2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

A Deep Learning-based Method for Tooth Segmentation on Panoramic Dental X-ray Images

Jin Kim†, Su Yang†, Min-Hyuk Choi, Sang-Jeong Lee, Bo-Soung Jeoun, Geonsoo Kim and Won-Jin Yi*

***Abstract*— Segmentation of the tooth region is essential in the dental field to aid clinical diagnosis and make an appropriate surgical plan. However, this process is very tedious, challenging, and time-consuming. To address this problem, we propose a self-attention U-Net (SANet) for the fully automated tooth segmentation on panoramic dental X-ray images. Experimental results show that the SANet achieves higher performance than the baseline segmentation method.**

I. INTRODUCTION

Accurate tooth segmentation is an important process in the dental field to aid clinical diagnosis and make an appropriate surgical plan. However, tooth segmentation mostly depends on manual or semi-automatic interactive segmentation by experts, which is very tedious, challenging and time-consuming, and it involves prior clinical knowledge.

In this study, we propose a self-attention U-Net (SANet) for the fully automated tooth segmentation on panoramic dental X-ray images.

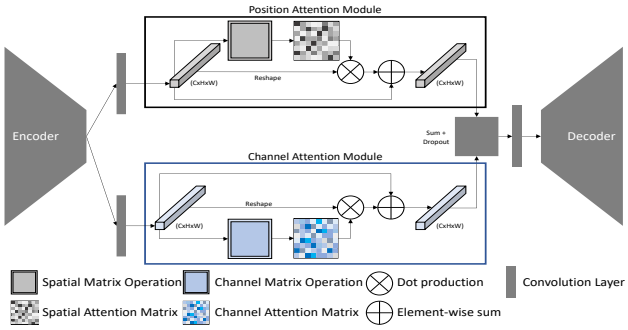

Figure 1) SANet deep network architecture

II. METHODS

Network Architecture. The proposed network architecture, described in Fig. 1, was inspired by vanilla U-Net [1] and self-attention modules [2], which consisted of encoder, decoder parts, position and channel attention modules to predict the segmentation mask. We embedded the self-attention modules after the last encoder layer for adaptively integrating high-level local features with their global dependencies.

Training setup. The proposed network was trained for 300 epochs with a mini-batch size of 32. Data augmentation was performed with rotation$(-30° - 30°)$, translation shift$(0 - 10\%)$, and zoom$(0 - 10\%)$. Soft dice loss was adopted for the multi-class segmentation. To train models, we employed Adam optimizer with learning rate of 10^{-4}.

Experimental setup. In this study, we used a total of 499 panoramic X-ray images for network training and testing. The ground truths were annotated by an expert. The number of training, validation, and test sets was randomly split into 300, 100, and 99 images, respectively. We used precision (PR), recall (RC), and dice similarity coefficient (DSC) for evaluating performance.

TABLE I. COMPARISION OF SEGMENTATION PERFORMANCE

Models	DSC	PR	RC
Simple UNet	0.924±0.01	0.911±0.02	0.944±0.02
SANet	0.928±0.01	0.921±0.02	0.942±0.02

DSC, dice similarity coefficient; PR, precision; RC, recall.

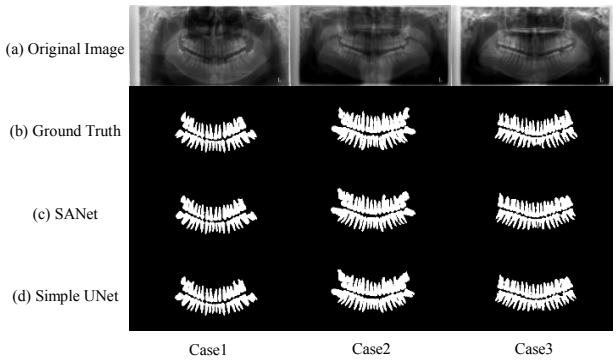

Figure 2) Examples of ground truth and segmentation results

III. RESULTS & CONCLUSION

We compared the performance of SANet with simple U-Net. The experimental results show that SANet achieves higher performance than simple U-Net (Table 1). Fig. 2 shows the representative segmentation results of SANet and simple U-Net. Experimental results demonstrate that SANet can obtain superior performance for tooth segmentation on panoramic X-ray images.

REFERENCES

[1] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.

[2] Fu, Jun, et al. "Dual attention network for scene segmentation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019

This work was supported by the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (1711138289, KMDF-PR-20200901-0147, 1711137883, KMDF_PR_20200901_0011-02)

Jin Kim is with the Interdisciplinary Program in Bioengineering, Seoul National University College of Engineering, Seoul, Korea (phone: (+82)10-2173-8802; e-mail: kimjin116@ snu.ac.kr).

S. Yang is with Medical Image Innovation Laboratory, Seoul National University, Seoul, Korea (phone: (+82)10-5100-1396; e-mail: s8431@snu.ac.kr).

†: Both authors contributed equally to this work.

W.J. Yi is with the Dep. of Oral and Maxillofacial Radiology, School of Dentistry, and the Dep. of Biomedical Radiation Sciences, GSCST, Seoul National University, Seoul, Korea (corresponding author to provide phone: (+82)2-2072-3049; e-mail: wjyi@snu.ac.kr)