DENTAL ANATOMY SEGMENTATION FROM CONE BEAM CT IMAGES

Minhui Tan[1,2], Yu Fang[1], Lei Ma[1], Yu Zhang[2], Zhiming Cui[1,], Dinggang Shen[1,3,4]*

[1] School of Biomedical Engineering, ShanghaiTech University, Shanghai, 201210, China
[2] School of Biomedical Engineering, Southern Medical University, Guangzhou, 510515, China
[3] Shanghai United Imaging Intelligence Co., Ltd., Shanghai, 200230, China
[4] Shanghai Clinical Research and Trial Center, Shanghai, 201210, China

ABSTRACT

Dental cone-beam computed tomography (CBCT) has been commonly used in digital dentistry, measuring complete dental anatomy information for diagnosis and treatment planning. Existing tooth segmentation from CBCT images has made adequate progress. However, there is no study proposed to make dental anatomical segmentation (i.e., enamel, pulp, and dentin), even if it is crucial in digital dentistry. Moreover, the limited CBCT resolution and large shape variance bring additional challenges. In this paper, we propose a novel learning-based method to automatically segment 3D tooth from CBCT images with structurally anatomical parts (i.e., enamel, pulp, and dentin). Furthermore, we utilize the tooth skeleton and Frangi filter to guide pulp segmentation precisely. Extensive experiments on our established dataset of 200 patients demonstrate the effectiveness and advantage of our method, which is helpful for clinical diagnosis and surgical treatment.

Index Terms— Dental cone-beam computed tomography (CBCT), Dental anatomy segmentation, Skeleton, Frangi filter

1. INTRODUCTION

Dental cone-beam computed tomography (CBCT) is a 3D volumetric data that has been widely used in digital dentistry, providing comprehensive oral information with detailed dental anatomical structures. Existing research mainly focuses on tooth segmentation from CBCT images, for downstream applications, such as path planning in orthodontic treatment. However, many important dental anatomical structures (e.g., enamel, pulp, and dentin) in CBCT images, as shown in Fig. 1, are completely ignored, even if they are of great significance in dental diagnosis (e.g., caries, and pulpitis) and treatment design (e.g., crown planning, and root canal therapy)[1]. Hence, in this paper, we aim to automatically segment 3D tooth from CBCT images with structurally anatomical parts, including enamel, pulp, and dentin.

Tooth segmentation from CBCT images, a crucial step in digital dentistry, has been a long-standing research problem.

---

* is the corresponding author
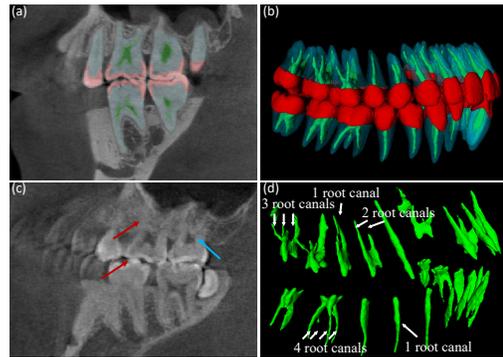
Fig. 1. Overview of dental anatomy structures and also challenges in segmentation.

Traditional methods mainly include level-set [2], but extensive manual interactions are usually involved due to lack of robustness. Recently, with the advance of deep learning, data-driven tooth segmentation frameworks have been developed rapidly with promising results, using either instance segmentation mechanism (e.g., ToothNet [3]). However, these methods only obtain individual tooth shapes, not those finer-layer anatomical structures (i.e., enamel, pulp, and dentin) embedded in CBCT images.

Automatic and accurate segmentation of dental anatomy is relatively difficult with many challenges. First, due to limitations of CBCT image resolution, the intensity values of enamel and dentin are similar to those of adjacent teeth and alveolar bone, respectively, leading to incomplete segmentation boundaries (Fig. 1(c)). Second, the pulp topologies of incisors, canines, and molars are complex and vary greatly with slender root canal(s), resulting in poor consistency of root canal on images, which in return affects the accuracy of pulp segmentation (Fig. 1(d)). Third, there are low contrast and ambiguous boundaries between pulp and dentin areas due to inhomogeneity of CBCT images (Fig. 1(c)).

To tackle the challenges mentioned above, we propose a novel framework for automatic dental anatomy segmentation from CBCT images. First, we detect and segment individual teeth from CBCT scans. Then, for each tooth, we design another network to decompose it with the enamel, pulp, and dentin structures. Specifically, we observe that the geometry

Fig. 2. Our proposed framework for dental anatomy segmentation from CBCT images. Given a CBCT image, we first detect and segment each tooth with instance segmentation. Then, we apply skeleton extraction on the tooth label and Frangi filter on the tooth image to obtain the tooth skeleton and Frangi image respectively. Eventually, we input the product of tooth image and tooth label, tooth skeleton, and Frangi image into the network to segment dental anatomy (i.e., enamel, pulp, and dentin).
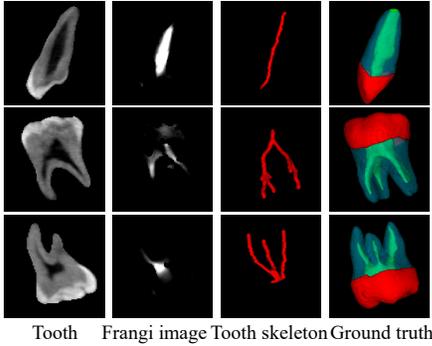
Fig. 3. Typical cases for Frangi images and tooth skeletons of incisors and molars. Notice that Frangi images highlight pulp-dentin boundaries, while all tooth skeletons share similar shapes with their corresponding pulps.

of pulp is consistent with the tooth skeleton, a coarse yet effective tooth shape representation, i.e., one pulp and skeleton branch for an incisor, and multi-branches for a molar. Thus, providing tooth skeleton as topological guidance can better delineate the pulp with accurate connectivity. Moreover, to address low-intensity contrast between pulp and surrounding dentin, we perform a multi-scale Frangi filter (a vascular structure extraction operation) to enhance the pulp boundaries, especially on the area with limited intensity contrast.

In summary, our main contributions are as follows:

1. We propose a framework to segment dental anatomical structures from CBCT images. To the best of our knowledge, this is the first work to segment dental anatomy structures from CBCT images.

2. We propose to exploit tooth skeleton and multi-scale Frangi filter to improve pulp segmentation, by providing topological guidance and vascular structure enhancement, respectively.

3. Extensive experiments have verified the effectiveness of our method on a reasonable size of dataset. The Dice scores of enamel, pulp, and dentin reach 97.8%, 93.7%, and 98.4% respectively, indicating high application potential in clinical practice.

2. METHOD

Our framework is shown in Fig. 2, divided into two stages: tooth instance segmentation and dental anatomy segmentation. Since the first stage has been well-studied with high accuracy, our work mainly focuses on the second stage, i.e., to automatically segment detailed dental anatomical structures. Particularly, we propose to utilize tooth skeletons and Frangi images to preserve the complex pulp structure and enhance the limited intensity contrast, respectively.

2.1. Base Network

In our base network, given a CBCT image, we aim to make an instance tooth segmentation, detecting and providing each complete tooth for the following dental anatomy segmentation. We follow the state-of-the-art method, i.e., HMG-Net [4], where the first step is to localize each tooth by detecting tooth centroid and then, for each tooth, we crop the tooth image to obtain its tooth label.

Based on the prediction of HMG-Net, we feed the tooth image and tooth label into the dental anatomy segmentation network, attaining the labels of enamel, pulp, and dentin. Notably, as tooth segmentation provides more reliable whole tooth boundaries, we directly crop the tooth image with the predicted tooth label to focus on anatomy segmentation

within individual teeth. In the dental anatomy segmentation network, we also use both position attention and channel attention to improve the segmentation of pulp and dentin.

2.2. Tooth Skeleton

Although the segmentation network can already provide reasonable results on enamel and dentin, it struggles with pulp segmentation for the following reasons: 1) the shape of pulp is of large variance for different types of teeth; 2) the fragile tubular topology makes it difficult to preserve the continuity. To tackle these problems, we propose to extract tooth skeleton, obtained from the whole segmentation results, for pulp segmentation, providing reliable morphology of the tooth. As shown in Fig. 3, we observe that the tooth skeleton shares extremely similar shapes with the pulp, especially in the numbers and trends of root canals, providing effective topological information for guiding pulp segmentation.

2.3. Frangi Filter

While tooth skeletons provide faithful topology to guide pulp segmentation and preserve its connectivity, limited intensity contrast between pulp and dentin in CBCT images still results in inaccurate segmentation boundaries. To solve this problem, we propose to utilize the Frangi filter [5], a filter with excellent performance in enhancing and segmenting blood vessel-like structures, to strengthen the pulp structures, highlighting the boundary between the pulp and dentin. Specifically, we first obtain the eigenvalues of the hessian matrix of the image, and then define the pulp response function according to the special relationship between eigenvalues and different structures in the image. Note that the vessels in the MRA image are bright-white tubular structures, while the pulp in the CBCT image is gray-black (as shown in Fig. 3). Hence, the pulp response function is as follows:

$$V_{pulp} = \begin{cases} 0, & \text{if } \lambda_2 < 0 \text{ or } \lambda_3 < 0 \\ (1 - exp(-\frac{R_A^2}{2\alpha^2}))exp(-\frac{R_B^2}{2\beta^2})(1 - exp(-\frac{S^2}{2c^2})), & else \end{cases} , \quad (1)$$

where the eigenvalues are ordered: $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$, $R_A = \frac{|\lambda_2|}{|\lambda_3|}$ is crucial to distinguish plate-like and line-like structures. $R_B = \frac{|\lambda_1|}{\sqrt{|\lambda_2 \lambda_3|}}$ accounts for the deviation from blob-like structures. And $S = ||H||_F = \sqrt{\sum_{j \leq D} \lambda_j^2}$ is the second order measure of image structure. In this work, we set $\alpha = 0.1$, $\beta = 0.5$, and $c = 0.1$.

To find the most suitable Frangi filter for teeth, we take the maximum response function $V_{pulp}(s)$ within the scale s:

$$V_{pulp}(s) = \max_{s_{\min} \leq s \leq s_{\max}} V_{pulp}, \quad (2)$$

where $s_{\min}$ and $s_{\max}$ are the maximum and minimum scales at which pulp is expected to be found. These two values are set to be 1 and 10, respectively.

2.4. Loss Function

In the training stage, we adopt a binary cross-entropy loss to supervise tooth instance segmentation in the first stage. As for the second stage of dental anatomy segmentation, we design separate Dice loss functions to supervise different structures (i.e., enamel, pulp, and dentin):

$$L_{anatomy} = \frac{1}{3}(L_{enamel_{dice}} + L_{pulp_{dice}} + L_{dentin_{dice}}). \quad (3)$$

3. EXPERIMENTS

3.1. Dataset

To validate the effectiveness of our method, we have collected a dataset consisting of 200 CBCT scans from 200 patients in routine clinical care. Each image has the same resolution with spacing $0.2 \times 0.2 \times 0.2$ mm^3. Note that, we exclude incomplete teeth, wisdom teeth, and implants from the dataset, and evaluate our framework on 3672 teeth with full enamel, pulp, and dentin substructures. In our experiments, we randomly select 120 scans (i.e., with 1994 teeth) for training and 80 scans (i.e., with 1678 teeth) for testing.

3.2. Implementation Details and Evaluation Metrics

To reduce the effect of metal artifacts and multi-center intensity distribution variance, we clip the intensity values of each CBCT scan to [0, 3000] before normalizing the intensities to [0, 1]. Moreover, the tooth patches predicted in the first stage are cropped with the same size of $96 \times 96 \times 144$. This network is implemented on the PyTorch platform with one NVIDIA Tesla V100 GPU. The batch size is set as 8 for training, and the Adam optimizer is used with an initial learning rate of 1×10^{-4} decayed by 0.5 every 50 epochs.

In our experiments, we adopt the Dice Similarity Coefficient (DSC, %) and Volumetric overlap error (VOE, %) to evaluate the overall dental anatomy segmentation performance, and 95% Hausdorff distance(95%HD, mm) to evaluate the boundary distance error.

3.3. Comparison

We compare our framework with the most commonly used segmentation networks (i.e., V-Net, U-Net, and nnUNet). Table 1 lists segmentation accuracy for each dental anatomical part. It can be observed that our proposed method consistently outperforms other competing methods. Specifically, with the dual attention module, our method achieves the best performance in enamel and dentin segmentation. As for pulp segmentation, our network, combining both tooth skeleton and Frangi filter, generates more accurate pulp labels. Visual comparisons, shown in Fig 4, further demonstrate the advantage of our method. More importantly, there are two interesting observations. First, the lengths and boundaries of enamel and dentin are segmented incorrectly by the competing methods (Fig 4(a)). Instead, the results of our method match with the ground truth well, especially of the severe wear of enamel
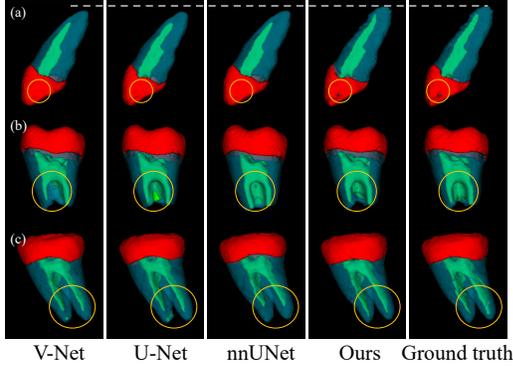
Fig. 4. Comparisons of dental anatomy segmentation by different methods.
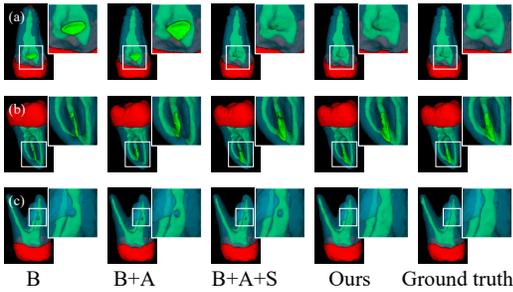
Fig. 5. Visualization of ablation study results. Our method performs better in pulp segmentation.

and the entire contour of dentin. Second, our method dramatically avoids over-segmentation, fractured, and short root canals in the pulp segmentation (Fig 4(b)(c)), verifying the advantages of the tooth skeleton and Frangi filter.

3.4. Ablation study

We also conduct ablation studies to demonstrate the effectiveness of our proposed components. As shown in Table 2, our method performs slightly better in the quantitative results. The reason is that pulp is a relatively small object, where the advance cannot be properly reflected statistically. Moreover, the segmentation accuracies of enamel and dentin are extremely high in the base network. However, the improvements in proposed components can be easily observed in the visual results. As shown in Fig 5, with the dual attention module, pulp over-segmentation can be reduced (Fig 5(a)), but remaining poor connectivity (Fig 5(b)). We adopt tooth skeletons to provide topological guidance, but incorrect boundaries between pulp and dentin and hole (Fig 5(c)) still exist due to blurred boundary. Furthermore, our method with the Frangi filter improves low-intensity contrast between pulp and dentin, attaining accurate dental anatomy segmentation.

4. CONCLUSION

In this work, we have proposed a framework to accurately segment dental anatomical structures (i.e., enamel, pulp, and dentin) from CBCT images. Specifically, tooth skeleton and multi-scale Frangi filter are introduced to improve pulp seg-

Table 1. Comparison of our method with the other three methods, i.e., V-Net, U-Net, and nnUNet.

Models	Enamel			Pulp			Dentin		
	DSC	VOE	95%HD	DSC	VOE	95%HD	DSC	VOE	95%HD
V-Net	94.2	4.2	0.311	92.0	8.3	0.420	95.2	3.3	0.334
U-Net	94.6	3.9	0.288	92.4	7.3	0.407	95.5	3.0	0.326
nnUNet	92.5	5.8	0.256	91.0	10.1	0.338	94.4	4.1	0.268
Ours	**97.8**	**2.0**	**0.132**	**93.7**	**6.0**	**0.295**	**98.4**	**1.1**	**0.126**

Table 2. Ablation studies. We report performances of our base network ('B'), and its three gradually-augmented versions such as by adding a dual attention module ('B+A'), tooth skeleton ('B+A+S'), and Frangi filter ('Ours').

Models	Enamel			Pulp			Dentin		
	DSC	VOE	95%HD	DSC	VOE	95%HD	DSC	VOE	95%HD
B	97.4	2.4	0.160	92.7	6.7	0.327	98.1	1.3	0.164
B+A	97.5	2.4	0.140	93.0	6.7	0.310	98.2	1.3	0.154
B+A+S	97.7	2.2	0.139	93.4	6.1	0.298	98.3	1.2	0.137
Ours	**97.8**	**2.0**	**0.132**	**93.7**	**6.0**	**0.295**	**98.4**	**1.1**	**0.126**

mentation. We have qualitatively and quantitatively evaluated our method, and compared it against the typical segmentation networks. The excellent performance demonstrates the advancement and potential applicability of the proposed method in clinical diagnosis and surgical treatment.

5. COMPLIANCE WITH ETHICAL STANDARDS

This study was performed in line with the principles of the Declaration of Helsinki and approved by the local institutional review board. Informed written consent was obtained from all institutional patients.

6. REFERENCES

[1] Galler et al., "Inflammatory response mechanisms of the dentine–pulp complex and the periapical tissues," *International journal of molecular sciences*, vol. 22, no. 3, pp. 1480, 2021.

[2] Hui Gao and Oksam Chae, "Individual tooth segmentation from ct images using level set method with shape and intensity prior," *Pattern Recognition*, vol. 43, no. 7, pp. 2406–2417, 2010.

[3] Zhiming Cui, Changjian Li, and Wenping Wang, "Toothnet: Automatic tooth instance segmentation and identification from cone beam ct images," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 6361–6370.

[4] Zhiming Cui, Yu Fang, Lanzhuju Mei, et al., "A fully automatic ai system for tooth and alveolar bone segmentation from cone-beam ct images," *Nature communications*, vol. 13, no. 1, pp. 1–11, 2022.

[5] Alejandro F Frangi, Wiro J Niessen, Koen L Vincken, and Max A Viergever, "Multiscale vessel enhancement filtering," in *International conference on medical image computing and computer-assisted intervention*. Springer, 1998, pp. 130–137.