TSNet: Integrating Dental Position Prior and Symptoms for Tooth Segmentation from CBCT Images

Author name(s) withheld EMAIL(S) WITHHELD
Address withheld

Editors: Under Review for MIDL 2023

Abstract

Automated dental diagnosis requires accurate segmentation of tooth from cone-beam computed tomography (CBCT) images. However, existing segmentation methods often neglect incorporating prior information and symptoms of tooth, which can cause unsatisfactory segmentation performance on tooth with symptoms. To this respect, we propose Tooth Symptom Network (TSNet), consisting of Dental Prior Guiding Data Augmentation (DPGDA) and Dental Symptom Shape Loss (DSSL), to improve segmentation performance for tooth with different clinical symptoms. Experiments show that TSNet outperforms all state-of-the-art methods across datasets with all kinds of symptoms with an average increase of 1.13% in Dice and 2.00% in IoU.

Keywords: Tooth symptoms, CBCT, Dental prior, Symptom shape loss

1. Introduction

Digital cone-beam computed tomography (CBCT) reconstruction has been shown to improve the effectiveness of dental treatment planning and management (Hao et al., 2022). A high-quality automated CBCT model reconstruction requires an accurate tooth segmentation from CBCT images (Weiss and Read-Fuller, 2019). Prior works such as U-Net (Ronneberger et al., 2015), U-Net++ (Zhou et al., 2018), MedT (Valanarasu et al., 2021), UCTransnet (Wang et al., 2022), and Semask Transformer (Semask T) (Jain et al., 2021), have been proposed to tackle various image segmentation tasks. However, these methods are general approaches that were not specifically developed for tooth segmentation, and as a result may not attain perfect accuracy as well as robustness on tooth CBCT images.

Figure 1(a) shows one normal tooth CBCT image and CBCT images with five symptoms (Decurcio et al., 2012; Fontenele et al., 2022; Liu et al., 2007; Hofmann et al., 2013; Kuo et al., 2016). Root canal therapy (RCT) leads to high-density filling imagery in the area of the Root canal, while filings lead to high density filling imagery and metal artifacts. The severity of metal artifacts differs between composite-metal and composite-resin. Multiply teeth usually appear in the middle of upper jaws, and their dental crowns are in the shape of small cones, while their roots are smaller than those of normal teeth. Permanent tooth germ lacks normal dental parts, such as dental crowns. Prosthesis leads to high-density imagery and artifacts, which can be more severe if the material is metal. Since the number of images with these symptoms is significantly smaller than that of normal images, models do not have the chance to learn symptoms thoroughly, these degradations are likely to result in shape distortion of the tooth segmentation result as well as the misclassification of high-density filling imagery or artifacts as tooth when encountering certain symptoms.

© 2023 CC-BY 4.0, Anonymous.

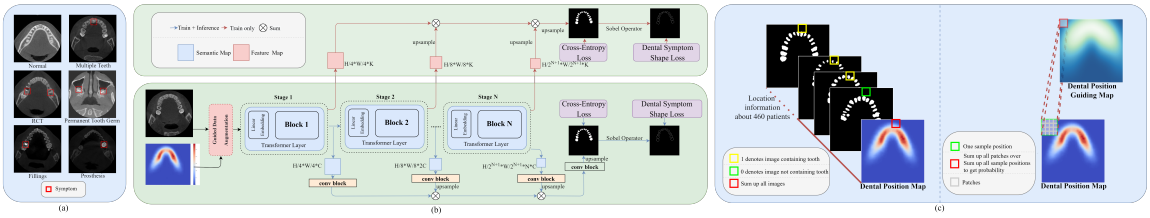

Figure 1: (a) include typical six symptoms of CBCT tooth. (b) is the pipeline of TSNet. (c) is the process of Dental Prior Guiding Data Augmentation.

Noticing that the region around the dental arch curve is both the most informative area for tooth segmentation and the area where different degradations occur, it is reasonable to let the model focus more on the region of interest. In this paper, we propose a novel method named Tooth Symptom Network (TSNet) consisting of two designs, where Dental Prior Guiding Data Augmentation (DPGDA) integrates tooth location information to draw the model's attention more towards the neighborhood of the dental arch curve, and Dental Symptom Shape Loss (DSSL) aims to minimize the difference between the predicted tooth boundary and the ground truth so that the model can be wiser when dealing with degraded images. Using the two modules, semantic priors are incorporated into the transformer layer to help extract a semantic map (Jain et al., 2021) which can improve the performance of tooth segmentation. We compare our method with state-of-the-art image segmentation methods on one normal tooth CBCT dataset and five datasets with certain symptoms. Experimental results demonstrate a superior performance of TSNet in tooth segmentation on both normal images and images with diverse symptoms.

2. Method

2.1. Dental Prior Guiding Data Augmentation (DPGDA)

Figure 1(b) exhibits the framework of TSNet. To integrate prior information about the distribution of teeth, we propose DPGDA which applies the distribution of tooth positions to guide the sampling of CBCT images for data augmentation, as is illustrated in Figure 1(c). First, we extract position information from the dataset to generate a Dental Position Map, by dividing all CBCT images into 128×128 patches and recording the number of images that have teeth in each patch. Then, instead of using pixels, patches (eg., 16×16 patches) are used to describe the size of sampling. In each position, we compute the sum of the corresponding patches in the Dental Position Map which represents the likeness that this position is sampled. After all samples are computed, sums are normalized to obtain a Dental Position Guiding Map which represents the probability distribution that guides the sampling of CBCT images in data augmentation.

2.2. Dental Symptom Shape Loss (DSSL)

DSSL is designed to constrain the shape of tooth segmentation results, especially for images that contain symptoms. Firstly, I_p is obtained by softargmax (Chapelle and Wu, 2010) on

Table 1: Comparison of TSNet with state-of-the-art methods

	Normal		RCT		Fillings		Multiply Teeth		Permanent Tooth Germ		Prosthesis	
	IoU	Dice	IoU	Dice	IoU	Dice	IoU	Dice	IoU	Dice	IoU	Dice
UNet	84.04	91.33	80.19	89.01	84.28	91.47	83.87	91.23	85.95	92.45	76.02	86.38
UNet++	84.25	91.45	77.5	87.32	83.75	91.16	83.24	90.85	85.42	92.14	69.79	82.21
UCTransNet	85.89	92.41	84.02	91.32	86.17	92.57	86.84	92.96	87.83	93.52	83.25	90.86
MedT	79.31	88.46	60.92	75.72	77.05	87.04	73.68	84.85	79.67	88.69	42.88	60.02
SemaskT	84.31	91.49	86.10	92.53	86.40	92.70	87.44	93.30	87.63	93.41	85.90	92.42
TSNet	$\mathbf{86.92}_{+1.03}$	$\mathbf{93.00}_{+0.59}$	$\mathbf{87.81}_{+1.71}$	$\mathbf{93.51}_{+0.98}$	$\mathbf{89.22}_{+2.82}$	$\mathbf{94.30}_{+1.60}$	$\mathbf{89.05}_{+1.61}$	$\mathbf{94.21}_{+0.91}$	$\mathbf{90.22}_{+2.39}$	$\mathbf{94.86}_{+1.34}$	$\mathbf{88.31}_{+2.41}$	$\mathbf{93.79}_{+1.37}$

the probability map that is outputted by the decoder. Then, the shape of the tooth is extracted in the ground truth I_g and the prediction I_p. DSSL is defined as follows:

$$loss_{DSSL} = \frac{(shape(I_g) - shape(I_p))^2}{N}, \ shape(I) = 255 \cdot \frac{(I_x)^2 + (I_y)^2}{max\{(I_x)^2 + (I_y)^2\}}, \tag{1}$$

where N denotes the number of pixels in the image, I_x and I_y denote the the Sobel operator results in the x-direction and y-direction, respectively (Pratt, 2007).

3. Experiment

We construct a tooth CBCT image dataset which consists of one training set and six test sets, containing images of 160 patient samples with varying symptoms obtained from hospitals in China during 2018-2021. The training set encompasses 1906 images from 100 **normal** individuals. The six test sets encompasses 196 images from 10 **normal** patients, 198 images from 10 **RCT** patients, 189 images from 10 **fillings** patients, 184 images from 10 **multiply teeth** patients, 97 images from 10 **permanent tooth germ** patients, and 203 images from 10 **prosthesis** patients.

To demonstrate the effectiveness of TSNet, we compare our method with five different methods: U-Net (Ronneberger et al., 2015), U-Net++ (Zhou et al., 2018), MedT (Valanarasu et al., 2021), UCTransnet (Wang et al., 2022), and Semask T (Jain et al., 2021). To ensure fairness, baseline methods are implemented in the original source code. For U-Net, U-Net++, MedT, and UCTransnet, training images are resized to $(512, 512)$. For Semask T, training images are cropped into $(512, 512)$ due to implementation details in the original source code (Jain et al., 2021). All the training hyperparameters are set to default, and evaluation metrics are computed in MMsegmentation (Contributors, 2020). Experiments show that TSNet achieves the best performance across all datasets, as shown in Table 1 (The highest value is boldfaced, and the red number in the lower right subscript indicates how much it outperforms the second-best result).

4. Conclusion

In this work, we design two modules for tooth segmentation tasks, DPGDA and DSSL, and propose an integrated tooth segmentation method TSNet. Compared with five advanced image segmentation methods on six datasets containing normal tooth CBCT images and images with different symptoms, TSNet shows a superior segmentation performance. In the future, we expect TSNet to be applied in clinical settings to assist diagnosis and treatment for dental diseases.

References

Olivier Chapelle and Mingrui Wu. Gradient descent optimization of smoothed information retrieval metrics. *Information retrieval*, 13:216–235, 2010.

MMSegmentation Contributors. MMSegmentation: Openmmlab semantic segmentation toolbox and benchmark. 2020.

Daniel Almeida Decurcio, Mike Reis Bueno, Ana Helena Gonçalves de Alencar, Olavo César Lyra Porto, Bruno Correa Azevedo, and Carlos Estrela. Effect of root canal filling materials on dimensions of cone-beam computed tomography images. *Journal of Applied Oral Science*, 20:260–267, 2012.

Rocharles Cavalcante Fontenele, Maurício do Nascimento Gerhardt, Jáder Camilo Pinto, Adriaan Van Gerven, Holger Willems, Reinhilde Jacobs, and Deborah Queiroz Freitas. Influence of dental fillings and tooth type on the performance of a novel artificial intelligence-driven tool for automatic tooth segmentation on cbct images–a validation study. *Journal of dentistry*, 119:104069, 2022.

Jin Hao, Jiaxiang Liu, Jin Li, Wei Pan, Ruizhe Chen, Huimin Xiong, Kaiwei Sun, Hangzheng Lin, Wanlu Liu, Wanghui Ding, et al. Ai-enabled automatic multimodal fusion of cone-beam ct and intraoral scans for intelligent 3d tooth-bone reconstruction and clinical applications. *arXiv preprint arXiv:2203.05784*, 2022.

Elisabeth Hofmann, Jürgen Medelnik, Martin Fink, Michael Lell, and Ursula Hirschfelder. Three-dimensional volume tomographic study of the imaging accuracy of impacted teeth: Msct and cbct comparison—an in vitro study. *The European Journal of Orthodontics*, 35 (3):286–294, 2013.

Jitesh Jain, Anukriti Singh, Nikita Orlov, Zilong Huang, Jiachen Li, Steven Walton, and Humphrey Shi. Semask: Semantically masked transformers for semantic segmentation. *arXiv preprint arXiv:2112.12782*, 2021.

Rong-Fu Kuo, Kwang-Ming Fang, Wong Ty, and Chia Yu Hu. Quantification of dental prostheses on cone-beam ct images by the taguchi method. *Journal of Applied Clinical Medical Physics*, 17(1):207–220, 2016.

Deng-gao Liu, Wan-lin Zhang, Zu-yan Zhang, Yun-tang Wu, and Xu-chen Ma. Three-dimensional evaluations of supernumerary teeth using cone-beam computed tomography for 487 cases. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 103(3):403–411, 2007.

William K Pratt. *Digital image processing: PIKS Scientific inside*, volume 4. Wiley Online Library, 2007.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.

Jeya Maria Jose Valanarasu, Poojan Oza, Ilker Hacihaliloglu, and Vishal M Patel. Medical transformer: Gated axial-attention for medical image segmentation. In *Medical Image Computing and Computer Assisted Intervention, 2021, Proceedings, Part I 24*, pages 36–46. Springer, 2021.

Haonan Wang, Peng Cao, Jiaqi Wang, and Osmar R Zaiane. Uctransnet: rethinking the skip connections in u-net from a channel-wise perspective with transformer. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, pages 2441–2449, 2022.

Robert Weiss and Andrew Read-Fuller. Cone beam computed tomography in oral and maxillofacial surgery: an evidence-based review. *Dentistry journal*, 7(2):52, 2019.

Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*, pages 3–11. Springer, 2018.