Dental Image Segmentation for Multi-Class Tooth Identification

Afnan Ghafoor, Bumshik Lee*

Chosun University

{afnanghafoor, bslee}@chosun.ac.kr

다중 분류 치아 식별을 위한 치과 이미지

아프난 가푸어, 이범식*
조선대학교

Abstract

Dental image segmentation is essential for accurate and efficient dental diagnostics. In this paper, we propose a dental image segmentation method that integrates a ResNet-50 encoder and a decoder based on the U-Net. The proposed approach is introduced for multi-class segmentation and shows outperforming segmentation performance for separating the 32 different dental classes. The high precision in dental feature identification is attributed to the ResNet-50's deep architecture, enabling the model to discern even the finest dental details. Moreover, the spatial localization characteristic of the U-Net in the decoder ensures accurate segmentation of each tooth and its associated structures. Experimental results show that our proposed method outperforms the existing state-of-the-art methods, achieving superior performance against various objective metrics in segmentation accuracy. A qualitative visual comparison for tooth segmentation also shows the outperforming segmentation capability of the proposed model.

I. Introduction

Dental imaging is an essential component of teeth diagnosis in dentistry. Panoramic X-ray images are beneficial to them and offer detailed information about the whole mouth, including teeth, upper and lower jaws, surrounding structures, and tissues, in contrast to standard intraoral X-rays concentrating on a single tooth or a small group of teeth. This comprehensive perspective aids in the discovery and diagnosis of dental diseases such as cavities, bone abnormalities, cysts, tumors, and impactions.

Despite their significance in dental diagnosis, panoramic X-ray images still have inherent issues with analysis and interpretation aspects. The inherent complexity and variance of X-ray images have frequently been shown in the existing tooth diagnosis methods in inadequate classification performance. Furthermore, classical manual interpretation of these images can be time-consuming, vulnerable to inter-observer variability, and can lead to the missed detection of minor abnormalities. In this context, the introduction of machine learning, notably deep learning, has signaled an evolution in medical imaging. Deep learning models driven by convolutional neural networks (CNNs) have shown excellent image identification skills, paving the way for novel applications in automated, accurate, and efficient image-based diagnostics.

This study aims to enhance the segmentation of panoramic X-ray dental images using advanced techniques. Unlike previous studies focusing on binary or multi-class segmentation, we propose a model for 32-class categorization, improving diagnostic accuracy and treatment efficiency. Our model combines a pre-trained ResNet-50 for feature extraction with a U-Net decoder for effective localization, outperforming other architectures in segmenting panoramic X-ray images.

II. Method

This study presents a method for segmenting dental X-ray images using a ResNet-50 encoder and a decoder inspired by the U-Net architecture. The encoder is chosen for its effectiveness in extracting hierarchical features from images, while the decoder is designed based on the U-Net architecture, known for its high accuracy in localizing objects. This approach addresses challenges in current dental image segmentation, such as inconsistent identification of dental structures and inaccuracies in boundary definitions.

The ResNet-50 encoder, pre-trained on the ImageNet1k dataset, extracts intricate features from dental panoramic images, recognizing boundary structures and irregularities, while the U-Net-like decoder addresses noise and variations.

III. Experimental Results

The dataset of 540 annotated dental panoramic X-ray images was split into training, validation, and testing sets. Using the Keras framework on the NVIDIA

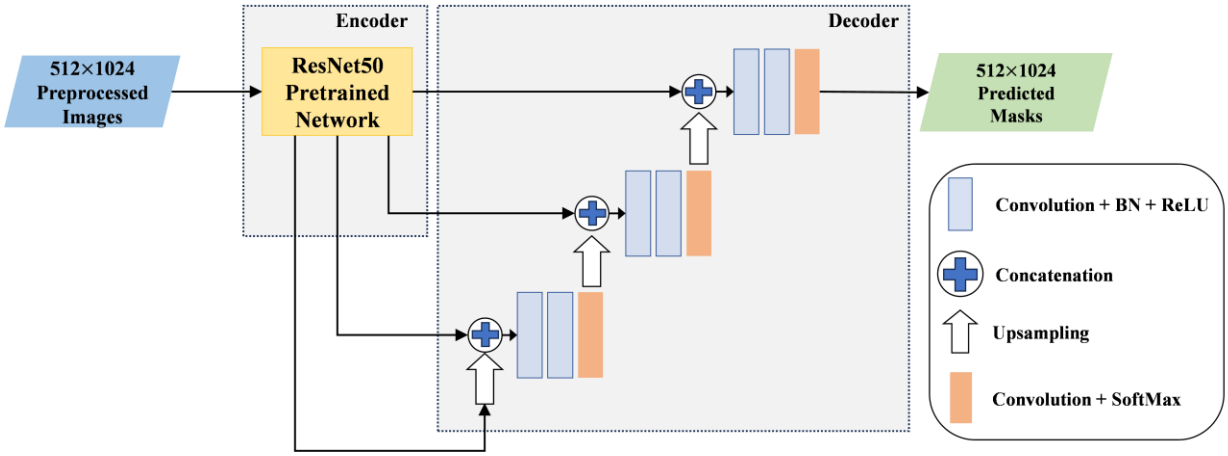

Figure 1. Overall architecture of the proposed teeth segmentation network

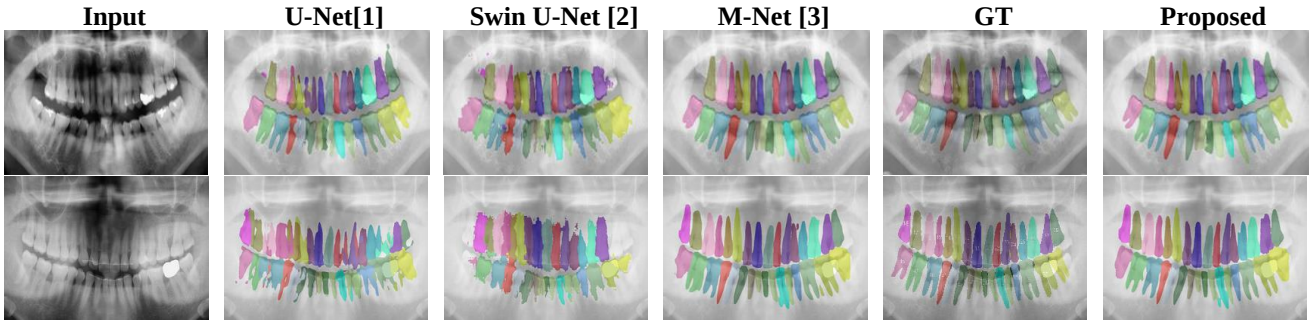

Figure 2. Visual comparison between the proposed model and the conventional model

Table 1. Comparison between the proposed model and the conventional ones.

Models	ACC	DSC	JI	Precision	Recall	Specificity
U-Net [1]	0.9720	0.7602	0.6871	0.7458	0.8366	0.9725
M-Net [3]	0.9721	0.7846	0.7132	0.7557	0.8391	0.9725
Swin U-Net [2]	0.9712	0.6348	0.5296	0.6107	0.7192	0.9721
Proposed	**0.9722**	**0.8012**	**0.7351**	**0.7602**	**0.8655**	**0.9726**

GeForce RTX 3090 GPU, our model, leveraging a ResNet-50 encoder and U-Net-like decoder, was trained and evaluated. It outperformed U-Net, M-Net[3], and Swin U-net across various metrics, achieving a Dice Coefficient of 0.8012 and Jaccard Index of 0.7351. Qualitative analysis showed sharper edges and clearer spacing between teeth, indicating superior segmentation. This balance between detail preservation and avoiding oversegmentation underscores the model's potential for dental image segmentation.

IV. Conclusion

This paper proposes a strategic approach for dental X-ray image segmentation, uniting the robust feature extraction capability of a ResNet-50 encoder with the precise localization capabilities of a U-Net-inspired decoder. The proposed approach consistently outperformed existing state-of-the-art methodologies in our analyses in terms of quantitative metrics and qualitative visual examination. These findings highlight its potential as an accurate means of dental diagnostics.

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] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015, arXiv:1505.04597.

[2] X. He, Y. Zhou, J. Zhao, D. Zhang, R. Yao and Y. Xue, "Swin Transformer Embedding UNet for Remote Sensing Image Semantic Segmentation," in IEEE Transactions on Geoscience and Remote Sensing

[3] R. Mehta and J. Sivaswamy, "M-net: A convolutional neural network for deep brain structure segmentation," in Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI), Apr. 2017, pp. 437- 440,