

# DEGAST3D: Learning Deformable 3D Graph Similarity to Track Plant Cells in Unregistered Time Lapse Images

Md Shazid Islam<sup>1</sup>, Graduate Student Member, IEEE, Arindam Dutta, Calvin-Khang Ta<sup>1</sup>, Kevin Rodriguez, Christian Michael, Mark Alber<sup>1</sup>, G. Venugopala Reddy<sup>1</sup>, and Amit K. Roy-Chowdhury<sup>1</sup>, Fellow, IEEE

**Abstract**—Tracking plant cells in three-dimensional (3D) tissue captured through light microscopy presents significant challenges due to the large number of densely packed cells, non-uniform growth patterns, and variations in cell division planes across different cell layers. In addition, images of deeper tissue layers are often noisy, and systemic imaging errors further exacerbate the complexity of the task. In this paper, we propose a novel learning-based method DEGAST3D: Learning Deformable 3D GrAph Similarity to Track Plant Cells in Unregistered Time Lapse Images exploits the tightly packed 3D cell structure of plant cells to create a three-dimensional graph for accurate cell tracking. We also propose a novel algorithm for cell division detection and an effective three-dimensional registration, improving state-of-the-art algorithms. On a public dataset, our novel cell pair matching method outperforms the baseline by 6.83%, 5.96%, 6.40% in precision, recall, and F-1 score, respectively. On the same dataset, our proposed novel cell division technique improves the results of the baseline method by 15.38% and 14.78% in terms of recall and F1-score, respectively.

**Index Terms**—Cell tracking, cell division, 3D segmentation, 3D registration, graph matching, deep learning.

## I. INTRODUCTION

**M**ORPHOGENETIC analysis is a cardinal topic of interest in computational biology, which analyzes the development of various biological forms, including cell growth and

Received 15 December 2023; revised 18 December 2024; accepted 19 December 2024. Date of publication 3 January 2025; date of current version 5 February 2025. This work was supported in part by NSF under Grant OAC-2411453, Grant OAC-1664172, Grant DMS-1762063, Grant DMS-2029814 and Grant IOS-2055690. (Corresponding author: Md Shazid Islam.)

Md Shazid Islam, Arindam Dutta, and Amit K. Roy-Chowdhury are with the Department of Electrical and Computer Engineering, University of California, Riverside, Riverside, CA 92521 USA (e-mail: misla048@ucr.edu; adutt020@ucr.edu; amitrc@ucr.edu).

Calvin-Khang Ta is with the Department of Computer Science, University of California, Riverside, Riverside, CA 92521 USA (e-mail: cta003@ucr.edu).

Kevin Rodriguez and G. Venugopala Reddy are with the Department of Botany and Plant Sciences, University of California, Riverside, 2311 EZ Leiden, The Netherlands (e-mail: krodr005@ucr.edu; venug@ucr.edu).

Christian Michael is with the Department of Mathematics of University of California, Riverside, CA 92521 USA (e-mail: cmich004@ucr.edu).

Mark Alber is with the Department of Mathematics, University of California Riverside, Riverside, CA 92521 USA, and also with Mathematical Institute, Leiden University, 2311 EZ Leiden, The Netherlands (e-mail: mark.alber@ucr.edu).

The implementation of our method can be found using this url path: <https://github.com/ShazidAraf/PLANT-CELL-TRACKING>.

This article has supplementary downloadable material available at <https://doi.org/10.1109/TCBBIO.2024.3525404>, provided by the authors.

Digital Object Identifier 10.1109/TCBBIO.2024.3525404

division patterns in both plant and animal tissues. Tracking the development of these cells over time provides a quantitative description of cell growth and division characteristics [1], [2], [3], [4], gene expression [5], [6], enzyme/protein localization [7], [8], [9] and analyzing bio signals [10].

The plant of interest for this study is *Arabidopsis thaliana*. This plant is a genetic model system and a representative system to study shoot apical meristem (SAM) morphology of dicotyledonous plants. Thus, its study aids in understanding the morphological processes of several other plants as discussed by the authors of [11], [12]. In order to analyze the cell growth and division characteristics, this work focuses on the Shoot Apical Meristems (SAM) [13] of the plant. Shoot Apical Meristems (SAM) are densely packed multi-layer tissues that provide cells for the development of leaves, stems, and branches. With the advancement of microscopy imaging techniques, time-lapse images of long time series can be collected from SAM by Confocal Laser Scanning Microscopy (CLSM) [14] based live cell imaging. In this imaging technique, a laser beam is moved along the depth of the SAM tissue, which results in a series of two-dimensional images. The two-dimensional images are referred to as *slices*, which, when put together, forms the image *stack* of the SAM. Fig. 1 shows some exemplar slices of the SAM region of *Arabidopsis thaliana* captured by CLSM.

The Shoot Apical Meristems (SAM) of *Arabidopsis Thaliana* is a tightly packed structure with many cells. Spatio-temporal tracking of such a large number of cells is computationally expensive and time-consuming. In addition, the images captured from the deep layers of SAM suffer from low signal-to-noise ratio (SNR), as mentioned in [15], which makes the problem of tracking deep-layer cells very challenging. Further, the plant has a natural tendency of growing towards the light source it gets exposed to [16]. As a result, when a plant is kept in a non-uniform lighting setup, the plant leans toward the light source and gets tilted. The live cell imaging technique can not capture this information as it only captures two-dimensional images on X-Y plane (or, top view), thus further complicating the problem of tracking plant cells. Hence, there's a need for an automated, fast, and robust algorithm that can perform tracking of plant cells in Shoot Apical Meristems (SAM) of *Arabidopsis Thaliana*. In this paper, we propose a novel deep learning-based method,

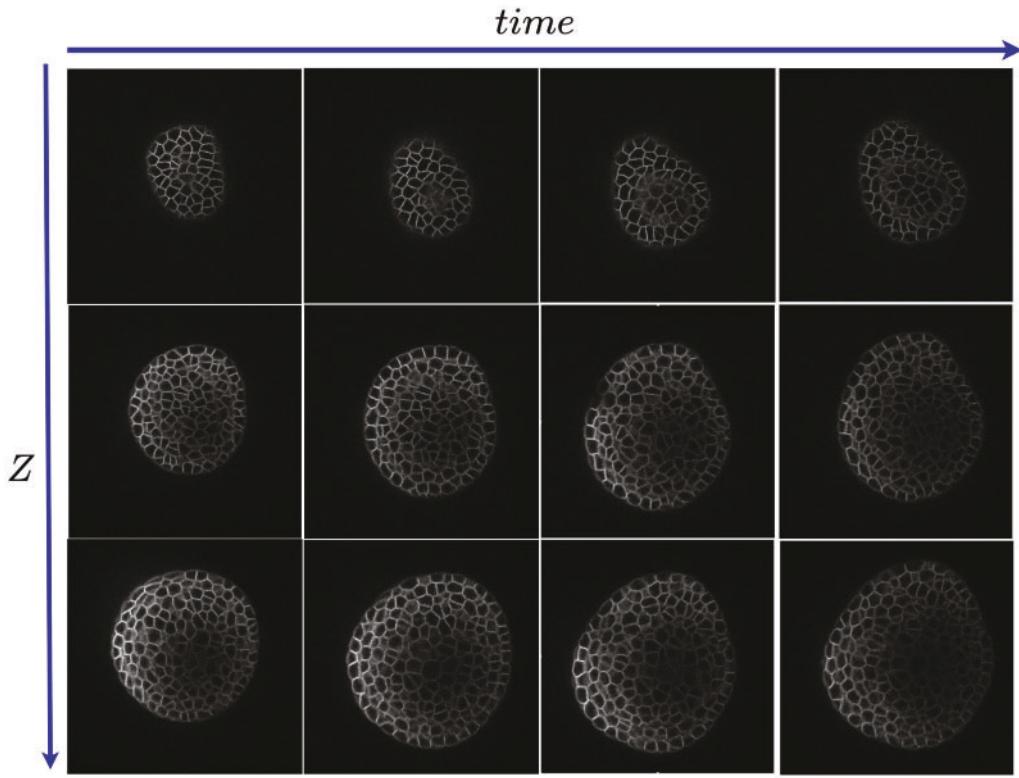


Fig. 1. The time series of microscopic plant cell images are shown. There is a time axis and a  $Z$ -axis.  $Z$ -axis which goes along the depth of the plant indicates the slice number and the time axis indicates the time passed during imaging. According to the dataset [4], time difference between two time points is 4 hours. On  $Z$ -axis, images of 3 slices are shown. Those slices are sampled from the top, middle, and bottom of the SAM, respectively.

## DEGAST3D: DEformable 3D GrAph Similarity to Track plant cells and detect cell division events.

### A. Related Work

There are several studies on plant cell tracking. Most of them are based on two-dimensional local-graph matching [17], [18], [19], [20], which exploits the tight spatial topology of neighboring cells. In those studies, a two-dimensional local star graph [21] is constructed by connecting the centroid of the cell of consideration to the centroids of its neighboring cells. This graph structure represents the neighboring structure of the cell of consideration. Two cells from different time points that have the closest graph structure are considered the corresponding cells (also known as matching cells). In these works, cell correspondences are done between a pair of slices of two consecutive time points, combined to obtain pairwise tracks over the entire stack. Chakraborty et al. [22] proposed a conditional random field (CRF) [23] based approach where all cells located on a slice are considered as nodes of the graph and all cells having common boundary are connected to form the graph. Marginal posteriors of each node are calculated using loopy belief propagation [24] followed by a graph labeling method to obtain the optimal correspondence.

Although this method gives very accurate results, it is not scalable for larger datasets as the graph labeling optimization step is fairly slow. A few recent works [25], [26] form a three-dimensional local graph of a cell by combining two-dimensional local star graphs of adjacent slices. However, these approaches

do not use the inter-slice connection among the cells, thus ignoring the three-dimensional spatial information of the cells. In addition, these works assume that two matching cells must have the same number of neighbors which makes these approaches very susceptible to segmentation error. Another recent method, DeepSeed [18], uses the weighted sum of shape similarity and neighboring structure similarity scores between the cells. The corresponding cell pairs (seed pairs) are then selected based on a similarity score threshold. This is followed by using relative position concerning the seed pairs to track the rest of the cells iteratively. Instead of using the shape similarity of individual cells, [27] compares the patches of cells of two different time points and uses the K-M algorithm [28] for patch association. However, these methods are highly dependent on segmentation accuracy and hence result in subpar performance for deeper layer cells of the plant.

Another limitation of the algorithms [18], [22], [27] is that they solve tracking problems only in a two-dimensional plane. This assumption ignores phenomena like the tilting of plants. Fig. 2 illustrates that when tilt occurs, the same cells change shape or size when the same cross-section is observed. As a result, the methods largely dependent on two-dimensional cell shape in tracking fail under tilting conditions of the SAM.

### B. Contributions

To address the limitations of existing approaches, our proposed method, **DEGAST3D**, makes the following contributions.

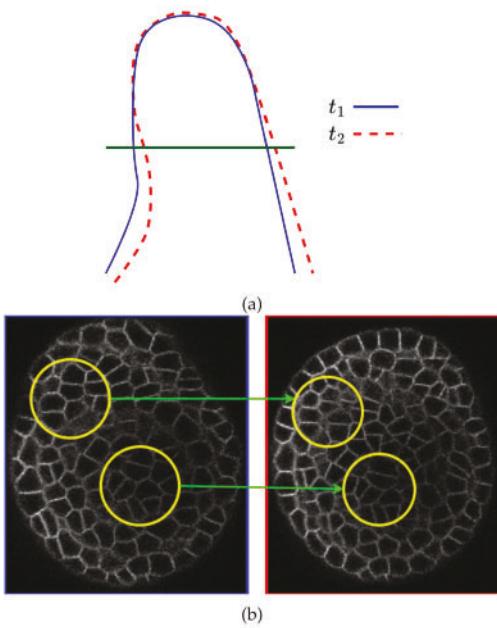


Fig. 2. (a) The overlay of SAM in two time-points when tilting occurs. The blue solid and red dotted lines indicate time  $t_1$  and  $t_2$ , respectively. (b) The cross-section view with respect to the green line in (a) is shown. The patches marked by yellow circles reveal that the shape and size of the same cells captured in the two-dimensional plane change due to tilt.

- We develop a novel learning-based tracking algorithm for plant cells that uses three-dimensional geometric information of the tightly packed plant cell structure.
- We develop a 3D registration technique that aligns the plant in two different time points.
- We propose a three-dimensional graph matching technique where graphs are connected to  $k$  nearest neighbors to extract contextual information of the cell structure, increasing tracking accuracy.
- We further develop a novel learning-based cell division detection technique that uses three-dimensional shape and local graph similarity to detect mother and daughter pairs.

The organization of this paper is as follows: Section II provides a detailed description of the proposed methodology. Experiments and results are discussed in Section III. Finally, Section IV concludes the paper.

## II. METHODOLOGY

This paper tackles the core problem of tracking SAM cells over long time periods. The inputs are unregistered plant image stacks of multiple time points with constant intervals. Over this time period, some cells divide. The objective is to track tightly packed plant cells and detect cell division events to determine cell lineages that develop over a period of time. In this section, we elaborate on our proposed pipeline for performing robust tracking of plant cells. First, three-dimensional segmentation of CLSM plant stack images is done. Then, three-dimensional registration aligns the SAM image stacks over time. Finally, a novel learning-based three-dimensional local graph matching technique is employed to find cell pair matches of tightly packed

plant cells and detect cell division events. Fig. 3 shows the entire workflow of our proposed algorithm.

### A. Segmentation

Cell segmentation is the first step for cell-tracking algorithms, which helps identify individual cells, present in the SAM tissue. State-of-the-art cell tracking methods such as [18], [22], [27] focused on watershed segmentation [29] technique that segments the cell border on every slice. However, watershed segmentation does not provide the three-dimensional spatial relationships of the cells among the slices. Recently, learning-based techniques such as spherical harmonics [30], U-Net watershed [31], Cellpose3D [32] have been proposed, which provide instances of every cell in three-dimensional space. Among these works, Cellpose3D [32] provides the best segmentation performance on the plant cells. Hence, this technique has been used in this paper. Cellpose3D learns the gradient map along the  $X$ ,  $Y$ , and  $Z$ -axis. Using that gradient map and some post-processing steps, it predicts the instances of the plant cells. Fig. 4 shows the application of Cellpose3D to get the instances of plant cells.

### B. Registration

Cell growth and division cause physiological changes in the structure of the plant, which result in slight orientation and shift along the  $X$ ,  $Y$ , and  $Z$  axes. Also, during the live imaging technique, the plant has to be physically moved between image acquisition time intervals. These incidents misalign the image stacks at different time points.

As plant cells are tightly packed, the positions of the cells do not change abruptly in a short time interval. Hence, if image stacks are aligned (or registered) properly, we can focus on a small region of interest in the image stack instead of searching the cell on the entire stack to track a particular cell. Thus, registration saves time and computational costs. Authors of [33] proposed a landmark-based registration algorithm that obtained state-of-the-art results on registering two-dimensional slices. However, the algorithm shows subpar performance where three-dimensional registration is needed (for example, the tilting case of Fig. 2). This motivates us to propose a novel 3D registration algorithm, an overview of which has been shown in Fig. 5.

Our proposed algorithm involves three steps. The first step obtains three-dimensional instance segmentation maps using pre-trained Cellpose3D models [32]. The second step obtains several high-confidence pairwise cell correspondences (also known as seed pairs) using the algorithm proposed in [18]. Finally, the Random Sample Consensus (RANSAC) algorithm [34] is used to determine the rotation and translation matrix from the point clouds of matching seed pair cells. The use of RANSAC is motivated by the fact that some false positive matches can be found in determining seed pairs in the second step. RANSAC can estimate the parameters of a mathematical model from a set of matching data containing false positive matches. The estimated rotation and translation matrix is used to perform registration between two plant image stacks. Fig. 6 shows the overlays of two-time points from the top before and after registration. However, the SAM is considered a rigid object in this registration

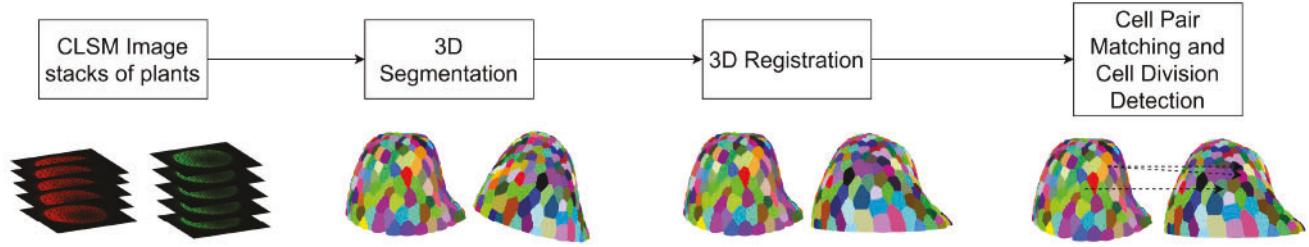


Fig. 3. The entire workflow is shown in the figure. Image stacks obtained by CLSM imaging are segmented using a 3D segmentation technique (Section II-A). Then, registration is done for pairwise time points using our proposed 3D registration method (Section II-B). Finally, our novel learning-based approach is used for cell pair matching (Section II-C) and cell division detection (Section II-D).

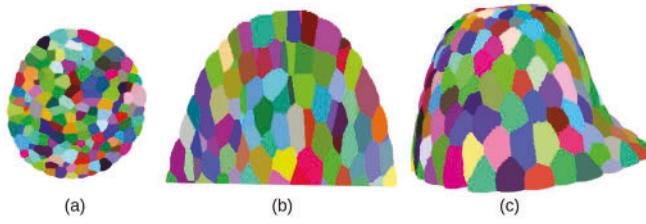


Fig. 4. Segmentation of plant cells using Cellpose3D. (a) Cell segmentation in a horizontal slice. (b) Cell segmentation on the vertical slice. (c) 3D reconstruction using the 3D instance segmentation.

technique, which is not practically true. We considered a region of interest for tracking cells described in the next sub-section to address this issue.

### C. Cell Pair Matching

The SAM cells are non-rigid due to cell growth and cell division. Hence, the segmentation map of a cell gets deformed as it grows over time. In addition, over-segmentation and under-segmentation may occur due to image noise. As a result, when we overlay the image stacks of two consecutive time points, the corresponding cells do not always overlap. To tackle this issue, we consider a region of interest to find a match for every cell. Let us consider a cell  $s_{t_1}$  of time point  $t_1$  which has centroid at coordinate  $c_{t_1} \in \mathbb{R}^3$ . There are total  $q$  number of cells  $s_{t_2}^{(1)}, s_{t_2}^{(2)}, \dots, s_{t_2}^{(q)}$  in the next time point  $t_2$  with their centroid coordinate at  $c_{t_2}^{(1)}, c_{t_2}^{(2)}, \dots, c_{t_2}^{(q)}$ , respectively.

In order to find the corresponding cell of  $s_{t_1}$  at time point  $t_2$ , we consider  $n$  number of candidate cells of set  $\mathcal{R} = \{s_{t_2}^{(1)}, s_{t_2}^{(2)}, \dots, s_{t_2}^{(n)}\}$  which have lowest distance from coordinate  $c_{t_1}$  among all other cells at time point  $t_2$ . Mathematically

$$\begin{aligned} \|c_{t_2}^{(1)} - c_{t_1}\| &\leq \|c_{t_2}^{(2)} - c_{t_1}\| \dots \leq \|c_{t_2}^{(n)} - c_{t_1}\| \dots \\ &\leq \|c_{t_2}^{(q)} - c_{t_1}\|. \end{aligned}$$

The region of interest comprises the cells in  $\mathcal{R}$  as shown in Fig. 7. We use a 3D geometric feature extractor [35] and propose a 3D graph matching algorithm to predict the corresponding cell of  $s_{t_1}$  out of those candidates in  $\mathcal{R}$ .

1) *3D Graph Matching*: The local graph for a particular cell comprises total  $n_p$  nodes where the centroids of that cell and its  $n_p - 1$  number of nearest neighbors in the three-dimensional

space are considered. All these nodes are connected to  $k$ -nearest neighbors, forming the local 3D graph as shown in Fig. 8. The formation of this type of graph helps us in two ways. First, our graph is capable of extracting enriched features in three-dimensional space. The local graphs presented in [18], [26] consider a two-dimensional star graph that mainly uses the positional information of neighboring cells with respect to the center cell only. On the other hand, our 3D local graph uses positional information with respect to not only the central cell but also the neighboring cells among themselves, which extracts rich three-dimensional spatial information. Second, our 3D local graph reduces computational cost compared to the graph proposed in [22] where the local graph is not formed. Rather, all the cells of a slice are considered nodes of the graph, and all nodes corresponding to the cells with common boundaries are connected. However, the formation of this kind of large graph makes the optimization computationally expensive and time-consuming. On the contrary, the 3D local graph is more inference-friendly with less computational expense.

We propose a learning-based 3D graph matching approach where the input is an anchor graph (the 3D local graph of the cell we want to track) and  $n$  number of local 3D graphs of potential matching candidate cells in the next time point. The objective is to point out the candidate graph which has the highest similarity to the anchor graph. Fig. 9 shows the graph similarity model where the inputs are the three-dimensional coordinates of two graphs. A geometric feature extractor [35] has been used in order to extract the three-dimensional geometric features of each node in a  $k$ -nearest neighbor connected graph. The outputs of the model are the point-to-point similarity score (shown as  $y_1$ ) and overall graph similarity score (shown as  $y_2$ ). As two graphs have the same number of nodes, we can make the one-to-one association of nodes between two graphs based on their distance. The point-to-point similarity score indicates how similar the node pairs are based on their spatial orientation. The overall graph similarity score indicates the overall similarity between two graphs.

2) *Training*: The 3D local graph matching involves joint training on both overall graph similarity and point-to-point similarity. Overall graph similarity indicates which candidate graph is the most similar to the anchor graph out of  $n$  candidates. This supervision can be obtained directly from the ground truth of the dataset where information regarding pairwise correspondent cells are provided.

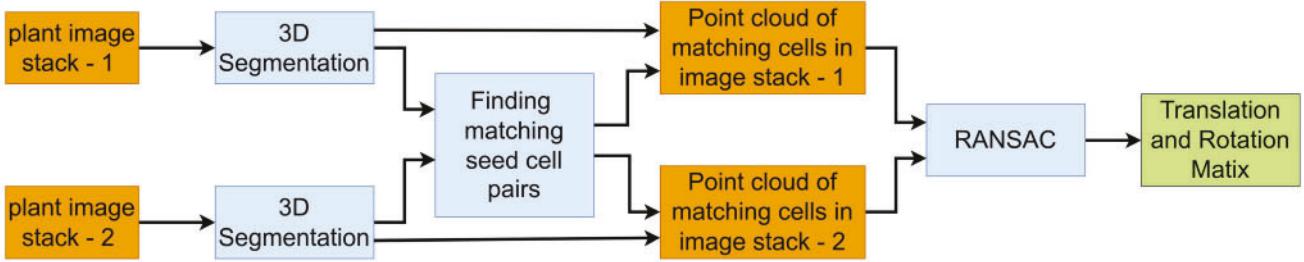


Fig. 5. Proposed method for registration of plant stacks. At first, the three-dimensional instance segmentation is obtained using Cellpose3D. Then, the number of correspondent cell pairs (seed pairs) is determined using the method described in [18]. The point cloud of seed pair cells of each stack is extracted. Using the FPFH (Fast Point Feature Histogram) feature of the point cloud, RANSAC algorithm is used to determine the translation and rotation matrix that is used for registration.

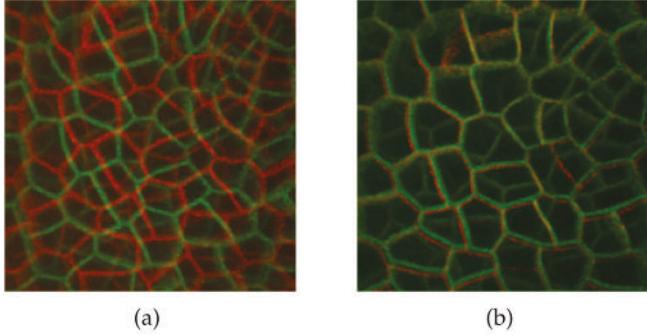


Fig. 6. Overlay of cell borders of plant image stacks are shown. The red and green channels indicate cell borders of two different time points of SAM. (a) When plant image stacks are unregistered, cell borders of two-time points are visible separately. (b) After registration, the cell borders of two-time points get overlapped in most of the cases.

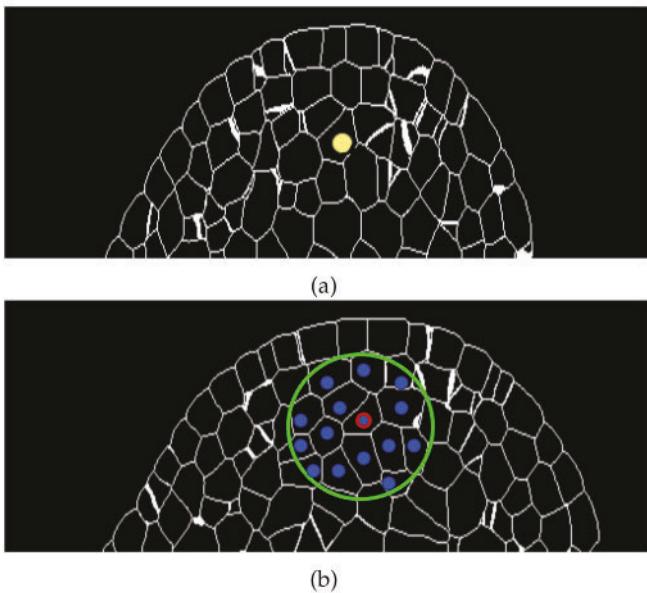


Fig. 7. Longitudinal section of plant for two consecutive time points where (a) and (b) indicate the former and later time point, respectively. In (a), the cell with the yellow circle is the cell that needs to be tracked. In (b), cells marked by the blue circle are the candidates of matching, and the cell with a red border is the true matching. The green circle indicates our region of interest.

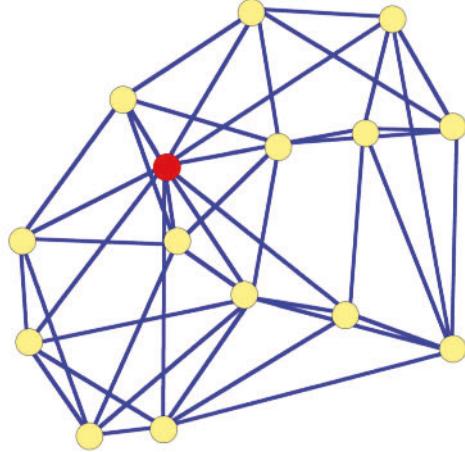


Fig. 8. A three-dimensional graph generated by connecting the  $k$ -nearest neighbors. The central cell is shown by the red circle, and its neighboring cells are shown by yellow circles. Every cell is connected to its  $k$ -nearest neighbors (in the figure  $k = 5$ ).

In the ground truth of our dataset, pairwise correspondences are not provided for all the cell pairs; instead, only a subset of cell pairs have them. Given a pair of correspondent cells in a tightly packed cell structure of two time points, we can assume that most of the neighbors of those correspondent cells exhibit one-to-one matching despite the matching information for all the neighboring cells is not provided in the ground truth. It leads to point-to-point similarity supervision, which is described in Fig. 10 in detail. Hence, training on point-to-point similarity is a semi-supervised process that involves pairwise node correspondence between two graphs.

Mathematically, there are  $n$  candidate graphs and  $(y_1^{(1)}, y_2^{(1)}), (y_1^{(2)}, y_2^{(2)}), \dots, (y_1^{(n)}, y_2^{(n)})$  are the outputs of the networks,  $p^{(1)}, p^{(2)}, \dots, p^{(n)}$  are the point-to-point similarity ground truth and  $GT_{graph}$  is the overall graph similarity ground truth.

Binary cross-entropy loss ( $BCE$ ) is used to train the point-to-point supervision. The point-to-point supervision loss ( $L_{pp}$ ) can be expressed by

$$L_{pp} = \sum_{k_1=1}^n \sum_{k_2=1}^{n_p} BCE \left( y_{1,k_2}^{(k_1)}, p_{k_2}^{(k_1)} \right). \quad (1)$$

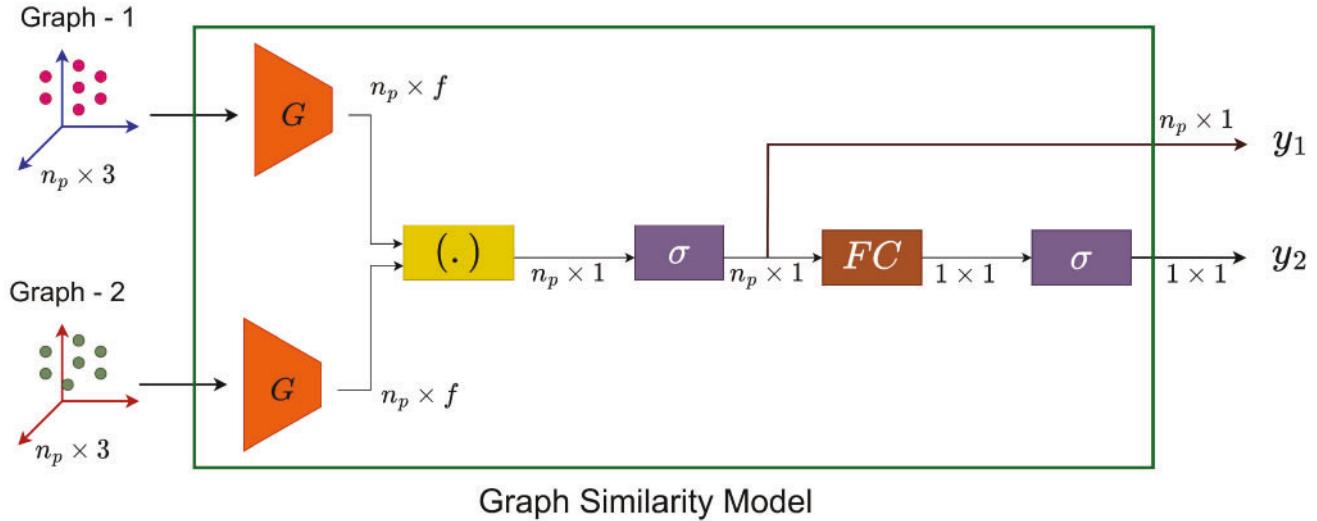


Fig. 9. Graph Similarity Model (GSM) shown by the green box. There are two inputs of this model: three-dimensional coordinates of two graphs where each graph has  $n_p$  number of points. The model provides two outputs. They are point-to-point similarity prediction ( $y_1$ ) for each point, and overall graph similarity prediction ( $y_2$ ). The block diagram  $G$ ,  $(\cdot)$ ,  $\sigma$ ,  $FC$  indicates geometric feature extractor, dot product, sigmoid activation function and fully connected layer, respectively. In the figure, the input and output shape of every block is shown.

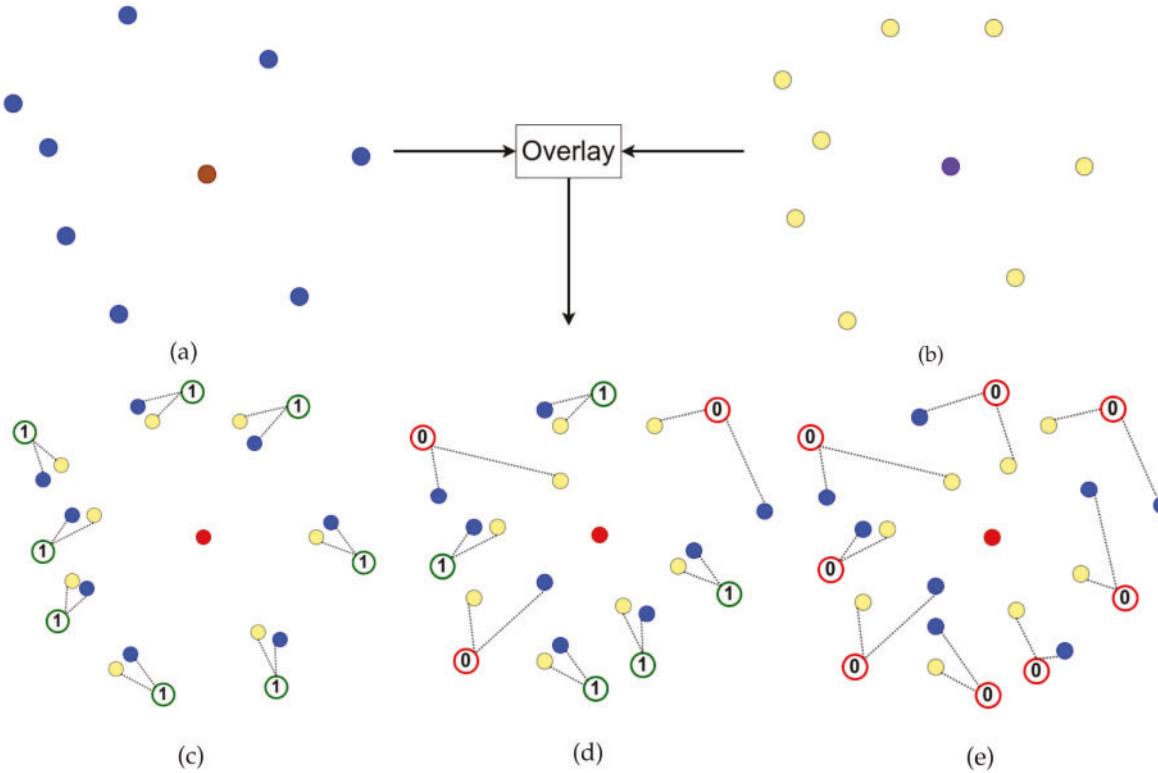


Fig. 10. This figure describes how point-to-point supervision is obtained. Every circle represents the centroid of a cell. (a) The brown and blue circles represent the center of the graph and its neighbors, respectively, at the one-time point. (b) The purple and yellow circles represent the center of the graph and its neighbors, respectively, at another time point. From the ground truth, we only know whether brown and purple cells are corresponding cells or not. The centers of two graphs (brown and purple circles) are overlaid at the red color circle, and there can be three cases as shown in (c)–(e). Case - 1 : (c) the cells are matching cells according to the ground truth, and two local graphs are less deformed. Point-to-point supervision is set to 1 for all point pairs (shown in the figure). Case 2: (d) the cells are matching cells, but local graphs are slightly different due to deformation. In this case, some point pairs are far apart. The point pairs with distances within a threshold value are set to 1. Otherwise, the ground truth is set to 0. Case -3 : (e) If the two cells are non-matching, all the pairwise point-to-point ground truths are set to 0.

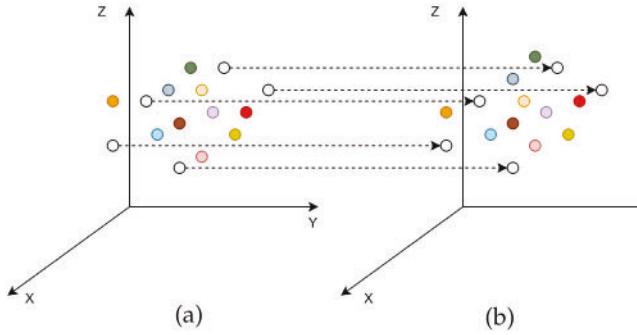


Fig. 11. (a) and (b) indicate the cells of former and later time points, respectively. The circles indicate centroids of the cells. Using the learning-based tracking method, most of the cells are tracked. The same colored cells indicate the matching pairs. However, the white color circles indicate the cells that our method could not track. Using the relative positions of the tracked cells (colored), we can associate the untracked cells (shown by the dotted arrow).

We note some cases where cell deformation or segmentation errors occur very much. In those cases, the region of interest shown in Fig. 7 does not always include the true matching cell. Hence, we considered an extra class  $y'$  for graph similarity supervision which indicates the anchor graph does not match with any candidate graph. The score for  $y'$  can be expressed by

$$y' = 1 - \max \left( y_2^{(1)}, y_2^{(2)}, \dots, y_2^{(n)} \right). \quad (2)$$

Hence, the anchor graph can match with any of the members of set  $Y_{\text{cand}} = \{y_2^{(1)}, y_2^{(2)}, \dots, y_2^{(n)}, y'\}$  which is guided by cross-entropy (CE) loss. The overall graph similarity supervision ( $L_g$ ) loss can be expressed by

$$L_g = CE(GT_{\text{graph}}, Y_{\text{cand}}). \quad (3)$$

Now the total loss can be written as

$$L_{\text{total}} = \lambda L_{\text{pp}} + L_g, \quad (4)$$

where  $\lambda$  is a non-negative hyper-parameter that controls the weight of point-to-point similarity supervision in the joint loss function.

The figure of overall training procedure is provided in the supplementary. Although this approach can find most of the matching cell pairs, some cells remain unmatched. Inspired by [18], we can track the unmatched cells using their relative positions with respect to matching pair cells (demonstrated in Fig. 11).

#### D. Cell Division Detection

Cell division [36] is the process by which a mother cell divides into two daughter cells. The state-of-the-art cell division detection techniques [37], [38] were based on two assumptions:

- The combined area of daughter cells is almost equal to the area of mother cells.
- The combined cell borders of daughter cells are almost identical to the mother cell's border.

However, the outcome from these assumptions is highly dependent on the segmentation accuracy, and it also does not consider cell growth over time. In this paper, we have introduced

a novel learning-based method for detecting cell division. This method uses two sources of information to detect cell division: the 3D shape similarity between mother and daughter cells and the 3D local graph similarity between mother and daughter cells. For every potential mother cell, a region of interest (same as Fig. 7) is selected in the next time point, and potential daughter cell pairs are selected by taking pairs of adjacent cells in that region of interest.

1) *3D Shape and Volume Similarity*: In three-dimensional space, the shape and volume of combined daughter cells are almost equal to the shape and volume of the mother cell. To use this physical property, we present a workflow shown in Fig. 12, which is totally data-driven. Voxelization [39] is done on the point clouds of mother and daughter cells. Then, subtraction is done between the voxelized mother cell and concatenated voxelized daughter cells. The subtracted result represents distinct patterns based on whether the input point clouds are from mother-daughter cell pairings. A 3D convolutional neural network [40] is trained to do binary classification between these two separate patterns, which is guided by a BCE loss.  $Y_{\text{div}}$  is the model output, which is a probability score on how similar the mother and daughter pairs are in a three-dimensional shape.  $GT_{\text{div}}$  is the ground truth for cell division. If the inputs are true, mother-daughters  $GT_{\text{div}}$  is 1, otherwise 0.

$$L_{\text{cell sim}} = BCE(Y_{\text{div}}, GT_{\text{div}}). \quad (5)$$

2) *Pairwise Local Graph Similarity*: Spatially, the mother and daughter cells occupy almost the same location in the SAM. Hence, the 3D local graph of the mother cell and each of the daughter cells are almost similar. To quantify how similar between two local 3D graphs, we train the network presented in Fig. 9 using the loss function

$$L_{\text{graph pair}} = \sum_{k=1}^{n_p} BCE(y_{1,k}, p_k) + BCE(y_2, GT_{\text{graph pair}}). \quad (6)$$

$GT_{\text{graph pair}}$  can be obtained from the ground truth where pairwise cell correspondences are provided. If a pair of cells are corresponding cells their 3D local graphs are similar and  $GT_{\text{graph pair}}$  is set to 1 otherwise, it is set to 0. The overall graph similarity score  $y_2$  predicts a score based on the similarity of two local 3D graph inputs.

Let  $m$  is the mother and  $(d_1, d_2)$  are daughter pairs. For each of the inputs of  $(m, d_1)$  and  $(m, d_2)$  the network of Fig. 9 predicts overall graph similarity score  $y_{2,d_1}$  and  $y_{2,d_2}$  respectively. The cell division score can be expressed by

$$Score_{\text{div}} = w_d Y_{\text{div}} + (1 - w_d) \frac{(y_{2,d_1} + y_{2,d_2})}{2}, \quad (7)$$

where  $w_d$  is a weighting factor for cell division which is set to 0.5. If the value of  $Score_{\text{div}}$  is higher than a threshold, then it is reasonable to assume that cell division occurred.

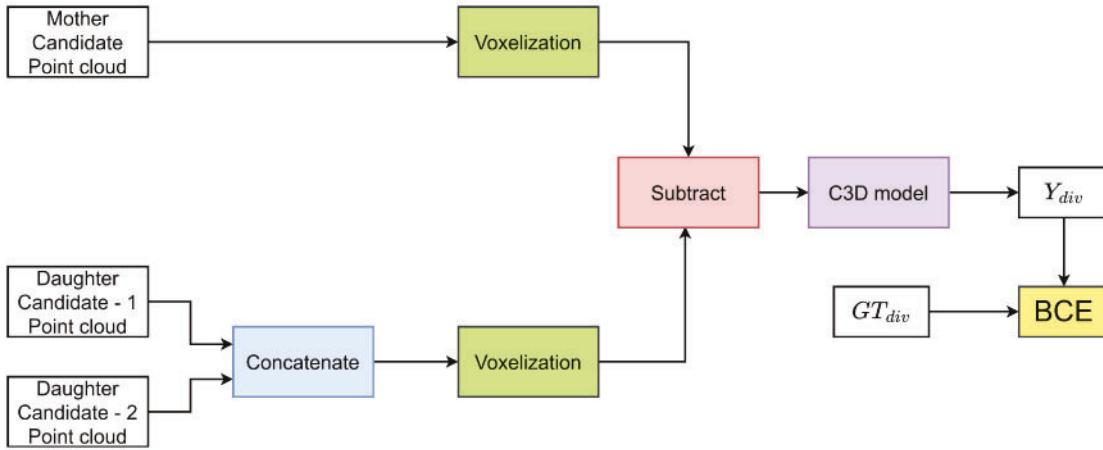


Fig. 12. Three-dimensional shape and volume similarity prediction between mother and combined daughter cells. Both the point cloud of mother cell and concatenated daughter cell pair is voxelized and subtracted. The subtracted result is trained by 3D convolutional model (C3D model). The training is guided by Binary Cross Entropy (BCE) loss.

### III. EXPERIMENTS AND RESULTS

#### A. Dataset

For the experiments, we have used a publicly available confocal imaging-based plant cell dataset [4] consisting of six plants. For each plant, on average, 20 time points data were provided with a gap of four hours between two consecutive time points. Each time point image stack has around 200 slices, and each slice has the size  $512 \times 512$  pixel. Two types of ground truth are provided with this dataset. One is for segmentation, and another is for tracking. In segmentation ground truth, instance segmentation of each cell was given. The tracking ground truth provides pairwise cell correspondence between two-time points.

#### B. Experimental Settings

Out of the available six plant data for cell tracking, four plants have been used for training, one for validation, and one for testing. We also engaged cross-validation to avoid bias. From the recent 3D segmentation techniques such as spherical harmonics [30], Cellpose3D [32], we have used Cellpose3D. The spherical harmonics approach does not preserve the polygonal shape of the cell, which makes this approach inappropriate for detecting cell division. On the other hand, Cellpose3D preserves the plant cell's polygonal shape, making this approach suitable for our work. The pre-trained model of Cellpose3D is publicly available, and according to [32], experiments on segmentation were done on the same dataset with the same data-split we have used.

#### C. Results

We compare the tracking performance of our proposed method with the method described in [22]. This method has used watershed segmentation [29], which segments the cell border of the plant cell. In our work, we have used Cellpose3D for segmentation. While comparing with [22], we have used the cell borders from the segmentation provided by Cellpose3D instead of using noisy watershed segmentation to make a fair comparison. In

TABLE I  
PERFORMANCE COMPARISON ON DIFFERENT METHODS FOR CELL PAIR  
MATCHING AND CELL DIVISION

Cell Pair Matching				
Method	Reg.	Precision	Recall	F1 Score
Chakraborty et al. [22]	2D [33]	0.7645	0.8118	0.7874
Chakraborty et al. [22]	3D (ours)	0.9209	0.8992	0.9099
DEGAST3D (Ours)	3D (ours)	<b>0.9894</b>	<b>0.9588</b>	<b>0.9739</b>
Cell Division				
Method	Reg.	Precision	Recall	F1 Score
Chakraborty et al. [22]	2D [33]	0.9302	0.2857	0.4372
Chakraborty et al. [22]	3D (ours)	<b>1.0000</b>	0.3521	0.5208
DEGAST3D (Ours)	3D (ours)	0.9862	<b>0.5059</b>	<b>0.6687</b>

addition, the method described in [22] uses landmark-based 2D registration [33], while we use 3D registration. Hence, we have provided a comparative results for both 2D and 3D registration.

1) *Performance of Cell Pair Matching*: We compare our novel method with [22] in Table I for pairwise cell matching. The comparison includes both the landmark-based two-dimensional registration [33] and proposed three-dimensional registration technique. The evaluation indices are precision, recall and F1 score. From Table I, We observe that proposed 3D registration has improved the performance of the method described in [22] compared to 2D registration. Overall, our method has secured 6.83%, 5.96% and, 6.40% improvement in terms of precision, recall and F1 score, respectively when 3D registration is considered. Fig. 13 shows some visual results of cell pair matching using our method.

2) *Performance on Cell Division Detection*: Table I shows the comparison of performances of our novel method with [22] for cell division. It is evident that our method outperforms the [22] (2D registration) in all metrics. When 3D registration is considered for [22], our method shows competitive results on precision and shows 15.38% and 14.78% improvement in recall and F1-score, respectively. Fig. 14 some visual results of cell division using our method.

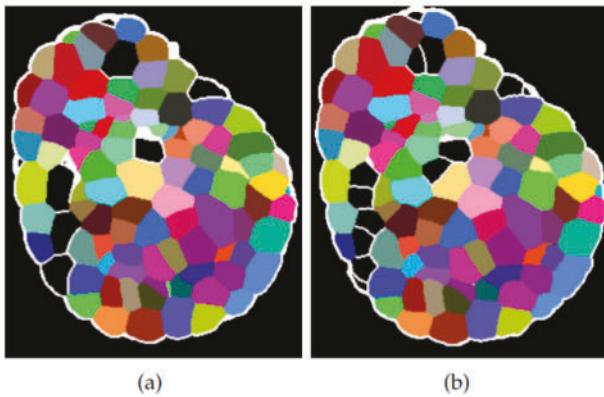


Fig. 13. (a) and (b) indicates the tracked cells of former and later time point, respectively. The same colored cells of these figures indicate the matching cell pairs.

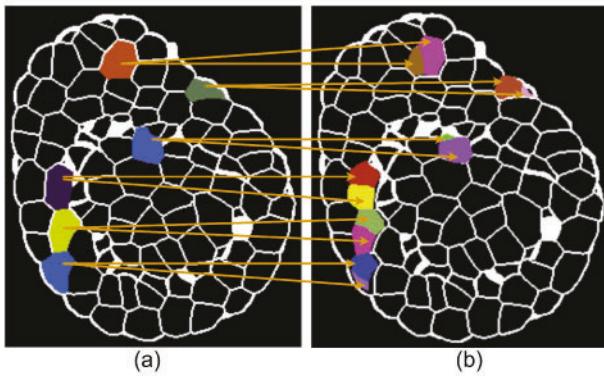


Fig. 14. Coloured cells of (a) indicate the mother cells and colored cells of (b) indicate the daughter cells. The arrow sign shows mother-daughter correspondence.

TABLE II  
PERFORMANCE COMPARISON ON DETERMINING LONG-TIME TRAJECTORY ACCURACY

Method	Registration	Accuracy (%)
Chakraborty et al. [22]	2D [33]	47.96
Chakraborty et al. [22]	3D (Ours)	58.29
DEGAST3D (Ours)	3D (Ours)	70.99

3) *Performance on 3D Registration:* Most of the previous works [18], [22], [27] are based on two-dimensional registration while we have proposed and implemented a 3D registration method. In Table I, we compare the performance of 2D registration [33] and our proposed 3D registration method for the same tracking approach. Our 3D registration significantly enhances the results of [22] in both cell pair matching and cell division detection. Specifically, for cell pair matching, precision, recall, and F1-score improve by 15.64%, 8.74%, and 12.25%, respectively. For cell division detection, precision, recall, and F1-score increase by 6.98%, 6.64%, and 8.36%, respectively.

4) *Cell Lineage Generation:* Inspired by [41], Xie et al. [27] proposed a node chaining method for plant cell trajectories, which we adopted for long-term lineage generation. Table II

TABLE III  
PERFORMANCE COMPARISON ON DIFFERENT SEGMENTATION TECHNIQUES

Cell Pair Matching			
Segmentation Method	Precision	Recall	F-1 Score
Cellpose3D [32]	0.9892	0.9588	0.9739
Spherical Harmonics (SH) [30]	0.9611	0.9715	0.9663
Unet Watershed (UW) [31]	0.9358	0.9113	0.9234
Cell Division			
Segmentation Method	Precision	Recall	F-1 Score
Cellpose3D [32]	0.9862	0.5059	0.6687
Spherical Harmonics (SH) [30]	1.0	0.0462	0.0882
Unet Watershed (UW) [31]	1.0	0.4125	0.5842

compares the accuracy of three methods for determining long-term cell lineages. We note that dataset [4] does not provide long-time lineages of all the cells. In the evaluation of long-time tracks, we have only considered those cells which have ground truth tracks of at least 16 hours. Accuracy, measured as the percentage of experimental lineages matching ground truth, shows our method outperforming the others by 23.03 % and 12.96 %, respectively.

#### D. Effects of Errors in Segmentation

The quality of segmentation can be judged in two aspects. First, whether the segmentation technique can detect cells correctly, and second, whether the segmentation technique preserves the shape of the cells. We used Cellpose3D [32] as a 3D segmentation method of plant cells. There are other 3D segmentation techniques such as Spherical Harmonics (SH) [30] and Unet Watershed (UW) [31]. The SH method, while avoiding under-segmentation, fails to preserve cell shapes. In contrast, UW is more effective at retaining cell shapes but suffers from under-segmentation, leading to fewer detected cells compared to the ground truth. Cellpose3D outperforms both SH and UW in terms of cell detection and shape representation. Our method focuses on two objectives: cell pair matching and cell division detection. Cell pair matching only uses the centroids of the cells in 3D, no shape information is utilized here. Conversely, cell division detection is highly dependent on accurate cell shape representation. As shown in Table III, SH segmentation exhibits a massive performance drop in cell division detection (Recall and F-1 Score) due to its failure in preserving cell shapes. SH segmentation exhibits only a slight performance drop in cell pair matching, as it is nearly as effective as Cellpose3D in detecting cells. Meanwhile, UW outperforms SH in cell division detection due to its shape-preserving characteristics but shows the worst performance in cell pair matching because of under-segmentation. Cellpose3D is efficient in both cell detection and shape preservation and so it exhibits the best results for both cell pair matching and cell division detection. In summary, the performance of DEGAST3D improves with higher-quality segmentation. While the algorithm remains robust in cell pair matching even without shape preservation, accurate cell division detection necessitates shape-preserving segmentation.

TABLE IV  
PERFORMANCE COMPARISON OF CELL PAIR MATCHING IN 2D SLICE TRACKING

Method	Precision	Recall	F-1 Score
Chakraborty et al. [22]	0.9209	0.8992	0.9099
Liu et al. [18]	0.4461	0.5910	0.5085
Liu et al. [27]	0.3730	0.4789	0.4194
DEGAST3D (ours)	<b>0.9461</b>	<b>0.9114</b>	<b>0.9285</b>

TABLE V  
EFFECT OF CHANGING THE NUMBER OF COLLECTIONS WITH NEIGHBORS ( $k$ )

$k$	One-step Tracking			Two-step Tracking		
	Prec.	Recall	F-1 Sc.	Prec.	Recall	F-1 Sc.
4	0.9801	0.9074	0.9423	0.9781	0.9567	0.9673
5	0.9880	0.9130	0.9491	0.9804	0.9575	0.9688
6	<b>0.9909</b>	<b>0.9219</b>	<b>0.9551</b>	<b>0.9894</b>	<b>0.9588</b>	<b>0.9739</b>

#### E. Performance With 2D Single Slice-Based Tracking

Most of the previous works [18], [22], [27] consider 2D star graph in order to perform cell pair matching. On the other hand we consider 3D  $k$ -nearest neighbor connected graph for the same purpose. In Table IV we investigate how our algorithm performs if we consider only 2D slices for cell pair matching without considering any 3D structure. To be specific we want to evaluate how does 2D  $k$ -nearest neighbor connected graph perform with respect to 2D star graph based methods. We consider that all the methods incorporate same registration method. We observe from Table IV our approach outperforms all the other methods in two-dimensional slice-by-slice cell matching. It indicates the superiority of proposed cell-matching approach over existing approaches.

#### F. Ablation Study

The proposed cell pair matching method has two steps. In the first step, we find pairwise cell matching using a deep neural network. However, some cells are left unmatched in this step. In the second step, we track those untracked cells using their relative position with respect to the tracked cells. Clearly, in the second step, no learning-based approach is involved. In our ablation study, we show the comparison between one-step and two-step tracking while changing different factors of tracking, such as the number of connections to neighbors in graph formation, the size of the region of interest, and the impact of modified loss function during training.

1) *Effect of Changing the Number of Connections With Neighbors in Graph Formation:* Given the centroids of  $n_p$  number of cells, we form the graph by joining  $k$  nearest points among themselves. In Table V, we show how the results of cell matching change with different values of  $k$ . We note that the precision value increases with the increase of  $k$ . Because an increase of  $k$  means more information is used regarding the spatial orientation of the nodes. Hence, cell matching accuracy increases with the increase of  $k$ . In addition, while using one-step tracking, some cells are left untracked. Some untracked cells can find their matching pairs in two-step tracking. As a result, we observe an increase in recall value in two-step tracking.

TABLE VI  
EFFECT OF CHANGING THE NUMBER OF MATCHING CANDIDATES ( $n$ )

$n$	One-step Tracking			Two-step Tracking		
	Precision	Recall	F1	Precision	Recall	F1
5	<b>0.9978</b>	0.8887	0.9401	<b>0.9937</b>	0.9694	0.9814
10	0.9858	0.9058	0.9441	0.9819	0.9503	0.9658
20	0.9909	<b>0.9219</b>	<b>0.9551</b>	0.9894	<b>0.9588</b>	<b>0.9739</b>

TABLE VII  
IMPACT OF INDIVIDUAL LOSS TERMS

Loss	One-step Tracking		Two-step Tracking	
	Precision	Recall	Precision	Recall
$L_{pp}$	0.9408	0.3605	0.8305	0.8819
$L_g$	0.9394	0.6928	0.9316	0.9132
$L_{pp} + L_g$	<b>0.9909</b>	<b>0.9219</b>	<b>0.9894</b>	<b>0.9588</b>

2) *Effect of Changing the Size of the Region of Interest:* The size of the region of interest is related to  $n$ , which indicates how many candidate graphs are compared against the anchor graph. Table VI shows when  $n$  increases, the precision value decreases while the recall value increases in one-step tracking. It is because when  $n$  is lower, the network learns to classify for fewer classes, which is easier. Hence, the precision value gets higher when a low value of  $n$  is used. On the other hand, with the lower value of  $n$ , the region of interest does not always cover the true matching cell. As a result, the number of unmatched cells increases, lowering the recall value.

3) *Effect of Individual Loss Terms:* Two losses guided the learning-based 3D graph matching. They are point-to-point similarity supervision loss ( $L_{pp}$ ) and overall graph similarity supervision loss ( $L_g$ ). In Table VII, we demonstrate the importance of both of the loss terms. If we train the network with only one loss term, all metrics show lower values in one- and two-step tracking. Hence, both of the loss terms are necessary for effective cell tracking.

#### G. Scalability to Larger Dataset

Larger datasets can be addressed in two aspects: the increased number of average cells per plant and the increased number of time points. If the region of interest (ROI) is fixed, the running time of the algorithm is linearly proportional to the number of average cells per plant and time points. Additionally, if we consider  $n$  number of cells in the ROI, the cell pair matching and cell division detection have time complexity of  $O(n)$  and  $O(n^2)$ , respectively. However, our efficient 3D registration, combined with the densely packed nature of plant cells, enables the selection of a small ROI rather than processing the entire stack. Consequently, a small value of  $n$  is sufficient for cell pair matching and detecting cell division. Hence, it is possible of scale our approach to larger datasets keeping the size of ROI ( $n$  cells) unchanged.

## IV. CONCLUSION

This paper proposes a novel learning-based method to automatically track plant cells in any three-dimensional,

unregistered condition. Unlike the state-of-the-art methods where tracking was done by two-dimensional graph matching, our proposed method constructs a three-dimensional graph that extracts the tight spatial features in a more sophisticated way, taking cell growth and deformity into consideration. The method also introduces a unique learning-based cell division detection algorithm, utilizing volumetric and structural data for improved accuracy. Additionally, the proposed 3D registration method can be applied to any computational biology related application. DEGAST3D's robustness against segmentation quality to track cells, scalability to larger datasets, and ability to handle 3D unregistered conditions make it a versatile tool for computational biology, particularly in scenarios involving tightly packed cell structures.

## REFERENCES

- [1] K.-K. Liu, C.-C. Wang, C.-L. Cheng, and J.-I. Chao, "Endocytic carboxylated nanodiamond for the labeling and tracking of cell division and differentiation in cancer and stem cells," *Biomaterials*, vol. 30, no. 26, pp. 4249–4259, 2009.
- [2] G. V. Reddy, M. G. Heisler, D. W. Ehrhardt, and E. M. Meyerowitz, "Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of *arabidopsis thaliana*," *Development*, vol. 131, pp. 4225–4237, 2004.
- [3] C. Michael et al., "Role of turgor-pressure induced boundary tension in the maintenance of the shoot apical meristem of *Arabidopsis thaliana*," *J. Roy. Soc. Interface*, vol. 20, no. 203, 2023, Art. no. 20230173.
- [4] L. Willis et al., "Cell size and growth regulation in the *Arabidopsis thaliana* apical stem cell niche," *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 51, pp. E8238–E8246, 2016.
- [5] H. Shen et al., "Automated tracking of gene expression in individual cells and cell compartments," *J. Roy. Soc. Interface*, vol. 3, no. 11, pp. 787–794, 2006.
- [6] M. Xie, M. Tataw, and G. V. Reddy, "Towards a functional understanding of cell growth dynamics in shoot meristem stem-cell niche," in *Seminars in Cell & Developmental Biology*, vol. 20, Amsterdam, Netherlands: Elsevier, 2009, pp. 1126–1133.
- [7] S. C. Fry, "Primary cell wall metabolism: Tracking the careers of wall polymers in living plant cells," *New Phytologist*, vol. 161, no. 3, pp. 641–675, 2004.
- [8] K. Rodriguez et al., "Concentration-dependent transcriptional switching through a collective action of cis-elements," *Sci. Adv.*, vol. 8, no. 31, 2022, Art. no. eab06157.
- [9] A. Plong, K. Rodriguez, M. Alber, W. Chen, and G. V. Reddy, "CLAVATA3 mediated simultaneous control of transcriptional and post-translational processes provides robustness to the WUSCHEL gradient," *Nature Commun.*, vol. 12, no. 1, 2021, Art. no. 6361.
- [10] M. Banwarth-Kuhn et al., "Combined computational modeling and experimental analysis integrating chemical and mechanical signals suggests possible mechanism of shoot meristem maintenance," *PLoS Comput. Biol.*, vol. 18, no. 6, 2022, Art. no. e1010199.
- [11] K. Koch, B. Bhushan, and W. Barthlott, "Diversity of structure, morphology and wetting of plant surfaces," *Soft Matter*, vol. 4, no. 10, pp. 1943–1963, 2008.
- [12] H. Zhang, H. Rong, and D. Pilbeam, "Signalling mechanisms underlying the morphological responses of the root system to nitrogen in *Arabidopsis thaliana*," *J. Exp. Botany*, vol. 58, no. 9, pp. 2329–2338, 2007.
- [13] E. Shani, O. Yanai, and N. Ori, "The role of hormones in shoot apical meristem function," *Curr. Opin. Plant Biol.*, vol. 9, no. 5, pp. 484–489, 2006.
- [14] S. W. Paddock, "Confocal laser scanning microscopy," *Biotechniques*, vol. 27, no. 5, pp. 992–1004, 1999.
- [15] A. Gupta, A. Aich, K. Rodriguez, G. V. Reddy, and A. K. Roy-Chowdhury, "Deep quantized representation for enhanced reconstruction," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. Workshops*, 2020, pp. 1–4.
- [16] B. M. Pollan, "The intelligent plant," *New Yorker*, p. 93, 2013.
- [17] W. Jiang, L. Wu, S. Liu, and M. Liu, "CNN-based two-stage cell segmentation improves plant cell tracking," *Pattern Recognit. Lett.*, vol. 128, pp. 311–317, 2019.
- [18] M. Liu, Y. Liu, W. Qian, and Y. Wang, "DeepSeed local graph matching for densely packed cells tracking," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 3, pp. 1060–1069, May/Jun. 2021.
- [19] M. Liu, A. K. Roy-Chowdhury, and G. V. Reddy, "Robust estimation of stem cell lineages using local graph matching," in *Proc. 2009 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, 2009, pp. 194–201.
- [20] M. Liu, P. Xiang, and G. Liu, "Robust plant cell tracking using local spatio-temporal context," *Neurocomputing*, vol. 208, pp. 309–314, 2016.
- [21] V. E. Menda and D. Sarkar, "Optimal broadcasting on the star graph," *IEEE Trans. Parallel Distrib. Syst.*, vol. 3, no. 04, pp. 389–396, Jul. 1992.
- [22] A. Chakraborty and A. K. Roy-Chowdhury, "Context aware spatio-temporal cell tracking in densely packed multilayer tissues," *Med. Image Anal.*, vol. 19, no. 1, pp. 149–163, 2015.
- [23] C. Sutton et al., "An introduction to conditional random fields," *Found. Trends Mach. Learn.*, vol. 4, no. 4, pp. 267–373, 2012.
- [24] A. T. Ihler, J. W. Fisher III, A. S. Willsky, and D. M. Chickering, "Loopy belief propagation: Convergence and effects of message errors," *J. Mach. Learn. Res.*, vol. 6, no. 5, pp. 905–936, 2005.
- [25] M. Liu, Y. He, W. Qian, Y. Wei, and X. Liu, "Cell population tracking in a honeycomb structure using an IMM filter based 3D local graph matching model," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 15, no. 5, pp. 1706–1717, Sep/Oct. 2018.
- [26] M. Liu, Y. Liu, and W. Qian, "A multi-seed 3D local graph matching model for tracking of densely packed cells," in *Proc. 2018 IEEE Int. Conf. Acoust. Speech Signal Process.*, 2018, pp. 1075–1079.
- [27] Y. Xie, M. Liu, S. Zhou, and Y. Wang, "A deep local patch matching network for cell tracking in microscopy image sequences without registration," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 19, no. 6, pp. 3202–3212, Nov./Dec. 2022.
- [28] H. Zhu, D. Liu, S. Zhang, Y. Zhu, L. Teng, and S. Teng, "Solving the many to many assignment problem by improving the Kuhn–Munkres algorithm with backtracking," *Theor. Comput. Sci.*, vol. 618, pp. 30–41, 2016.
- [29] S. Beucher and F. Meyer, "The morphological approach to segmentation: The watershed transformation," *Math. Morphol. Image Process.*, vol. 34, no. 1993, p. 433–481, 1993.
- [30] D. Eschweiler, M. Rethwisch, S. Koppers, and J. Stegmaier, "Spherical harmonics for shape-constrained 3D cell segmentation," in *Proc. IEEE 18th Int. Symp. Biomed. Imag.*, 2021, pp. 792–796.
- [31] D. Eschweiler, T. V. Spina, R. C. Choudhury, E. Meyerowitz, A. Cunha, and J. Stegmaier, "CNN-based preprocessing to optimize watershed-based cell segmentation in 3D confocal microscopy images," in *Proc. IEEE 16th Int. Symp. Biomed. Imag.*, 2019, pp. 223–227.
- [32] D. Eschweiler, R. S. Smith, and J. Stegmaier, "Robust 3D cell segmentation: Extending the view of cellpose," in *Proc. 2022 IEEE Int. Conf. Image Process.*, 2022, pp. 191–195.
- [33] K. Mkrtchyan, A. Chakraborty, and A. K. Roy-Chowdhury, "Automated registration of live imaging stacks of *Arabidopsis*," in *Proc. IEEE 10th Int. Symp. Biomed. Imag.*, 2013, pp. 672–675.
- [34] R. B. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (FPFH) for 3D registration," in *Proc. 2009 IEEE Int. Conf. Robot. Autom.*, 2009, pp. 3212–3217.
- [35] Z. Zhang and W. S. Lee, "Deep graphical feature learning for the feature matching problem," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 5087–5096.
- [36] D. Mazia, "Mitosis and the physiology of cell division," in *The Cell*, Amsterdam, Netherlands: Elsevier, 1961, pp. 77–412.
- [37] S. Basu, D. Bhattacharya, and R. Banerjee, "Mapping the distribution of packing topologies within protein interiors shows predominant preference for specific packing motifs," *BMC Bioinf.*, vol. 12, no. 1, pp. 1–26, 2011.
- [38] A.-A. Liu, J. Tang, W. Nie, and Y. Su, "Multi-grained random fields for mitosis identification in time-lapse phase contrast microscopy image sequences," *IEEE Trans. Med. Imag.*, vol. 36, no. 8, pp. 1699–1710, Aug. 2017.
- [39] Q. Chen, L. Sun, E. Cheung, and A. L. Yuille, "Every view counts: Cross-view consistency in 3D object detection with hybrid-cylindrical-spherical voxelization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 21224–21235.
- [40] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning spatiotemporal features with 3D convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 4489–4497.
- [41] J. Peng et al., "Chained-tracker: Chaining paired attentive regression results for end-to-end joint multiple-object detection and tracking," in *Proc. 16th Eur. Conf. Comput. Vis.*, Glasgow, U.K., Springer, 2020, pp. 145–161.



**Md Shazid Islam** (Graduate Student Member, IEEE) received the bachelor's and master's degrees in electrical and electronic engineering from the Bangladesh University of Engineering and Technology. Currently, he is working toward the PhD degree with the Department of Electrical and Computer Engineering, University of California, Riverside. His research interests include computer vision with a focus on domain adaptation, segmentation, and tracking.



**Christian Michael** received the PhD degree in mathematics from the University of California Riverside. Currently, he is working as a postdoctoral researcher with the University of Michigan. His research interest includes computational biology, mathematical modelling of virtual TB hosts.



**Arindam Dutta** received the bachelor's degree from NIT Durgapur and master's degree from IISC Bangalore. Currently, he is working toward the PhD degree with the Department of Electrical and Computer Engineering, University of California, Riverside. His research interest include domain adaptation, pose estimation and semantic segmentation.



**Mark Alber** received the PhD degree in mathematics from the University of Pennsylvania. He is currently a distinguished professor with the Department of Mathematics and director of the Center for Quantitative Modeling in Biology with the University of California, Riverside. He was elected a fellow of the American Association for the Advancement of Science (AAAS), and received the 2024 US Fulbright Scholar Award, The Netherlands. He currently serves as a Section editor in *Systems Biology of PLoS Computational Biology* and an associate editor of *Bulletin of Mathematical Biology*. His research interests include combined experimental and computational multi-scale modeling study of blood clot formation, tissue maintenance in plants and epithelial tissue growth.



**Calvin-Khang Ta** received the bachelor's degree in computer science from the University of California, Riverside. Currently, he is working toward the PhD degree in computer science with the University of California, Riverside. His research interests broadly include computer vision and machine learning with a focus on image restoration, pose estimation and vision-language models.



**G. Venugopala Reddy** received the PhD degree from the Tata Institute of Fundamental Research, India focused on Developmental Neurobiology and post-doctoral training on Plant Development with Caltech. He is currently a professor with the Department of Botany and Plant Sciences, University of California, Riverside. His broad research interests are on the regulatory aspects of cell-cell communication and transcriptional dynamics in the shoot apical meristem.



**Kevin Rodriguez** received the bachelor's degree in biology and the PhD degree in cell, molecular and developmental biology from the University of California, Riverside. Currently he is an assistant professor of biology with The University of the South. His broad research interest focuses on the molecular mechanisms that regulate stem cell maintenance. He investigates gene expression and protein dynamics critical to cell-cell communication.



**Amit K. Roy-Chowdhury** (Fellow, IEEE) received the PhD degree from the University of Maryland, College Park (UMCP), in 2002 and joined the University of California, Riverside (UCR), in 2004 where he is a professor and UC presidential chair of Electrical and Computer Engineering, cooperating faculty in computer science and engineering, and co-director of the UC Riverside AI Research and Education Institute. He leads the Video Computing Group with UCR, working on foundational principles of computer vision, image processing, and machine learning, with applications in cyber-physical, autonomous and intelligent systems. He has published more than 250 papers in peer-reviewed journals and conferences and two monographs. He is a fellow of the IAPR, received the Doctoral Dissertation Advising/Mentoring Award from UCR, and the ECE Distinguished Alumni Award from UMCP.