RESEARCH ARTICLE
Refined tooth and pulp segmentation using U-Net in CBCT image

[1]Wei Duan, [1]Yufei Chen, [2]Qi Zhang, [2]Xiang Lin and [1]Xiaoyu Yang

[1]College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China; [2]Department of Endodontics, School and Hospital of Stomatology, Tongji University, Shanghai Engineering Research Center of Tooth Restoration and Regeneration, Shanghai 200072, China

Objectives: The aim of this study was extracting any single tooth from a CBCT scan and performing tooth and pulp cavity segmentation to visualize and to have knowledge of internal anatomy relationships before undertaking endodontic therapy.

Methods: We propose a two-phase deep learning solution for accurate tooth and pulp cavity segmentation. First, the single tooth bounding box is extracted automatically for both single-rooted tooth (ST) and multirooted tooth (MT). It is achieved by using the Region Proposal Network (RPN) with Feature Pyramid Network (FPN) method from the perspective of panorama. Second, U-Net model is iteratively performed for refined tooth and pulp segmentation against two types of tooth ST and MT, respectively. In light of rough data and annotation problems for dental pulp, we design a loss function with a smoothness penalty in the network. Furthermore, the multi-view data enhancement is proposed to solve the small data challenge and morphology structural problems.

Results: The experimental results show that the proposed method can obtain an average dice 95.7% for ST, 96.2% for MT and 88.6% for pulp of ST, 87.6% for pulp of MT.

Conclusions: This study proposed a two-phase deep learning solution for fast and accurately extracting any single tooth from a CBCT scan and performing accurate tooth and pulp cavity segmentation. The 3D reconstruction results can completely show the morphology of teeth and pulps, it also provides valuable data for further research and clinical practice.

Dentomaxillofacial Radiology (2021) **50**, 20200251. doi: 10.1259/dmfr.20200251

Introduction

Dental pulp diseases and periapical disease are common and frequently occurring in the oral cavity. Root-canal therapy is the most commonly used and effective treatment. It is important to visualize and to gain knowledge of internal anatomy relationships before conducting endodontic therapy, which is the prerequisite for accurate diagnosis and preoperative assessment.[1]

Compared with the two-dimensional images generated by conventional radiography, the three-dimensional characteristic of CBCT provides more information about the teeth and their surrounding structures, which makes the application of it in endodontics rapidly growing worldwide.[2] However, the image presented to the dentists is blurred and two-dimensional, imagining the 3D structure of the teeth highly depends on the clinical experience, especially the dental pulp reconstruction is one of the most challenging problems for all dentists.

Various traditional methods for tooth CBCT image segmentation have been proposed. These methods can be divided into two categories: threshold based approaches[3–5] and active contour model related methods.[6–8] The former methods lack of robustness because a single threshold is hard to distinguish the alveolar bone and the root region. Besides, the threshold value is difficult to adapt to different applications. The latter methods are sensitive to initialization that

requires tedious user interactions. Few works studied on the topic of tooth pulp cavity segmentation in CBCT. Wang et al. in 2019 proposed a simple threshold-based approach for pulp cavity segmentation.[9] The method cannot guarantee an accurate and robust result. It requires manual selection of a single tooth region for subsequent processing. And it becomes worse if the intensity of the pulp cavity is significantly different from the background or the pulp cavity region is too small. Therefore, this method is not suitable for complex and variable clinical applications.

Recently, many deep learning methods for tooth image analysis have demonstrated promising performance over traditional methods in various tasks, such as classification of teeth in CBCT or in panoramic radiographs,[10,11] detection and diagnosis of dental disease in radiographic images,[12–14] caries detection with near-infrared transillumination.[15,16] As for images segmentation, based on U-Net[17] which have achieved good performance on variety biomedical segmentation applications, Tian et al used Hierarchical Deep Learning Networks, to achieve automatic classification and segmentation of teeth on 3D dental model.[18] Gou M et al integrated U-Net with level set model for tooth CT image segmentation.[19] Besides U-Net, Cui et al proposed Tooth-Net to achieve automatic and accurate tooth instance segmentation and identification from CBCT images for digital dentistry.[20] However, most existing methods target on orthodontic diagnosis rather than endodontic therapy, so they aim to segment whole teeth instead of a single tooth. None of the above work has carried out related research on pulp cavity.

Segmentation of tooth and its pulp cavity from CBCT by using deep supervised neural network is still a challenging task for following reasons:

(1) Extracting any single tooth from a CBCT scan is necessary but very difficult, since the varied shape and complex distribution of teeth no matter in axial or sagittal views, as shown in Figure 1(a). (2) Manually labeling of the tooth and pulp slice-by-slice is time-consuming and labor-intensive, which limits the number of labeled data so that makes the training of network inefficiency. Moreover, as shown in Figure 1(b),

the crown-to-root topology in traditional axial-view is constantly changing between slices, causing the axial-view cannot exhibit a complete morphological structure of the tooth as in sagittal view. Thus, it is indispensable to utilize the image features in different views to improve the training ability of the model on limit labeled samples. (3) As the Figure 1(c) shows, due to the low contract and intensity inhomogeneity in CBCT, the boundaries of the region-of-interest (ROI) are extremely ambiguous especially at the pulp cavity. This causes the doctors to be unable to exactly defined boundaries of labels.

In this work, we propose a refined tooth and pulp segmentation using U-Net in CBCT images. The main contributions of the work are as follows: (1) From the perspective of panorama, each tooth area can be automatically pinpointed as bounding boxes from CBCT by using deep learning method. (2) According to the tooth type, multi-view data are generated to improve the training ability of the model on limit labeled samples. (3) For refined pulp cavity segmentation, the U-Net model with smoothness penalty loss function are executed iteratively to remove the interference of the irrelevant tissue around the teeth, which improves the accuracy of the segmentation results with less mottled and zigzag effects on the boundaries.

Methods and materials

In this study, 20 sets of CBCT image data were collected for the segmentation experiment. All CBCT data were scanned by an ORTHOPHOS SL 3D CBCT machine (Dentsply Sirona, Germany). The field of view (FOV) size is 8 cm × 8 cm × 8 cm, and the voxel size is 0.16 × 0.16 × 0.16 mm. According to the equipment manufacturer's instructions, the exposure parameters were set as follows: voltage 85kVp, current 10mA, time 5000 ms. The original image matrix of each slice is 501 × 501, the number of each volume is 501, and the file format is DICOM.
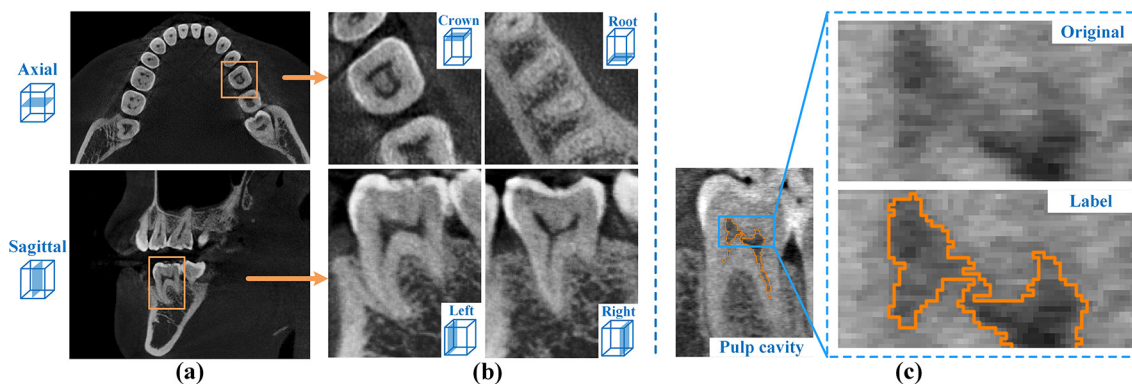
Figure 1 The segmentation challenges of tooth and its pulp cavity from CBCT.
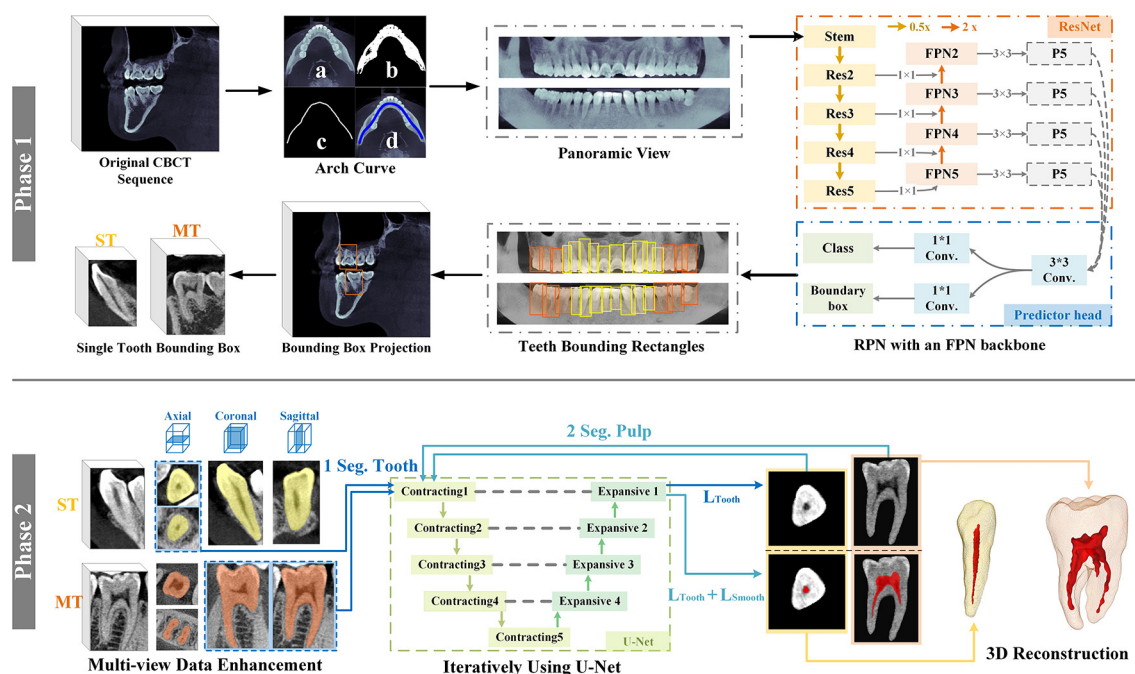
Figure 2 A two-phase framework for tooth and pulp cavity segmentation. **Phase1**: Single Tooth Bounding Box Extraction. **Phase2:** Refined Tooth and Pulp Segmentation.

Workflow

Our goal is to visualize and study the internal anatomy of the dental pulp to serve endodontic therapy instead of orthodontic diagnosis. Therefore, the main emphasis becomes the morphology of each individual tooth and its pulp cavity. As illustrated in Figure 2, in this paper, we propose a two-phase framework for tooth and pulp cavity segmentation. In Phase 1, using original CBCT sequence, the panoramic view is generated and put into the RPN with FPN to pinpoint each tooth, which helps us locate each single tooth and obtain the bounding box of it. In Phase 2, data in multiple views are generated for training, and then using U-Net iteratively with different loss functions for segmentation. The details of each phase will be introduced below.

Single tooth bounding box extraction

Extracting any single tooth from an original CBCT scan is very difficult, since the varied shape and complex distribution of teeth. Our proposed method is shown in Figure 2(**Phase1**). Using original CBCT sequence, we first get the dental arch curve by several steps: (a) calculate the maximum intensity projection (MIP), (b) binarize, (c) skeletonize, (d) get result. Then, we project along the tangent of each point on the dental arch curve to generate a panoramic view of the teeth sequence.

The panoramic image will be put into the pre-trained RPN with FPN[21] to identify the type of each tooth and to get the bounding rectangle of it. According to the

number of tooth roots[22] and considering the feasibility of subsequent work, the teeth are separated into two types: single-rooted tooth (ST) and multirooted tooth (MT), marked as yellow and orange, respectively.

Finally, the rectangles are projected into original CBCT to obtain the bounding box for each single tooth. For single-rooted tooth, the boxes size are $64 \times 94 \times l$, where l is the length of the bounding rectangle. For multirooted tooth, the boxes size are set as $128 \times 128 \times 176$.

Multi-view data enhancement

The doctors often label the tooth and pulp in single axial view. Manually labeling is a time-consuming and labor-intensive process, thus only limited labeled data is available for training. As illustrated in Figure 2 (**Phase2**), the single-view of a 2D slice sequence cannot reflect the actual morphological structure of the tooth. Particularly, for MT, the crown-to-root topology of traditional axial-view of 2D slices is constantly changing. Thus, it is indispensable to utilize the image features of multiple views to improve the training ability of the model on small samples. According to the morphological characteristics of two types of teeth, we choose different views to generate more batches.

For simple single-rooted teeth, the axial view is chosen. Compared with the other two perspectives, the axial view of ROI is approximating a circle which has more uniform features whether the crown or the root.

For complex multirooted teeth, the coronal and sagittal views are chosen. Due to the limited number of existing annotations, the traditional axial view cannot reflect the variable target structures from the crown to the root. While the other views can complement more spatial information and better reflect the structure of the teeth. In this way, one multirooted teeth bounding box can generate two sets of batches to improve learning ability of the model on limited labeled samples.

Refined tooth and pulp segmentation

The network architecture for tooth and pulp segmentation is shown in Figure 2(**Phase2**). The core part adopts the U-Net,[17] which consists of five contracting blocks in encoding path, four expanding blocks in decoding path and four jump connections with the correspondingly cropped feature map between convolutional and up-convolutional blocks. Through four expanding blocks and four jump connections, U-Net is thought to combine low-resolution deep information for object recognition and high-resolution shallow information for boundary segmentation, enabling it to achieve significant performance in medical image segmentation.

Through using U-Net iteratively, the network first implements tooth segmentation. Because of the intensity of teeth is close to the surrounding tissues in CBCT, the tooth boundary is too ambiguous to be exactly defined especially at the root. We use a combination of binary cross-entropy and dice coefficient as the loss function,[23] denoted as L_{Tooth}, to make the results as accurate as possible.

$$L_{Tooth} = -\frac{1}{N} \sum_{b=1}^{N} \left(\frac{1}{2} \cdot Y_b \cdot log\hat{Y}_b + \frac{2 \cdot Y_b \cdot \hat{Y}_b}{Y_b + \hat{Y}_b} \right) \quad (1)$$

where $\hat{Y}_b$ and Y_b denote the predicted result and the ground truth of bth image, respectively, and N indicates the batch size.

After extracting the tooth from CBCT images, the results feed into the network again for pulp cavity segmentation. The slender and faint pulp cavity leads to rough edges in original images, even the doctors are not able to exactly define its boundaries. General method will have zigzag effects on the boundaries, tend to be mottled and have plenty of mis-segmentation points. To tackle this problem, a smoothness penalty term is added into the loss function, denoted as follows.

$$L_{pulp\ cavity} = L_{Tooth} + L_{smooth}$$

$$= L_{Tooth} + \frac{\lambda}{N} \sum_{b=1}^{N} \sum_{i,j} \sqrt{\left(\hat{Y}_{(i,j+1,b)} - \hat{Y}_{(i,j,b)} \right)^2 + \left(\hat{Y}_{(i+1,j,b)} - \hat{Y}_{(i,j,b)} \right)^2}$$

$$(2)$$

where (i, j, b) is the coordinate of the bth image pixel. λ is the smoothness penalty weight. Specifically, given

a coordinate (i, j, b), $\hat{Y}_{(i,j\ b)}$ equals to the value of a pixel in the bth predicted image. $\left(\hat{Y}_{(i,j+1,b)} - \hat{Y}_{(i,j,b)} \right)^2$ and $\left(\hat{Y}_{(i+1,j,b)} - \hat{Y}_{(i,j,b)} \right)^2$ calculate the variance between the pixel with its two neighbors (right and down). In this way, the degree of difference between each pixel and its surrounding area is integrated into the loss function. If the pixel is isolated whose intensity is varying from the surroundings, the variance value is large, otherwise is small. L_{smooth} calculates the sum of the above differences in the whole image. When L_{smooth} is reduced, it will make the pixels of predicted boundary more unified so that reducing the number of independent pixels and decreasing the zigzag effects.

After the tooth and its pulp cavity segmentation is finished, 3D reconstruction of the results is performed to show the internal and external structure of the teeth more intuitively.

Performance evaluation

The network is implemented in PyTorch and trained on the server with NVIDIA Tesla P4. All training processes separate dataset randomly into three batches for each epoch, and 10 epochs were run at a learning rate of 0.0001 with the Adam Optimization. The ground truth was labeled by two doctors from Department of Endodontics, School and Hospital of Stomatology, Tongji University after extracted single tooth bounding boxes from original CBCT. We select 30 sets of single-rooted teeth and 30 sets of multirooted teeth separately for the next segmentation.

For each type of tooth, ten-fold cross-validation was used to evaluate the model performance from the following criteria: dice similarity coefficient (DSC), average symmetric surface distance (ASD) metric, and relative volume difference (RVD).[24] For pulp cavity, we add another criterion: surface area (SA). They can be defined as follows:

$$DSC = \frac{2|MS \cap GT|}{|MS| + |GT|}$$

$$\begin{aligned} ASD = 0.16 \times &\frac{1}{|B_{MS}| + |B_{GT}|} \\ \times &\left(\sum_{x \in B_{MS}} d\left(x, B_{GT}\right) + \sum_{y \in B_{GT}} d\left(y, B_{MS}\right) \right) \end{aligned} \quad (3)$$

$$RVD = \pm \frac{|MS| - |GT|}{|GT|}$$

$$SA = 0.16 \times 0.16 \times B_{MS}$$

where MS and GT denote machine segmented and ground truth sets of voxels, respectively, B_{MS} and B_{GT} denote the boundaries of MS and GT, respectively.

Table 1   The performance of tooth segmentation

Tooth Type	Approach	View	Dice	ASD (mm)	RVD
ST	Threshold-based[4]	/	0.583±0.072	0.909±0.162	0.898±0.463
	FCN[25]	Axial	0.872±0.054	0.373±0.153	0.159±0.128
	U-Net[17]		**0.957±0.005**	**0.104±0.019**	**0.049±0.017**
MT	Threshold-based[4]	/	0.568±0.062	1.152±0.228	0.703±0.554
	FCN[25]	Axial	0.863±0.054	0.445±0.127	0.146±0.187
	U-Net[17]	Axial	0.908±0.040	0.439±0.244	0.106±0.080
		Coronal + Sagittal	**0.962±0.002**	**0.137±0.019**	**0.053±0.010**

$d\left(x, A\right) = \min_{y \in A} d\left(x, y\right)$ is a distance measure for a voxel x from a set of voxels A, where $d(x, y)$ is the Euclidean distance between voxel x and y. As mentioned before, the voxel size of CBCT image is $0.16 \times 0.16 \times 0.16$ mm, to better present segmentation results, we convert the voxel values to actual values.

Results

Table 1 summarizes the performance of tooth segmentation using different approaches for two types of teeth. Using the traditional threshold-based approach, the segmentation results are quite unsatisfactory with dice values of both types of teeth less than 59%. If using FCN approach, the results are also not as promising as U-net. For ST, using training batches generated from the axial view to train U-Net, the average dice value of segmentation result reaches 95.7%. ASD and RVD also show that the surface shape and volume of the result are close to the label.

For MT, compared with the single axial view, choosing the other two views to train U-Net increases the average dice value of the segmentation result from 90.8% to 96.2% and the average ASD decreases from 0.439 mm to 0.137 mm, the average RVD decreases from 10.6% to 5.3% and the standard deviation of all indicators has been reduced.

As shown in Figure 3(a), different views are chosen to perform segmentation on ST and MT using U-Net, the segmentation result have accurate boundaries in 2D slices and complete morphology in 3D reconstruction.

Obtaining a good tooth segmentation result is crucial for the next pulp segmentation. Therefore, the U-net was chosen for the next pulp cavity segmentation to test the effectiveness of different loss functions, the results of which are presented in Table 2.

For pulp cavity segmentation, after adding the smoothing term in loss function, the dice value of ST and MT are both increase while the RVD and SA are decrease significantly which means smoothness penalty term can improve surface smoothness of segmentation results. As Figure 3(b) shows, the 3D reconstruction of segmentation results demonstrates the internal and external structure of the teeth intuitively.

Discussion

Tooth and pulp cavity segmentation in full mouth CBCT is a challenging task due to several aspects. Firstly, extracting any single tooth from a CBCT scan is difficult for locating the position of each tooth in 3D space and defining the boundaries. To address this problem,

(a) Tooth segmentation results

(b) 3D reconstruction of final results
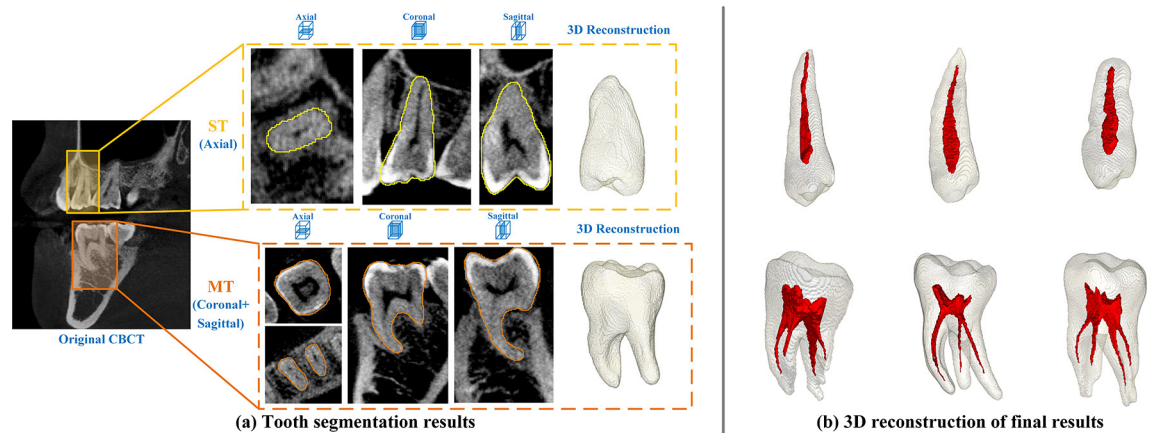
Figure 3   Segmentation results of tooth and pulp cavity using U-Net.

Refined Tooth and Pulp Segmentation using U-Net in CBCT Image
Duan *et al*
6 of 8

Table 2 The performance of pulp cavity segmentation

Tooth Type	Approach	Loss function	Dice	ASD (mm)	RVD	SA (mm²)
ST	U-Net[17]	L_{Tooth}	0.877±0.021	0.062±0.028	0.111±0.051	49.937±7.527
		$L_{Tooth} + L_{Smooth}$	**0.886±0.025**	**0.049±0.011**	**0.102±0.043**	**46.460±8.837**
MT		L_{Tooth}	0.865±0.020	**0.079±0.019**	0.143±0.086	125.180±28.696
		$L_{Tooth} + L_{Smooth}$	**0.876±0.025**	0.081±0.026	**0.080±0.033**	**102.060±24.826**

Cui et al identified each tooth in 3D image by encoding additional spatial and shape features in Tooth-Net.[20] However, identifying each tooth directly in 3D images requires complicated calculation methods and consumes more computing resources. Our proposed method uses the panoramic view to project the 3D image into 2D to reduce the spatial complexity of the image, and then use deep learning method to accurately identify each tooth and get the bounding rectangles. Compared with methods proposed by Hosntalab et al who calculated the dividing line between each tooth on panorama in 2D space,[6] our bounding rectangles contain more information about the length and width of the teeth.

Secondly, for type of MT, the single axial view of 2D slices cannot reflect the actual morphological structure of the tooth, cannot provide enough spatial information and the ROI in this view varies in shape. Choosing the suitable tooth segmentation method and using appropriate data augmentation and is essential to achieving good segmentation results. While it can be seen from the Table 1 that using U-Net for the axial view already yields better results than the threshold-based and the FCN-based approaches, the U-Net cannot learn such complex and changeable features from limited labeled samples, as Figure 4 shows, so that fragmented and unsatisfactory results are obtained inevitability. Generating batches on coronal and sagittal views simultaneously can better reflect the morphological characteristics of teeth, complements spatial information with each other, facilitates the division of tooth roots and surrounding tissues so that the boundaries of the segmentation results are accurate and clear. For ST, the structure of the tooth type is relatively simple and only the axial view is chosen for training. It can obtain satisfactory results with an average dice of 95.7%.

Thirdly, due to the low resolution and similar intensity distribution of CBCT, boundaries of the pulp cavity in the image are extremely ambiguous, as we can see in Figure 5(1)(4). It is difficult for doctors to exactly defined the boundaries when manually label the ground truth of pulp cavity. Besides, the labeling is mainly performed on traditional axial view that the result lacks of continuity and smoothness. As given in Figure 5(2)(5), the method using general loss function results in mottled and independent regions. After adding smooth term incorporating the intensity difference into the loss function, the modified model not only wipe incorrectly independent areas that shown in Figure 5(3), but also reduce zigzag effects and improve edge smoothness on pulp cavity boundaries that can be seen from Figure 5(6) and the 3D reconstruction results in Figure 5. Compared with the method proposed by Wang et al for pulp cavity segmentation,[9] our proposed method does not rely on a certain threshold calculation, it further improves the smoothness of the segmentation results to better meet clinical needs.
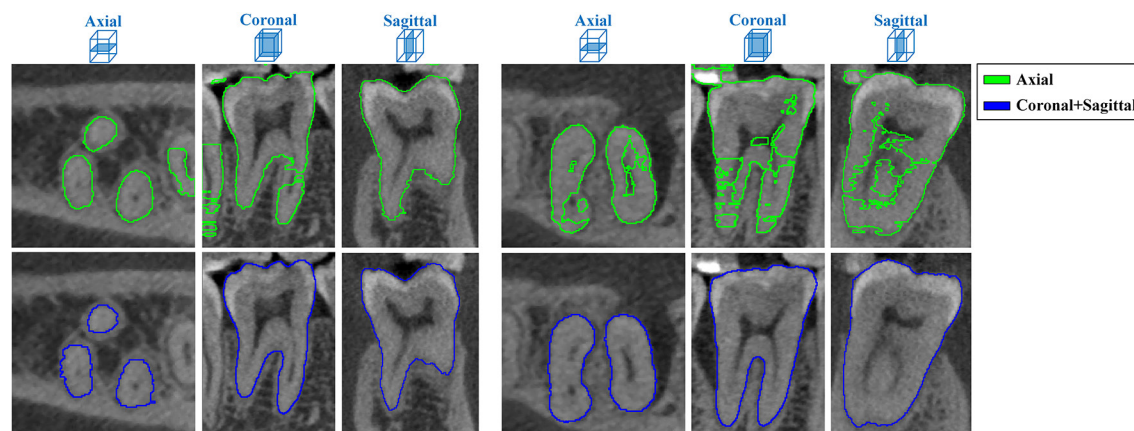
Figure 4 Comparing segmentation results of MT trained on different views.
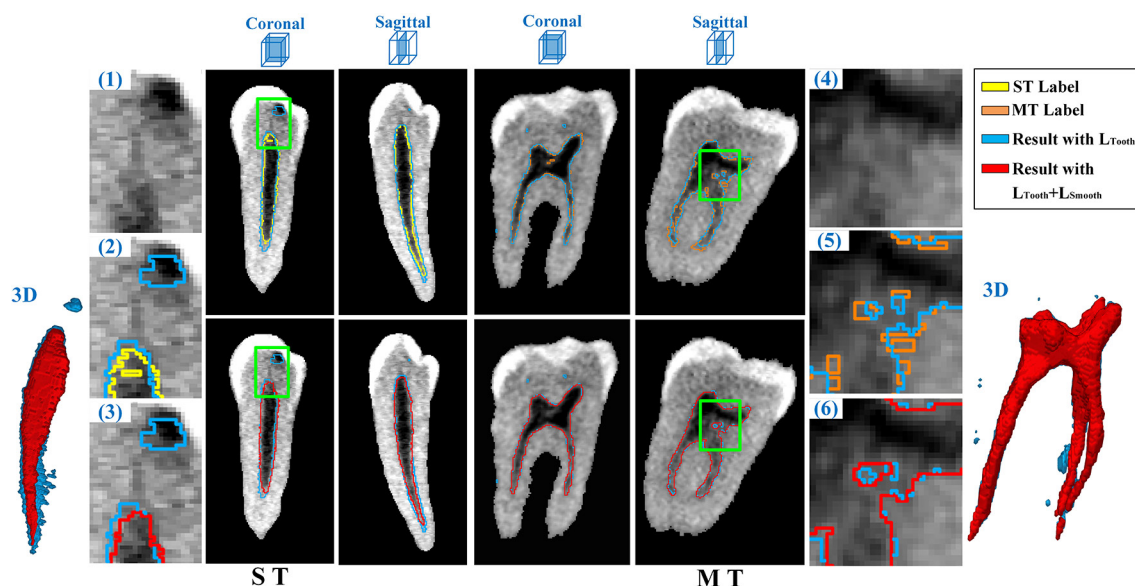
Figure 5 Segmentation and 3D reconstruction results of pulp cavity for ST and MT.

Fast and accurately achieving tooth and its pulp cavity segmentation and visualization is important in helping doctors understand the morphology and structure of dental pulp before implementing appropriate endodontic treatment in patients. Using our solution, we are able to achieve considerable accuracy and efficiency of refined tooth and pulp segmentation. In the future, we plan to test the model robustness on larger data set, perform attribute calculation on tooth and its pulp cavity, locate initial penetration point and calculate the path to access the pulp chamber and find the orifices of the root canals.

Conclusion

This paper proposed a two-phase deep learning solution for accurate tooth and pulp cavity segmentation. First, the single tooth bounding box is extracted automatically for both ST and MT. Second, U-Net model is iteratively performed for refined tooth and pulp segmentation against two types of tooth. In light of rough data and annotation problems for dental pulp, we design a loss function with smoothness penalty in the network. Furthermore, the multi-view data enhancement is proposed to solve the few shot data challenge and morphology structural problems. The experimental results show that the proposed method can obtain an average dice 95.7% for ST, 96.2% for MT, and 88.6% for pulp of ST, 87.6% for pulp of MT.

Acknowledgment

REFERENCES

1. Vertucci FJ. Root canal morphology and its relationship to endodontic procedures. *Endodontic topics* 2005; **10**: 3–29.
2. Patel S, Brown J, Pimentel T, Kelly RD, Abella F, Durack C. Cone beam computed tomography in Endodontics - a review of the literature. *Int Endod J* 2019; **52**: 1138–52. doi: https://doi.org/10.1111/iej.13115
3. Akhoondali H, Zoroofi R, Shirani G. Rapid automatic segmentation and visualization of teeth in CT-scan data. *Journal of Applied Sciences* 2009; **9**: 2031–44.
4. Naumovich SS, Naumovich SA, Goncharenko VG. Three-Dimensional reconstruction of teeth and jaws based on segmentation of CT images using watershed transformation. *Dentomax-illofac Radiol* 2015; **44**: 20140313. doi: https://doi.org/10.1259/dmfr.20140313
5. Kang HC, Choi C, Shin J, Lee J, Shin Y-G. Fast and accurate semiautomatic segmentation of individual teeth from dental CT images. *Comput Math Methods Med* 2015; **2015**: 810796. doi: https://doi.org/10.1155/2015/810796
6. Hosntalab M, Zoroofi RA, Tehrani-Fard AA, Shirani G. Segmentation of teeth in CT volumetric dataset by panoramic projection and variational level set. *International Journal of Computer Assisted Radiology and Surgery* 2008; **3**(3-4): 257–65.
7. Ji DX, Ong SH, Foong KWC. A level-set based approach for anterior teeth segmentation in cone beam computed tomography

images. *Comput Biol Med* 2014; **50**: 116–28. doi: https://doi.org/10.1016/j.compbiomed.2014.04.006

8. Wang Y, Liu S, Wang G, Liu Y. Accurate tooth segmentation with improved hybrid active contour model. *Phys Med Biol* 2018; **64**: 015012. doi: https://doi.org/10.1088/1361-6560/aaf441

9. Wang L, Li J-P, Ge Z-P, Li G, Jp L, Zp G. Cbct image based segmentation method for tooth pulp cavity region extraction. *Dentomaxillofac Radiol* 2019; **48**: 20180236. doi: https://doi.org/10.1259/dmfr.20180236

10. Miki Y, Muramatsu C, Hayashi T, Zhou X, Hara T, Katsumata A, et al. Classification of teeth in cone-beam CT using deep convolutional neural network. *Comput Biol Med* 2017; **80**: 24–9. doi: https://doi.org/10.1016/j.compbiomed.2016.11.003

11. Tuzoff DV, Tuzova LN, Bornstein MM, Krasnov AS, Kharchenko MA, Nikolenko SI, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofac Radiol* 2019; **48**: 20180051. doi: https://doi.org/10.1259/dmfr.20180051

12. Lee J-H, Kim D-H, Jeong S-N, Choi S-H. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J Dent* 2018; **77**: 106–11. doi: https://doi.org/10.1016/j.jdent.2018.07.015

13. Lee J-H, Kim D-H, Jeong S-N, Choi S-H. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J Periodontal Implant Sci* 2018; **48**: 114–23. doi: https://doi.org/10.5051/jpis.2018.48.2.114

14. Lee J-S, Adhikari S, Liu L, Jeong H-G, Kim H, Yoon S-J. Osteoporosis detection in panoramic radiographs using a deep convolutional neural network-based computer-assisted diagnosis system: a preliminary study. *Dentomaxillofac Radiol* 2019; **48**: 20170344. doi: https://doi.org/10.1259/dmfr.20170344

15. Casalegno F, Newton T, Daher R, Abdelaziz M, Lodi-Rizzini A, Schürmann F, et al. Caries detection with near-infrared transillumination using deep learning. *J Dent Res* 2019; **98**: 1227–33. doi: https://doi.org/10.1177/0022034519871884

16. Schwendicke F, Elhennawy K, Paris S, Friebertshäuser P, Krois J. Deep learning for caries lesion detection in near-infrared light transillumination images: a pilot study. *J Dent* 2020; **92**: 103260. doi: https://doi.org/10.1016/j.jdent.2019.103260

17. Ronneberger O, Fischer P. Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical image computing and computer-assisted intervention*: Springer; 2015. pp. 234–41.

18. Tian S, Dai N, Zhang B, Yuan F, Yu Q, Cheng X. Automatic classification and segmentation of teeth on 3D dental model using hierarchical deep learning networks. *IEEE Access* 2019; **7**: 84817–28.

19. Gou M, Rao Y, Zhang M, Sun J, Cheng K. Automatic Image Annotation and Deep Learning for Tooth CT Image Segmentation. In: *International Conference on Image and Graphics*: Springer; 2019. pp. 519–28.

20. Cui Z, Li C, Wang W. ToothNet: automatic tooth instance segmentation and identification from cone beam CT images. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2019. pp. 6368–77.

21. Lin TY, Dollar P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2017. pp. 2117–25.

22. Ahmed HMA, Dummer PMH. A new system for classifying tooth, root and canal anomalies. *Int Endod J* 2018; **51**: 389–404. doi: https://doi.org/10.1111/iej.12867

23. Zhou Z. *Unet++: A nested u-net architecture for medical image segmentation*. Springer, Cham: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support; 2018. pp. 3–11.

24. Yeghiazaryan V, Voiculescu I. *An overview of current evaluation methods used in medical image segmentation*. University of Oxford: Department of Computer Science; 2015.

25. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C. *Proceedings of the IEEE conference on computer vision and pattern recognition* 2015;: 3431–40.