Detegmentation of Organs-at-Risk for Head and Neck Cancer

Shan Wang[1] 202300172023@mail.sdu.edu.cn
Jax Luo[2] jluo5@mgh.harvard.edu
Xu Qiao[1] qiaoxu@sdu.edu.cn
Nazim Haouchine[3] nhaouchine@bwh.harvard.edu
Scott Raymond[2] raymons3@ccf.org

[1] *School of Control Science and Engineering, Shandong University, Shandong, China*
[2] *Neurological Institute, Cleveland Clinic, OH, USA*
[3] *Brigham and Women's Hospital, Harvard Medical School, MA, USA*

Abstract

Segmentation of organs-at-risk (OARs) is a critical step in radiation therapy planning for head and neck cancer (HNC). We present a practical Detegmentation framework that integrates a detection network to autonomously generate box prompts for training and testing a foundation segmentation model for OARs in HNC. Our method achieves state-of-the-art performance without clinician intervention, demonstrating its strong potential for clinical application.

Keywords: Detegmentation, Detection, Segmentation, Foundational Model, Head and Neck Cancer

1. Introduction

Recent advancements in fine-tuned foundation models have demonstrated the potential to surpass conventional segmentation methods when provided with appropriately sized box prompts—rectangular regions that indicate the location of the target object (Kirillov et al., 2023; Xu et al., 2024, 2025; Fan et al., 2019; Chen et al., 2024). However, manually delineating these boxes for OARs is impractical, as it requires substantial clinician input and is susceptible to intra- and inter-rater variability. Head and neck (HNC) cancer is among the most common cancers worldwide. In radiation therapy, organs at risk (OARs) must be accurately segmented to minimize post-treatment complications. In this study, using the example of OARs for head and neck cancer(Ye et al., 2022), we propose a Detegmentation framework that integrates a detection network to autonomously generate fitted box prompts for each target OAR, enabling streamlined training of a foundation segmentation model without clinician intervention. Our method achieves state-of-the-art performance on a public benchmarks(Podobnik et al., 2023).

2. Methods

Detegmentation's framework comprises an nnDetect detection model(Baumgartner et al., 2021) and a 3D adapted Segment Anything Model (SAM)(Chen et al., 2024), which can be fine-tuned(Hu et al., 2021) for specific OARs. Initially, we train the nnDetect model to automatically predict the location and generate bounding boxes for target OARs. Subsequently, we utilize the coordinates from the detection outputs to fine-tune the 3D SAM for enhanced OAR segmentation. During the inference phase, the tested case is first processed by nnDetect to generate the corresponding box prompt, which, along with the image, is then fed into the segmentation model. Figure 1 presents an overview of the Detegmentation framework.

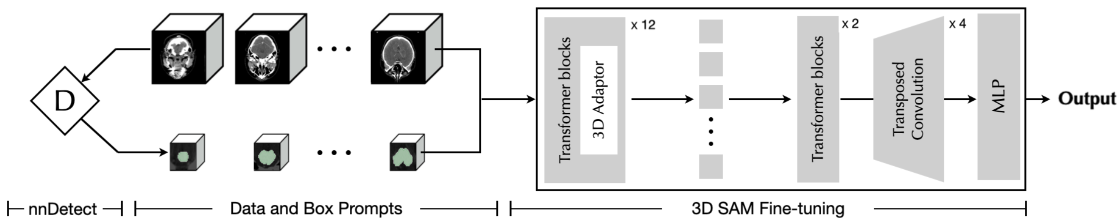

Figure 1: An overview of the proposed Detegmentation framework

3. Experiments

We assessed the proposed framework using the brainstem and left parotid gland structures from the publicly available HaN-Seg challenge dataset(Podobnik et al., 2023), which was partitioned into 30 cases for training and 12 cases for testing.

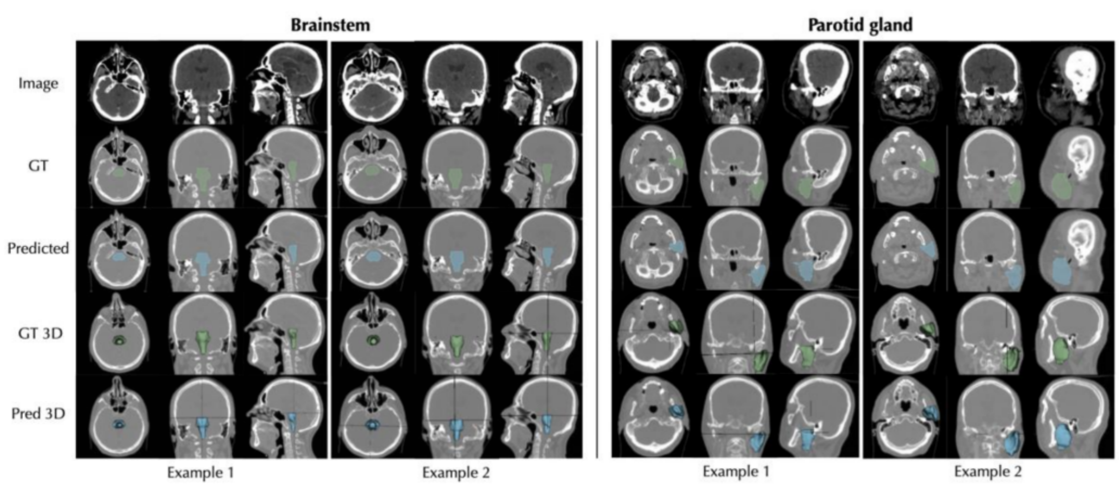

Figure 2: Qualitative results showcasing the predicted segmented brainstem and the left parotid gland labels, along with their 3D renderings. GT denotes the ground truth label

In the detection stage, the sensitivity for both the brainstem and left parotid gland is 1, with an satisfying average Dice score of 0.78 between the ground truth and predicted bounding boxes. As summarized in Table 1, the proposed Detegmentation framework achieves a mean 0.895 Dice and 1.8 HD95 of for brainstem segmentation, and a 0.886 Dice with a 2.0 HD95 for left parotid gland segmentation, surpassing the top performing nnU-Net(Isensee et al., 2018) and DynUNet(Yang et al., 2025) methods in the HaN-seg challenge. Figure 2 presents qualitative results for the segmented brainstem and left parotid gland.

Table 1: Quantitave segmentation results comparing the proposed Detegmentation framework with leading methods in the HaN-seg challenge

OAR	**Mean**	**nnU-Net**	**DynUNet**	**Ours**
Brainstem	DSC	0.885	0.849	**0.895**
	HD95	3.9	4.7	**1.8**
Parotid gland(L)	DSC	0.867	0.851	**0.886**
	HD95	5.1	5.1	**2.0**

4. Discussion and Conclusion

We present a practical Detegmentation framework that integrates a detection network to autonomously generate box prompts for training and testing a foundation segmentation model for organs at risks for head and neck cancer. Our method achieves state-of-the-art performance without clinician intervention, demonstrating its strong potential for clinical application. Future work will focus on expanding the framework to additional OARs, particularly smaller structures, and developing strategies to address cases where the detection model exhibits suboptimal performance.

References

Michael Baumgartner, Paul F Jäger, Fabian Isensee, and Klaus H Maier-Hein. nndetection: a self-configuring method for medical object detection. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part V 24*, pages 530–539. Springer, 2021.

Cheng Chen, Juzheng Miao, Dufan Wu, Aoxiao Zhong, Zhiling Yan, Sekeun Kim, Jiang Hu, Zhengliang Liu, Lichao Sun, Xiang Li, et al. Ma-sam: Modality-agnostic sam adaptation for 3d medical image segmentation. *Medical Image Analysis*, 98:103310, 2024.

Guoxin Fan, Huaqing Liu, Zhenhua Wu, Yufeng Li, Chaobo Feng, Dongdong Wang, Jie Luo, WM Wells, and Shisheng He. Deep learning–based automatic segmentation of lumbosacral nerves on ct for spinal intervention: a translational study. *American Journal of Neuroradiology*, 40(6):1074–1081, 2019.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021.

Fabian Isensee, Jens Petersen, Andre Klein, David Zimmerer, Paul F Jaeger, Simon Kohl, Jakob Wasserthal, Gregor Koehler, Tobias Norajitra, Sebastian Wirkert, et al. nnu-net: Self-adapting framework for u-net-based medical image segmentation. *arXiv preprint arXiv:1809.10486*, 2018.

Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. Segment anything. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4015–4026, 2023.

Gašper Podobnik, Primož Strojan, Primož Peterlin, Bulat Ibragimov, and Tomaž Vrtovec. Han-seg: The head and neck organ-at-risk ct and mr segmentation dataset. *Medical physics*, 50(3):1917–1927, 2023.

Guoping Xu, Xiaoxue Qian, Hua-Chieh Shao, Jax Luo, Weiguo Lu, and You Zhang. A segment anything model-guided and match-based semi-supervised segmentation framework for medical imaging. *Medical Physics*, 2025.

Zhe Xu, Donghuan Lu, Jie Luo, Yefeng Zheng, and Raymond Kai-yu Tong. Separated collaborative learning for semi-supervised prostate segmentation with multi-site heterogeneous unlabeled mri data. *Medical Image Analysis*, 93:103095, 2024.

Jin Yang, Daniel S Marcus, and Aristeidis Sotiras. Dynamic u-net: adaptively calibrate features for abdominal multiorgan segmentation. In *Medical Imaging 2025: Computer-Aided Diagnosis*, volume 13407, pages 326–334. SPIE, 2025.

Xianghua Ye, Dazhou Guo, Jia Ge, Senxiang Yan, Yi Xin, Yuchen Song, Yongheng Yan, Bing-shen Huang, Tsung-Min Hung, Zhuotun Zhu, et al. Comprehensive and clinically accurate head and neck cancer organs-at-risk delineation on a multi-institutional study. *Nature communications*, 13(1):6137, 2022.