Tooth Region Extraction from CBCT Images Using 3D Binary Segmentation

Muhammad Asif Jamal, Bumshik Lee*

Chosun University, Department of Information and Communication Engineering

{asif.jamal, bslee}@chosun.ac.kr, *Corresponding author

3D 이진 영상분할을 이용한 CBCT 의 치아영역 추출 방법

무함마드 아시프 자말, 이범식
조선대학교 정보통신공학과

Abstract

The three-dimensional (3D) anatomical anatomy of the human oral cavity is essential for clinical studies. In digital dentistry, precise tooth delineation from cone-beam computed tomography (CBCT) images is a crucial step toward precision dental treatment. In this work, we introduce a deep learning method for automatic, accurate, and efficient binary teeth segmentation of CBCT images. Using a modest dataset (e.g.100 images), this independent system outperforms the state-of-the-art (SOTA) methods, which requires a dataset ten times larger or more for training. Additionally, even in challenging cases with varying dental defects, the method consistently produces reliable outcomes, achieving an average dice score of 96% and an average pixel-wise accuracy of 95.78% for teeth segmentation.

Ⅰ. Introduction

A number of studies have attempted to apply deep neural networks for the segmentation of teeth and/or bony structures, motivated by the remarkable success of deep learning in computer vision and medical image processing [1-3]. Two factors make existing methods complex: first, they struggle to handle complex clinical cases, like CBCT images with different structures from individuals who have dental problems (e.g., missing teeth, misalignment, and metal artifacts), and second, they rely on large datasets that are hard to collect in medical fields. In this study, we have develop a deep-learning-based system that finds the Region of Interest (ROI) from CBCT images and segment the teeth accurately.

Our work mostly contributes to the following:

1) We have introduced a novel preprocessing function that ensure to crop teeth area.

2) The model does not need extensive dataset for training.

Ⅱ. Proposed Method

A. Dataset

We have used 100 CBCT images. The format of the images are Neuroimaging Informatics Technology Initiative (NIfTI). For training, we used 80 images and 20 for testing. Our CBCT images range in physical resolution from 0.2 to 1.0 mm. They are all normalized to a resolution of $0.4\times0.4\times0.4mm^3$.

B. Threshold-Based Tooth Region (TBTR) Cropping for Training Data

Before being fed into the V-Net [4] model, the input image is arbitrarily chopped into patches. A function called TBTR was implemented to guarantee that a part of the tooth region is present in every patch. The matching label is arbitrarily cropped, and the reduced patch is kept for training if it includes a dental region. The amount of tooth area in the patch is measured in relation to a threshold of 5000 pixels. Only the patches that satisfy this requirement are added to the training dataset; the remaining patches are eliminated. Excluding toothless patches makes the training process more focused and efficient, which improves the overall performance of the model.

C. Model Implementation

We have chosen V-Net [4] for our segmentation task after preprocessing step. Figure 1 presents the summary of our deep-learning-based system.

The network's left section is a compression path, while its right section decompresses the image till it reaches its original size. Every convolution is applied with the proper amount of padding. Each stage's convolutions are carried out using volumetric kernels made of 5x5x5 voxels. The resolution of the data decreases as it moves through the various phases of the compression process. Convolution using kernels that are 2×2×2 voxels broad and applied with stride 2 is used.

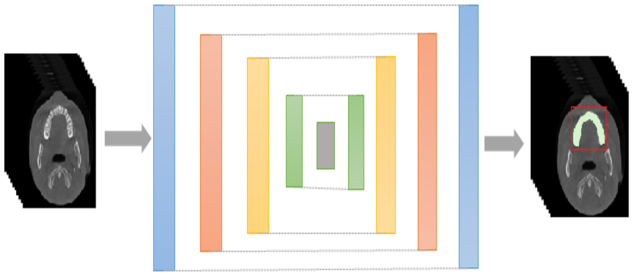

Figure 1. Illustration of our network architecture. The network model used in the middle is V-Net [4]. The input to the network is volumetric CBCT Volume, while the output is tooth region segmentation.

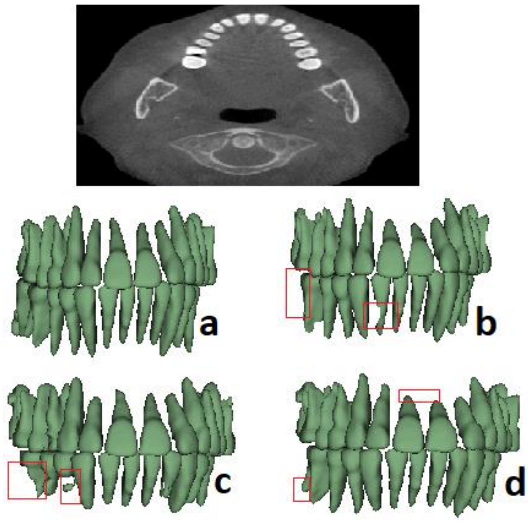

Figure 2. Qualitative results. Top is input CBCT image a) ground truth label. Segmentation result with b) U-Net[5] c)V-Net [4] d) V-Net+TBTR (ours). The red square shows the missing teeth area.

III. Experimental Results

Dice score and pixel-wise accuracy are two widely used metrics to thoroughly analyze segmentation accuracy. The results are presented in Table 1 and Figure 2, offering both quantitative and qualitative assessments, respectively. According to Table 1, the application of U-Net [5] yielded good outcomes, with an average Dice score of 93.89% and an average pixel-wise accuracy of 91.21%. The performance of the basic V-Net [4] was low, but after incorporating our novel function (TBTR), the performance improved significantly, achieving an average Dice score of 96% and an average pixel-wise accuracy of 95.76%.
As shown in Figure 2, our method successfully segments all teeth, unlike U-Net [5] and basic V-Net [4], where the segmentation results exhibit some missing teeth.

IV. Conclusion

In summary, our AI system achieves high accuracy with a limited training dataset by automating binary segmentation of CBCT images in an efficient manner.

Table 1: Segmentation Result

Algorithm	Avg. Dice (%)	Avg. Pixel wise Accuracy (%)
U-Net [5]	93.89	91.21
V-Net [4]	92.90	90.83
V-Net + TBTR (Ours)	**96.00**	**95.76**

It exhibits potential to advance digital dentistry as it outperforms the state-of-the-art techniques and retains an average Dice score of 96% for teeth segmentation.

ACKNOWLEDGMENT

This research was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea (NRF) funded by the Ministry of Education(MOE) (2021RIS-002) and funded by the Korea government (MSIT) (No. RS-2023-00217471).

REFERENCES

[1] Cui Z, Fang Y, Mei L, Zhang B, Yu B, Liu J, Jiang C, Sun Y, Ma L, Huang J, Liu Y. A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images. Nature communications. 2022 Apr 19;13(1):2096.

[2] Cui Z, Li C, Wang W. ToothNet: automatic tooth instance segmentation and identification from cone beam CT images. InProceedings of the IEEE/CVF conference on computer vision and pattern recognition 2019 (pp. 6368-6377).

[3] Cui Z, Zhang B, Lian C, Li C, Yang L, Wang W, Zhu M, Shen D. Hierarchical morphology-guided tooth instance segmentation from CBCT images. InInformation Processing in Medical Imaging: 27th International Conference, IPMI 2021, Virtual Event, June 28– June 30, 2021, Proceedings 27 2021 (pp. 150-162). Springer International Publishing.

[4] Milletari F, Navab N, Ahmadi SA. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In2016 fourth international conference on 3D vision (3DV) 2016 Oct 25 (pp. 565-571). Ieee.

[5] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. InMedical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 2015 (pp. 234-241). Springer International Publishing.