Topology-aware method to segment 3D plant tissues images.

Minh Ôn Vũ Ngọc
EPITA Research Laboratory
Le Kremlin-Bicetre, France

Nicolas Boutry
EPITA Research Laboratory
Le Kremlin-Bicetre, France

Jonathan Fabrizio
EPITA Research Laboratory
Le Kremlin-Bicetre, France

Abstract

The study of genetic and molecular mechanisms underlying tissue morphogenesis has received a lot of attention in biology. Especially, accurate segmentation of tissues into individual cells plays an important role for quantitative analyzing the development of the growing organs. However, instance cell segmentation is still a challenging task due to the quality of the image and the fine-scale structure. Any small leakage in the boundary prediction can merge different cells together, thereby damaging the global structure of the image. In this paper, we propose an end-to-end topology-aware 3D segmentation method for plant tissues. The strength of the method is that it takes care of the 3D topology of segmented structures. The keystone is a training phase and a new topology-aware loss - the *CavityLoss* - that are able to help the network to focus on the topological errors to fix them during the learning phase. The evaluation of our method on both fixed and live plant organ datasets shows that our method outperforms state-of-the-art methods (and contrary to state-of-the-art methods, does not require any post-processing stage). The code of CavityLoss is freely available at https://github.com/onvungocminh/CavityLoss.

1 Introduction

Learning structural biology from bioimages is a hot topic during the last decade (1; 2; 3; 4). Especially, segmentation of individual cells is a key step in the study of morphogenesis in a multicellular organism, i.e. counting the cell number of ovule primordia (5), or analysis the asymmetric division of lateral root founder cells (6). However, volumetric segmentation is still a challenging task and may cause some problems for the downstream analysis of the development of the plants. In the literature, most of the cell segmentation methods are based on their boundary prediction (5; 4; 7). Accurate boundary prediction is extremely important because neighboring cells share mutual information on the boundary location. Thus, even a small leakage on the boundary can merge adjacent regions and damage the global structure.

In light of the very successful trend toward deep learning-based image analysis, various methods have been proposed to leverage the ability of the convolutional neuron networks (CNN) to learn better image representations, thus improving the performance of the boundary prediction (7; 8; 9). In particular, U-Net (10), which consists of an encoder-decoder architecture, is proved to be useful for biomedical images and has later been further extended to process volumetric data (11). Although these methods achieve high accuracy, they do not contain any geometrical constraints. As a consequence, they usually provide discontinuities and broken connections in the segmentation results. Moreover,

NeurIPS 2022 AI for Science Workshop.

due to the limitation of the receptive field of the convolutional neural network (12), it can not capture global information of the image.

To address the problem of boundary prediction, several post-processing methods have been proposed to support grouping voxels into cells. Particularly, the watershed algorithm (13; 14; 15), which is a well-known method in the field of mathematical morphology, has been used to produce the segmentation result from the delineated boundaries. However, the drawback of the watershed algorithm is when the seed pixels are missing or misclassified. In (16), the author proposed a learned watershed algorithm in a reinforcement learning (17) style. Nevertheless, this method is greedy in terms of training data and computing time. Another problem with post-processing methods is that they do not integrate into the model as an end-to-end framework, which may slow down the segmentation process.

Recently, several topology-aware segmentation methods have been proposed to preserve the global structure of the image (18; 19; 20). Commonly, the networks in these methods are trained using a specific loss - a topological loss - that takes care of the topology of the structures. A topological loss should satisfy the two following conditions: (1) tolerate with a minor displacement of the boundary; (2) penalize the topological errors. The idea of topological loss started in (18), then developed in Topoloss (19), BALoss (20). However, the detected topological errors from these methods are usually neither relevant nor accurate. Nevertheless, these methods are mostly designed for 2D images and do not guarantee to preserve the topology. Especially, a new connectivity-preserving metric clDice (21) has been proposed leveraging the homotopy equivalent between the prediction and its skeleton. However, clDice treats equally every pixel on the skeleton, thus does not adequately penalize the broken connections. Therefore, we need a powerful method which is able to maintain the global structure for 3D image segmentation.

In order to tackle the aforementioned topology-preserving problem and provide a better solution to localize the broken connections in object boundaries, we propose a new method based on a 3D network. To train this network, we propose a new topological loss function called CavityLoss that pays attention to the structure of the image. Our method identifies topological errors in an efficient and accurate manner while we attempt to reconstruct the ground truth image from the prediction. Indeed, we leverage the efficiency of mathematical morphology in the general and watershed algorithm (13), in particular, to correctly detect the leakages on the 3D boundary locations. Penalizing these topological errors encourages the network to learn the connectivity of neighboring voxels on the boundary and also boosts up the value of the broken locations while training the network. The contributions of our method are following:

- We propose a topology-aware method for 3D segmentation of plant tissues using a CNN.
- To train this network, we propose a new topological loss (CavityLoss) to penalize the broken connections and encourage the network to improve the segmentation performance.
- To penalize the broken connections, we propose a new method inspired by mathematical morphology to efficiently localize topological errors in the boundary of the prediction.

2 Proposed Method

2.1 Overview

State-of-the-art methods still generate discontinuities and broken connections in the 3D segmentations due to the fact that these methods do not take into account the underlying topology of the image. In this paper, we present here a simple-yet-effective method, which integrates the geometrical information of the structures using the mathematical morphology method into the convolutional neural network. We aim to preserve the topology and voxel connectivity on the cell boundary. In particular, our method focuses on detecting and penalizing the topological errors in the boundary of regions.

Given an input image I and a ground truth image gt, our goal is to train the network to produce the predicted map u that has the same topology as the ground truth gt. In other words, the predicted map u has the same Betti numbers[1] as the ground truth. Especially in this article, for the task of 3D

[1]Betti numbers: β_0 represents the number of distinct connected components, β_1 represents the number of holes, and β_2 represents the number of cavities (only in 3D)

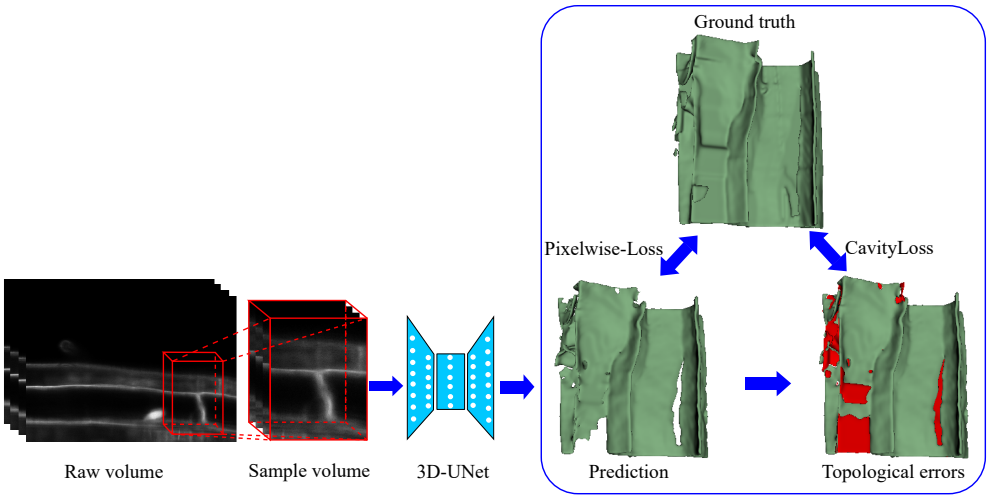


Figure 1: An overview of our proposed method. The boundary prediction which is the output of the network still generates a lot of broken connections (red zones). The topological errors are detected and our CavityLoss is computed to enforce the network correcting the leakage.

segmentation of plant tissues into individual cells, we focus on the number of cavities (β_2) in the image. However, the main challenge here is that the Betti number is a discrete value, which can not be differentiable for backpropagating through the network. Therefore, we need an alternative way to design a differentiable loss function which is able to preserve the topology information. Instead of computing directly the number of cavities, we detect the topological errors which cause the leakage on the boundary prediction. And then we propose a novel topological loss which targets to fix topological errors of the model. The overall framework is illustrated in Fig. 1. In practice, in this article, we use a 3D U-Net (11) as a network to train our topological loss because its architecture efficiently adapts for a 3D segmentation. Moreover, using 3D U-Net is also adequate for us to demonstrate that the CavityLoss outperforms state-of-the-art methods. Note that, theoretically our loss can be used with any other networks able to segment 3D images.

2.2 Detection of the topological errors

In this section, we propose our method to detect the topological errors. The accuracy of the detection of the topological errors is essential for the training process to improve the performance of the segmentation, especially for the 3D bioimage structure. We use 26-connectivity (shortly C_{26}) to represent the connectivity between neighboring voxels in 3D image. Each voxel has 26 voxels surrounding it. So that, 26-connected voxels are neighbors to every voxel that touches one of their faces, edges, or corners. The identification of the topological errors is depicted in Fig. 2. Note that, for better visualization, we present the process on the 2D slice of the volume, but the whole algorithm works on the 3D volume. Our method starts with thresholding the prediction image u at 0.5. We call u_B the thresholded image. Different from other state-of-the-art methods (19; 20) which localize the topological errors on the continuous likelihood image, our method computes the leakages on the binary image. The main reason for that is to remove noise from the likelihood image. Then, the task now is to train the network so that the binary image u_B can be topology equivalent to the ground truth gt.

We also dilate the ground truth gt to take into account the tolerance of minor displacements in boundary location. Indeed, minor differences in the boundary have little effect on the shape of the regions. The dilated ground truth $\hat{gt}$ is then multiplied with the thresholded binary u_B to get u_M to remove the false positives in the prediction. It is because we only focus on the leakages in the boundary region. We then subtract the ground truth gt to prediction u_M to get the different zone u_{Dif} where the two maps gt and u_M disagree. We only choose the parts of the u_{Dif} zone that are positive. Note that we should not compute the loss function directly on the voxel-level disagreement because this naive metric is over-sensitive to the boundary locations. However, we also realize that the

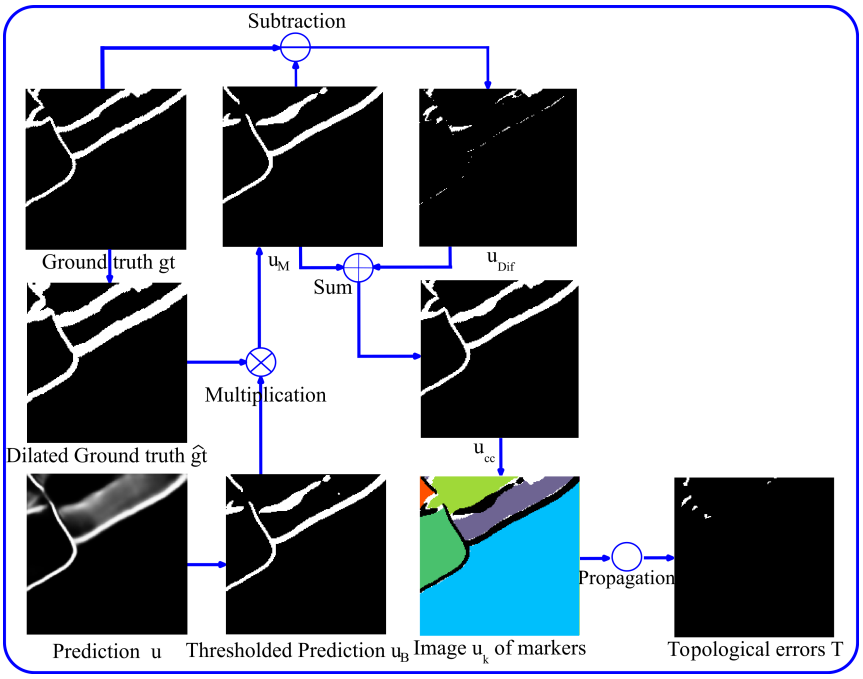


Figure 2: Illustration of the CavityLoss $\mathcal{L}_{CL}$ computation. $\mathcal{L}_{CL}$ is first implemented on the binary boundary prediction, then is multiplied with the ground truth to remove the false positives. The topological errors are detected in the difference zone (between the ground truth and prediction) by using the distance transform watershed algorithm (15).

topological errors have to belong to the different zone u_{Dif} and can not be elsewhere. By detecting these topological errors, we are able to preserve the topology of the structure.

To correctly extract the topological errors, we apply a segmentation method based on the watershed algorithm (15). However, to implement the watershed algorithm, we need to initiate the set of markers. Instead of finding local maxima of the distance map from the prediction image, we simply compute the set of connected components S from u_{cc}, the sum of the map u_M and the difference zone u_{Dif} to get u_k. The labels S in u_k also corresponds to different basin regions of the watershed algorithm (15). We use this label image as a initiated set of markers for the watershed algorithm. We then apply the watershed algorithm but the propagation is restricted only to the white areas of u_k (which correspond to the area of u_{Dif}) and we exclude the black areas as the errors can only occur in u_{Dif}. The detected watershed lines (zones) are the topological error. We can then deduce T, the image of the topological errors. In the image, we can see that all the white area of u_{Dif} are linked with more than one label in u_k.

To efficiently implement the distance transform watershed process (15), we use Dijkstra's algorithm to spread the distance values from the markers to other voxels in the image (voxels in the difference zone u_{Dif}). We use C_{26} to define the voxel connectivity in the 3D image. We first initialize the distance of the set of markers to zero. We also use a priority queue to store distance values of the voxels in the propagation process. After each iteration, we update the priority queue based on their distance costs, and update the labels of the voxels. Different label waves meet at the boundary of the regions. The process stops when every voxel in the image has been visited. The propagation is also fast as it is performed in the restriction u_{Dif}. Finally, we are able to extract the leakages on the boundary.

2.3 Topological loss

Let's call T the mask of the topological errors in the prediction image. The topological loss function $\mathcal{L}_{CL}$ of the network is defined as follows:

$$\mathcal{L}_{CL}(u, gt) = \mathcal{L}_{BCE}(u \odot T, gt \odot T) \tag{1}$$

where $\odot$ is the element-wise product and $\mathcal{L}_{BCE}$ is the binary cross entropy loss function.

The final loss of our method, $\mathcal{L}_{Total}$, is given by:

$$\mathcal{L}_{Total}(u, gt) = \mathcal{L}_{voxel}(u, gt) + \alpha \times \mathcal{L}_{CL}(u, gt), \tag{2}$$

where

$$\mathcal{L}_{voxel} = \mathcal{L}_{BCE}(u, gt) + \mathcal{L}_{Dice}(u, gt), \tag{3}$$

$\mathcal{L}_{voxel}$ is the voxel-wise loss function used to regulate the global information of the segmentation, and $\mathcal{L}_{CL}$ penalizes the topological errors on the prediction image and encourages the network to recover the broken connection. The hyper-parameter $\alpha \geq 0$ is used to balance the two losses. Note that the mask T is spare and not fixed. Indeed, the likelihood map is updated after every iteration leading to the re-computation of the set of topological errors T. Then, the new $\mathcal{L}_{CL}$ is computed. The training process continues until convergence.

3 Experiments

In this section, we test the efficiency of our proposed method and carry out an investigation on the impact of the hyper-parameter α on the performance of our method.

3.1 Datasets

The experiments are conducted on both fixed and live publicly available microscope datasets. These two datasets along with the ground truth were provided using different types of microscopes and differ in image quality.

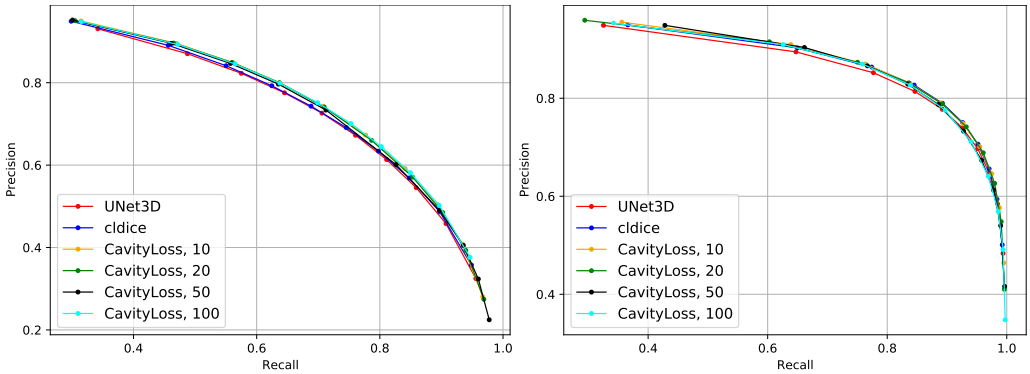
- **Arabidopsis thaliana ovules** (5) is a fixed dataset that contains 31 volumetric stacks with a voxel size of 0.075 x 0.075 x 0.235 μm to capture all the developmental stages by using confocal laser scanning microscopy. The dataset is split into 22 training stacks, 2 validating stacks, and a test set of 7 stacks.
- The second dataset contains 3 time-lapse videos of the development of **Arabidopsis thaliana lateral root primordia (LRP)** (6) using a light sheet fluorescence microscope (LSFM) with a voxel size of 0.1625 x 0.1625 x 0.250 μm. The sizes of the movies are respectively 2048 x 1050 x 486, 1940 x 1396 x 403, and 2048 x 1195 x 566 voxels. For this dataset, 28 stacks are used in total, where 22 stacks are used for training, 2 for validating, and 4 for testing.

To improve the performance of image segmentation in the training process, we also performed an on-the-fly data augmentation following the implementation in (11), including rotation, scaling, gray value augmentation, and a smooth dense deformation field on both data and ground truth labels.

3.2 Evaluation Metrics

Boundary detection and segmentation metrics are used to evaluate the performance of our proposed method. The evaluation metrics are presented below:

- **IoU** evaluates the overlap between the segmentation boundary u and the ground truth gt: IoU $= |u \cap gt|/|u \cup gt|$.
- **Precision and recall** describe respectively the purity and completeness of the positive predictions relative to the ground truth: Precision $=$ TP$/$(TP $+$ FP) and Recall $=$ TP$/$(TP $+$ FN), where TP, FP, FN are respectively the true positive, false positive, and false negative.
- **F-score** is the harmonic mean of **precision** and **recall**: $F = 2 \times P \times R/(P + R)$.
- **Adapted Rand error (ARand)** is the fraction of image voxel pairs on which the two segmentations disagree: ARand $= 1 - \frac{\sum_{ij} p_{ij}^2}{(\sum_k s_k^2 + \sum_k t_k^2)}$, where p_{ij} is the joint probability distribution that a randomly chosen voxel belong to region i in the segmentation image and region j in the ground truth, $\sum_{ij} p_{ij} = 1$, $s_i = \sum_j p_{ij}$, and $t_j = \sum_i p_{ij}$.



(a) LRP (b) Ovule

Figure 3: Precision-recall curves for different CNN variants on the ovule (a) and lateral root primordia (LRP) (b) datasets.

Greater IoU, precision, recall, and F-score scores indicate better performance. However thinner or thicker detection does not match entirely the ground truth leading to a penalty in IoU, precision, recall, and F-score even if the topological structure of the segmentation is preserved. Moreover, a missing voxel which has the same penalty can even change the topology of structures. That is the reason why we use the ARand score to capture the respect of cavity segmentation. A smaller ARand corresponds to a method with better segmentation results. Even if we want to get good results according to IoU, precision, recall, and F-score, our goal is to improve ARand score. Such orders are accompanied by an upward ($\uparrow$) or downward arrow ($\downarrow$) in upcoming quantitative result tables.

3.3 Training Details

We pre-train the network using the linear combination of BCE and Dice and the learning rate is originally set to 2.10^{-4}. Then we set the learning rate to 2.10^{-5} to fine-tune our CavityLoss function. The α value is chosen empirically depending on the dataset. As such, we experimented with different values to choose the best α for the network. Except in Sec. 3.4 the prediction is thresholded with a threshold set to 0.5. The proposed method was implemented using the Pytorch framework. All the experiments are performed on an Nvidia Quadro RTX 8000 GPU with 48GB of VRAM.

3.4 Boundary evaluation

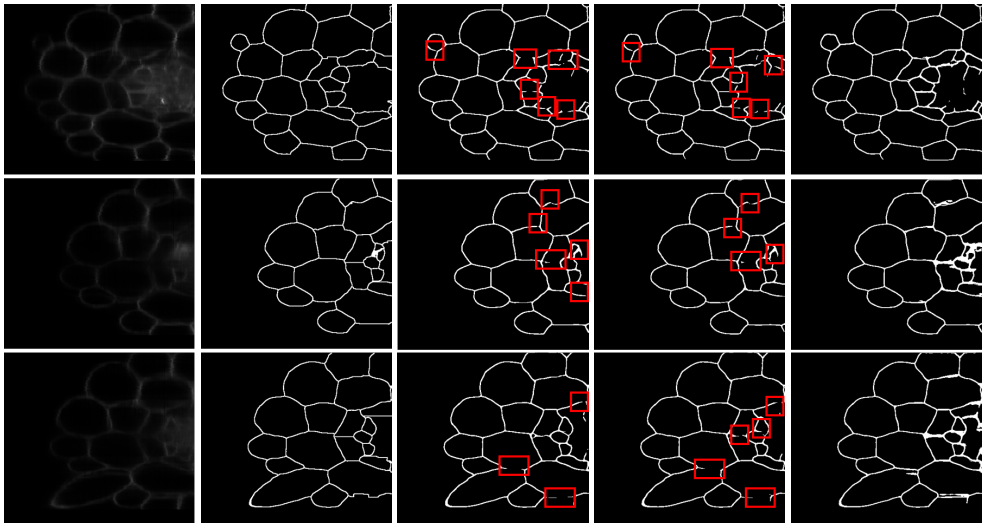
We conducted a comparison of the boundary detection between our proposed method with different values of the hyper-parameter α and state-of-art methods including 3D U-Net (11) and clDice (21) from the voxel-wise point of view. To do that, we evaluate the precision/recall curves at different threshold levels of the predicted boundary probability maps. The best performances on the LRP and Ovule datasets are illustrated in Fig. 3. At a glance, our proposed method outperforms the 3D U-Net on both datasets across different values of α. Compared to the *clDice* method, our method achieves better performance on the LRP dataset and comparable results on the Ovule dataset. In conclusion, our proposed method, which is intently designed to improve the topology for image segmentation, also ameliorates the performance of boundary detection. It is worth noting that when α increases, the boundary detection performance slightly decreases. It is because the CavityLoss focuses on the topology which can generate some noises on the segmentation image.

3.5 Segmentation evaluation

This section focuses only on the evaluation of the segmentation of tissues into cells between our method and state-of-the-art methods. To evaluate fairly when analyzing the strength of every method, we do not use any post-processing procedure (except thresholding the prediction at 0.5). We then use the Adapted Rand error (ARand) for the overall segmentation quality between different segmentation methods. The quantitative results are summarized in Tab. 1. Our first observation is that our proposed method achieves extremely high results for maintaining the topology of the image segmentation

Table 1: Quantitative results for different methods on Ovule and lateral root primordia (LRP) datasets. The best score in each metric is in bold.

Datasets	Methods	IoU↑	Precision↑	Recall↑	Fscore↑	ARand↓
LRP	UNet3D (11)	0.5611	0.7000	0.7338	0.7155	0.5078
	clDice (21)	0.5626	**0.7175**	0.7175	0.7164	0.5118
	CavityLoss, α = 10	**0.5724**	0.7097	0.7427	**0.7250**	0.4347
	CavityLoss, α = 20	0.5713	0.7035	0.7476	0.7241	0.3268
	CavityLoss, α = 50	0.5656	0.6827	0.7623	0.7195	0.1879
	CavityLoss, α = 100	0.5565	0.6550	**0.7685**	0.7160	**0.1224**
Ovule	UNet3D (11)	0.7039	0.7587	0.9102	0.8235	0.3928
	clDice (21)	0.7152	0.7704	0.9106	0.8308	0.4191
	CavityLoss, α = 10	**0.7178**	**0.7763**	0.9068	**0.8327**	0.3340
	CavityLoss, α = 20	0.7127	0.7656	0.9139	0.8293	0.3154
	CavityLoss, α = 50	0.7057	0.7618	0.9078	0.8244	**0.2998**
	CavityLoss, α = 100	0.6985	0.7445	**0.9204**	0.8192	0.3152

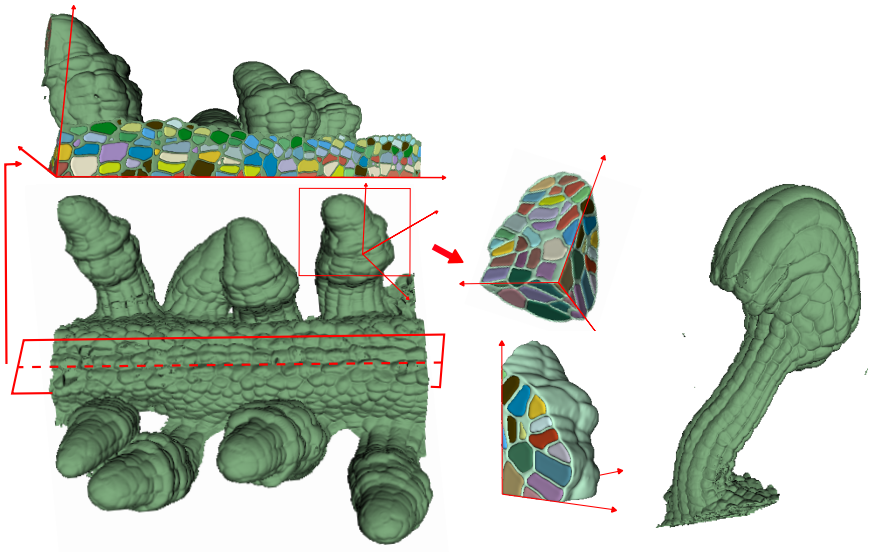


(a) Raw (b) GT (c) U-Net (d) clDice (e) Ours

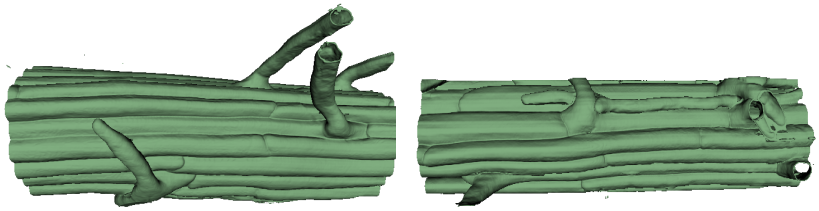
Figure 4: Qualitative results between state-of-the-art and our proposed method. (a,b) The original images and their ground truths, (c-e) segmentation results of U-Net, clDice, and CavityLoss with α equals to 50. The topological errors are highlighted in the red rectangles.

compared to 3D U-Net and clDice. For example, with the value of α equal to 50, our method improves **-0.3199** in ARand (in LRP dataset) and **-0.093** in ARand (in Ovule dataset) compared to 3D U-Net method. The second observation is that the clDice method improves the boundary detection compared to 3D U-Net for voxel-wise evaluations such as IoU, Precision, and F-score , but can not fix the cavity problem on the 3D segmentation. Lastly, we observe the impact of the hyper-parameter α on the segmentation results. The segmentation scores ARand decrease (which is better) when we increase the value of α, but it also decreases slightly the voxel-wise evaluation. This can be explained by the fact that our method focuses only on the topology of the image, in particular, the cavities in the predicted segmentation. When the value of α is too big, our CavityLoss dominates the BCE-Dice Loss, thereby generating some false positive noises on the segmentation results.

Fig. 4 illustrates the qualitative results of our proposed method with α equal to 50 compared to other state-of-the-art methods. Our method which combines the voxel-wise BCE-Dice and the CavityLoss is able to fix the broken connection problem highlighted by the red rectangles. These broken connections are because of the quality of the image (blur and noises) and also the fine-scale structure and thin connections. Our method leverages the efficiency of mathematical morphology to provide topological information, thereby accurately detecting the leakages on the boundary, thus



(a) 3D reconstruction of the ovule primordia. The middle images show the front and inside views of the individual label of the ovule primordia.



(b) 3D reconstruction of the lateral root

Figure 5: 3D reconstruction images of the cell of plant tissues using our proposed method on ovule development and lateral root datasets. An example of 3D individual label of the ovule primordia.

regulating the network behavior. As a result, our method produces more accurate image segmentation and maintains the consistency of the boundary of the cell on both fixed and live datasets. Fig. 5 shows the ovule development using confocal laser scanning microscopy and the lateral root in Arabidopsis thaliana. The 3D images are reconstructed from the segmentation of our method. Our method correctly segments the plant tissues into cells, which is important for the downstream analysis.

4 Conclusion

In this paper, we propose a simple-yet-effective topology-aware method for the correct segmentation of 3D plant tissues into cells. Our method relies on networks and especially we propose a way to train the network to detect topological issues during the learning phase, in order to focus on these errors and help the network to fix them. To do so, we introduce the CavityLoss which leverages the efficiency of the mathematical morphology to extract the topological information for regulating the training process of the network. By using the well-known watershed algorithm, we are able to detect the broken connections in the boundary prediction and recover the closeness image segmentation. For the experiments, we used the 3D U-Net but our CavityLoss can be used with any other network able to segment 3D images.

Experimenting results on the 3D ovule and lateral root datasets show that our method outperforms the state-of-the-art methods for preserving the topology of the image and also maintaining the voxel connectivities on the boundary of cells. In the future, we will extend our method to different plant and animal organs. Furthermore, we would design a Topology-Layer in the network to directly learn the structural representation.

References

[1] Romain Fernandez, Pradeep Das, Vincent Mirabet, Eric Moscardi, Jan Traas, Jean-Luc Verdeil, Grégoire Malandain, and Christophe Godin, "Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution," *Nature methods*, vol. 7, no. 7, pp. 547–553, 2010.

[2] Tamily A Weissman and Y Albert Pan, "Brainbow: new resources and emerging biological applications for multicolor genetic labeling and analysis," *Genetics*, vol. 199, no. 2, pp. 293–306, 2015.

[3] David A Van Valen, Takamasa Kudo, Keara M Lane, Derek N Macklin, Nicolas T Quach, Mialy M DeFelice, Inbal Maayan, Yu Tanouchi, Euan A Ashley, and Markus W Covert, "Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments," *PLoS computational biology*, vol. 12, no. 11, pp. e1005177, 2016.

[4] Adrian Wolny, Lorenzo Cerrone, Athul Vijayan, Rachele Tofanelli, Amaya Vilches Barro, Marion Louveaux, Christian Wenzl, Sören Strauss, David Wilson-Sánchez, Rena Lymbouridou, et al., "Accurate and versatile 3D segmentation of plant tissues at cellular resolution," *Elife*, vol. 9, pp. e57613, 2020.

[5] Rachele Tofanelli, Athul Vijayan, Sebastian Scholz, and Kay Schneitz, "Protocol for rapid clearing and staining of fixed arabidopsis ovules for improved imaging by confocal laser scanning microscopy," *Plant Methods*, vol. 15, no. 1, pp. 1–13, 2019.

[6] Amaya Vilches Barro, Dorothee Stöckle, Martha Thellmann, Paola Ruiz-Duarte, Lotte Bald, Marion Louveaux, Patrick von Born, Philipp Denninger, Tatsuaki Goh, Hidehiro Fukaki, et al., "Cytoskeleton dynamics are necessary for early events of lateral root initiation in arabidopsis," *Current Biology*, vol. 29, no. 15, pp. 2443–2454, 2019.

[7] Saining Xie and Zhuowen Tu, "Holistically-nested edge detection," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1395–1403.

[8] Iasonas Kokkinos, "Pushing the boundaries of boundary detection using deep learning," *arXiv preprint arXiv:1511.07386*, 2015.

[9] Jianzhong He, Shiliang Zhang, Ming Yang, Yanhu Shan, and Tiejun Huang, "Bi-directional cascade network for perceptual edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.

[10] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[11] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2016, pp. 424–432.

[12] Matthew D Zeiler and Rob Fergus, "Visualizing and understanding convolutional networks," in *European conference on computer vision*. Springer, 2014, pp. 818–833.

[13] Serge Beucher and Fernand Meyer, "The morphological approach to segmentation: the watershed transformation," in *Mathematical morphology in image processing*, pp. 433–481. CRC Press, 2018.

[14] Laurent Najman and Michel Schmitt, "Geodesic saliency of watershed contours and hierarchical segmentation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 18, no. 12, pp. 1163–1173, 1996.

[15] Jos BTM Roerdink and Arnold Meijster, "The watershed transform: Definitions, algorithms and parallelization strategies," *Fundamenta informaticae*, vol. 41, no. 1-2, pp. 187–228, 2000.

[16] Steffen Wolf, Lukas Schott, Ullrich Kothe, and Fred Hamprecht, "Learned watershed: End-to-end learning of seeded segmentation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Oct 2017, pp. 2011–2019.

[17] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore, "Reinforcement learning: A survey," *Journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996.

[18] Agata Mosinska, Pablo Marquez-Neila, Mateusz Koziński, and Pascal Fua, "Beyond the pixel-wise loss for topology-aware delineation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3136–3145.

[19] Xiaoling Hu, Fuxin Li, Dimitris Samaras, and Chao Chen, "Topology-preserving deep image segmentation," in *Proc. NIPS*, 2019, pp. 5658–5669.

[20] Minh Ôn Vû Ngoc, Yizi Chen, Nicolas Boutry, Joseph Chazalon, Edwin Carlinet, Jonathan Fabrizio, Clément Mallet, and Thierry Géraud, "Introducing the boundary-aware loss for deep image segmentation," in *British Machine Vision Conference (BMVC) 2021*, 2021.

[21] Suprosanna Shit, Johannes C Paetzold, Anjany Sekuboyina, Ivan Ezhov, Alexander Unger, Andrey Zhylka, Josien PW Pluim, Ulrich Bauer, and Bjoern H Menze, "clDice - a novel topology-preserving loss function for tubular structure segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 16560–16569.