used for selection (pseudo code given in Supplementary Figure 2). Figure 7 visually shows the outcomes of different seed generation approaches with random walk. It was observed that Algo1 and Algo2 over-segments the non-occluded leaves whereas Algo3 performs comparable to the slice based ap-

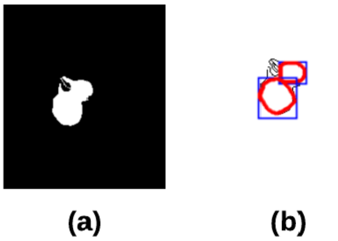


Fig. 9. Sample result of multi-leaf template matching algorithm (a) Modified mask and (b) Output.

proach. Similarly, with respect to segmentation step based on the seeds generated, different methods namely Power Watershed [41], Kruskal's [41] and Prim's Algorithm [41] for Maximum Spanning Forest computation were employed. Based on the quantitative results shown in Table I, it was observed that due to over-segmentation of non-overlapped leaves, the instance output failed the symmetry threshold (empirically estimated to be 0.1 based on few sample leaf and non-leaf candidates) and were labeled as non-leaf candidates, this is evident in DiC. However, the instances from these approaches that were labelled as leaf candidates showed good agreement with corresponding ground truth non-overlapped leaves. It is to be noted that the number of leaf candidates extracted plays an important role in extracting representative shapes especially at later growth stages. In such scenarios, the proposed framework performed well in extracting non-overlapped leaves in contrast to other approaches (shown in Supplementary Figure 3 [12]). It was observed that in the early and mid growth stages, 10 representative shapes were extracted and in later growth stage 3 shapes were obtained for plant phenotyping dataset [34]. In addition, the symmetry test utilised for selection of leaf candidates also discarded non-overlapped leaves which were deformed due to the bending in the top-view images (shown in Figure 1(b)). This further facilitates the extraction of potential leaf shapes for template generation.

Based on the plant phenotyping image database [34], for early and mid growth stage image, multi-leaf template matching was not frequently utilised since the leaf candidates generated comprised of 90% of the input segmented mask. This was due to the non-overlapping scenarios observed in these images. However, it also led to the neglection of very small leaves (resulting in negative DiC). In addition, for the images that did not satisfy the aforementioned criteria, the leaf candidates obtained are then removed from the input segmented mask to extract the modified mask (result shown in Figure 9). This is done to minimise the analysis region for multi-leaf template matching. In this regard, the avergae execution time for multi-leaf template matching was computed to be 1.4 sec/image as compared to MSU with 2.58 sec/image. The time was computed in $MATLAB^{TM}$ on LR dataset [15], [34].

In addition, we also performed a comparative study on PRL dataset (A1 data class). To evaluate the segmentation accuracy on this dataset in addition to DiC, we also employed Symmetric Best Dice (SBD) [6] (results shown in Table II). With respect to this dataset, MSU shows a superior performance as compared to the proposed framework, however the template database utilised in their framework was manually generated. Similarly, in the Wageningen approach the threshold was empirically set for this dataset. Since, the proposed framework is unsupervised, comparison with the supervised deep learning algorithms proposed for instance segmentation has not been performed.

In regards to the salad dataset [36], where the challenges of non-overlapping scenario is minimal, authors [36] utilised Deep Coloring model [42] for leaf instance segmentation. However, based on experimental results (shown in Supplementary Figure 4), we show that the utilisation of the proposed framework with no supervision provides good results.

Conducting these comparative study demonstrates the efficiency of our proposed framework in regards to different plant species. The strategy to extract non-occluded leaves provides for automatic generation of shape priors, that can be further utilised for segmenting individual leaves in overlapping regions of the image. In this regard, it was observed in another study [7] that the LSTM model trained on annotated dataset (PRL) for the task of leaf instance segmentation, handles occlusion by segmenting the non-occluded leaves first. This empirical evidence further encourages the proposed strategy.

TABLE II
SEGMENTATION RESULTS ON PRL DATASET (A1)

Method	SBD(%)	DiC
IPK	74.2	-1.9
Nottingham	68.0	-3.6
Wageningen	72.8	0.4
MSU	78.0	-2.3
Proposed framework	**74.7**	**-2.5**

IV. CONCLUSION

In this paper, we present a novel framework for leaf instance segmentation in image based phenotyping experiments. To relieve the current bottleneck of annotated data corresponding to different plant species, we propose to utilise graph based segmentation on automatically computed seeds for generating leaf shape priors. To propagate the priors to the rest of the image, multi-leaf template matching algorithm is used. Experimental results show the effectiveness of our proposed framework to produce natural leaf shape boundaries. For future work, we will investigate unsupervised deep learning algorithms for further refinement of results.

Acknowledgement. This work is supported by National Agricultural Science Fund (NASF) under Indian Council of Agricultural Research (ICAR), Delhi, India [Phenomics of moisture deficit stress tolerance and nitrogen use efficiency in rice and wheat-Phase II].

REFERENCES

[1] E. P. Spalding and N. D. Miller, "Image analysis is driving a renaissance in growth measurement," *Current Opinion in Plant Biology*, vol. 16, no. 1, pp. 100–104, 2013.

[2] M. E. Ghanem, H. Marrou, and T. R. Sinclair, "Physiological phenotyping of plants for crop improvement," *Trends in Plant Science*, vol. 20, no. 3, pp. 139–144, 2015.

[3] A. Singh, B. Ganapathysubramanian, A. K. Singh, and S. Sarkar, "Machine learning for high-throughput stress phenotyping in plants," *Trends in plant science*, vol. 21, no. 2, pp. 110–124, 2016.

[4] M. Minervini, M. M. Abdelsamea, and S. A. Tsaftaris, "Image-based plant phenotyping with incremental learning and active contours," *Ecological Informatics*, vol. 23, pp. 35–48, 2014.

[5] Y. Chen, S. Baireddy, E. Cai, C. Yang, and E. J. Delp, "Leaf segmentation by functional modeling," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0–0.

[6] H. Scharr, M. Minervini, A. P. French, C. Klukas, D. M. Kramer, X. Liu, I. Luengo, J.-M. Pape, G. Polder, D. Vukadinovic *et al.*, "Leaf segmentation in plant phenotyping: a collation study," *Machine vision and applications*, vol. 27, no. 4, pp. 585–606, 2016.

[7] B. Romera-Paredes and P. H. S. Torr, "Recurrent instance segmentation," in *European conference on computer vision*. Springer, 2016, pp. 312–329.

[8] M. Ren and R. S. Zemel, "End-to-end instance segmentation with recurrent attention," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 6656–6664.

[9] A. Salvador, M. Bellver, V. Campos, M. Baradad, F. Marques, J. Torres, and X. Giro-i Nieto, "Recurrent neural networks for semantic instance segmentation," *arXiv preprint arXiv:1712.00617*, 2017.

[10] D. Milner, H. Hel-Or, D. Keren, S. Raz, and E. Nevo, "Analyzing symmetry in biological systems," in *IEEE International Conference on Image Processing 2005*, vol. 1. IEEE, 2005, pp. I–361.

[11] M. V. Giuffrida, P. Doerner, and S. A. Tsaftaris, "Pheno-deep counter: a unified and versatile deep learning architecture for leaf counting," *The Plant Journal*, vol. 96, no. 4, pp. 880–890, 2018.

[12] D. Ward and P. Moghadam, "Scalable learning for bridging the species gap in image-based plant phenotyping," *Computer Vision and Image Understanding*, p. 103009, 2020.

[13] M. Minervini, M. V. Giuffrida, and S. A. Tsaftaris, "An interactive tool for semi-automated leaf annotation," 2016.

[14] J.-M. Pape and C. Klukas, "3-d histogram-based segmentation and leaf detection for rosette plants," in *European Conference on Computer Vision*. Springer, 2014, pp. 61–74.

[15] X. Yin, X. Liu, J. Chen, and D. M. Kramer, "Multi-leaf alignment from fluorescence plant images," in *IEEE Winter Conference on Applications of Computer Vision*. IEEE, 2014, pp. 437–444.

[16] H. Barrow, J. Tenenbaum, R. Bolles, and H. Wolf, "Parametric correspondence and chamfer matching: Two new techniques for image matching," in *Proceedings: Image Understanding Workshop*. Science Applications, Inc Arlington, VA, 1977, pp. 21–27.

[17] Y. Chen, J. Ribera, C. Boomsma, and E. J. Delp, "Plant leaf segmentation for estimating phenotypic traits," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 3884–3888.

[18] Y. Wang and L. Xu, "Unsupervised segmentation of greenhouse plant images based on modified latent dirichlet allocation," *PeerJ*, vol. 6, p. e5036, 2018.

[19] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[20] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*, vol. 2. IEEE, 2000, pp. 142–149.

[21] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital image processing using MATLAB*. Pearson Education India, 2004.

[22] S. de Sousa, N. M. Artner, and W. G. Kropatsch, "On the evaluation of graph centrality for shape matching," in *International Workshop on Graph-Based Representations in Pattern Recognition*. Springer, 2013, pp. 204–213.

[23] E. W. Dijkstra *et al.*, "A note on two problems in connexion with graphs," *Numerische mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[24] L. Grady, "Random walks for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 11, pp. 1768–1783, 2006.

[25] D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape," *International journal of computer vision*, vol. 72, no. 2, pp. 195–215, 2007.

[26] K. McGuinness and N. E. O'connor, "A comparative evaluation of interactive segmentation algorithms," *Pattern Recognition*, vol. 43, no. 2, pp. 434–444, 2010.

[27] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, no. 4, pp. 509–522, 2002.

[28] M. P. Shah, S. Singha, and S. P. Awate, "Leaf classification using marginalized shape context and shape+ texture dual-path deep convolutional neural network," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 860–864.

[29] F. L. Bookstein, *Morphometric tools for landmark data: geometry and biology*. Cambridge University Press, 1997.

[30] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.

[31] G. Borgefors, "Hierarchical chamfer matching: A parametric edge matching algorithm," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 10, no. 6, pp. 849–865, 1988.

[32] M. Minervini, A. Fischbach, H. Scharr, and S. A. Tsaftaris, "Finely-grained annotated datasets for image-based plant phenotyping," *Pattern recognition letters*, vol. 81, pp. 80–89, 2016.

[33] H. Scharr, M. Minervini, A. Fischbach, and S. A. Tsaftaris, "Annotated image datasets of rosette plants," in *European Conference on Computer Vision. Zürich, Suisse*, 2014, pp. 6–12.

[34] J. A. Cruz, X. Yin, X. Liu, S. M. Imran, D. D. Morris, D. M. Kramer, and J. Chen, "Multi-modality imagery database for plant phenotyping," *Machine Vision and Applications*, vol. 27, no. 5, pp. 735–749, 2016.

[35] H. Uchiyama, S. Sakurai, M. Mishima, D. Arita, T. Okayasu, A. Shimada, and R.-i. Taniguchi, "An easy-to-setup 3d phenotyping platform for komatsuna dataset," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2017, pp. 2038–2045.

[36] D. G. Shadrin, V. Kulikov, and M. Fedorov, "Instance segmentation for assessment of plant growth dynamics in artificial soilless conditions." in *BMVC*, 2018, p. 329.

[37] H. Permuter, J. Francos, and I. Jermyn, "A study of gaussian mixture models of color and texture features for image classification and segmentation," *Pattern Recognition*, vol. 39, no. 4, pp. 695–706, 2006.

[38] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using k-means clustering algorithm and subtractive clustering algorithm," *Procedia Computer Science*, vol. 54, pp. 764–771, 2015.

[39] H. Laga and S. J. Miklavcic, "Curve-based stereo matching for 3d modeling of plants," 2013.

[40] C. G. Harris, M. Stephens *et al.*, "A combined corner and edge detector." in *Alvey vision conference*, vol. 15, no. 50. Citeseer, 1988, pp. 10–5244.

[41] C. Couprie, L. Grady, L. Najman, and H. Talbot, "Power watershed: A unifying graph-based optimization framework," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 7, pp. 1384–1399, 2010.

[42] V. Kulikov, V. Yurchenko, and V. Lempitsky, "Instance segmentation by deep coloring," *arXiv preprint arXiv:1807.10007*, 2018.