Prompt Guided Adaptation of the Segment Anything Model Enhances Small Organ Segmentation In Head and Neck Radiotherapy

Y. Xia[1,2], S. Jabor[3], P. New[1,2], R. Salazar[1,2], A.Roy[1,2] , N. Kirby[1,2], N. Papanikolaou[1,2]

[1]University of Texas Health Science Center at San Antonio, [2]University of Texas at San Antonio, [3] St. Mary's University

Mays Cancer Center
UT Health San Antonio MD Anderson Cancer Center
UT San Antonio

INTRODUCTION

- Accurate segmentation of head and neck organs at risk is essential for high precision radiotherapy, yet small, low contrast structures (e.g., optic nerves, cochleae, chiasm) remain challenging for existing methods.
- While 3D nnUNet achieves strong performance, it requires extensive training data, high computational cost and limited interaction; SAM-Med3D enables flexible prompt-based segmentation but underperforms on very small organs.
- We fine-tuned a 3D SAM model (SAM-FT-HN) using domain specific adaptation, interactive prompting, and ROI guided tiny organ enhancement module, and evaluated its accuracy, efficiency, and clinical applicability against SAM-Med3D and 3D nnUNet.

METHODS

- A total of 211 anonymized head and neck CT datasets with 24 delineated organs at risk and targets were standardized, with 80% used for training, validation and 20% for testing.
- SAM-FT-HN was fine-tuned using domain-specific adaptations including adaptive small-organ thresholding, spherical ROI extraction, and binary search threshold calibration to enhance small-organ segmentation, while 3D nnUNet was trained using its automated configuration pipeline.
- Segmentation performance was assessed using Dice similarity coefficient, Jaccard index, precision, recall, specificity, and accuracy.

FIGURES

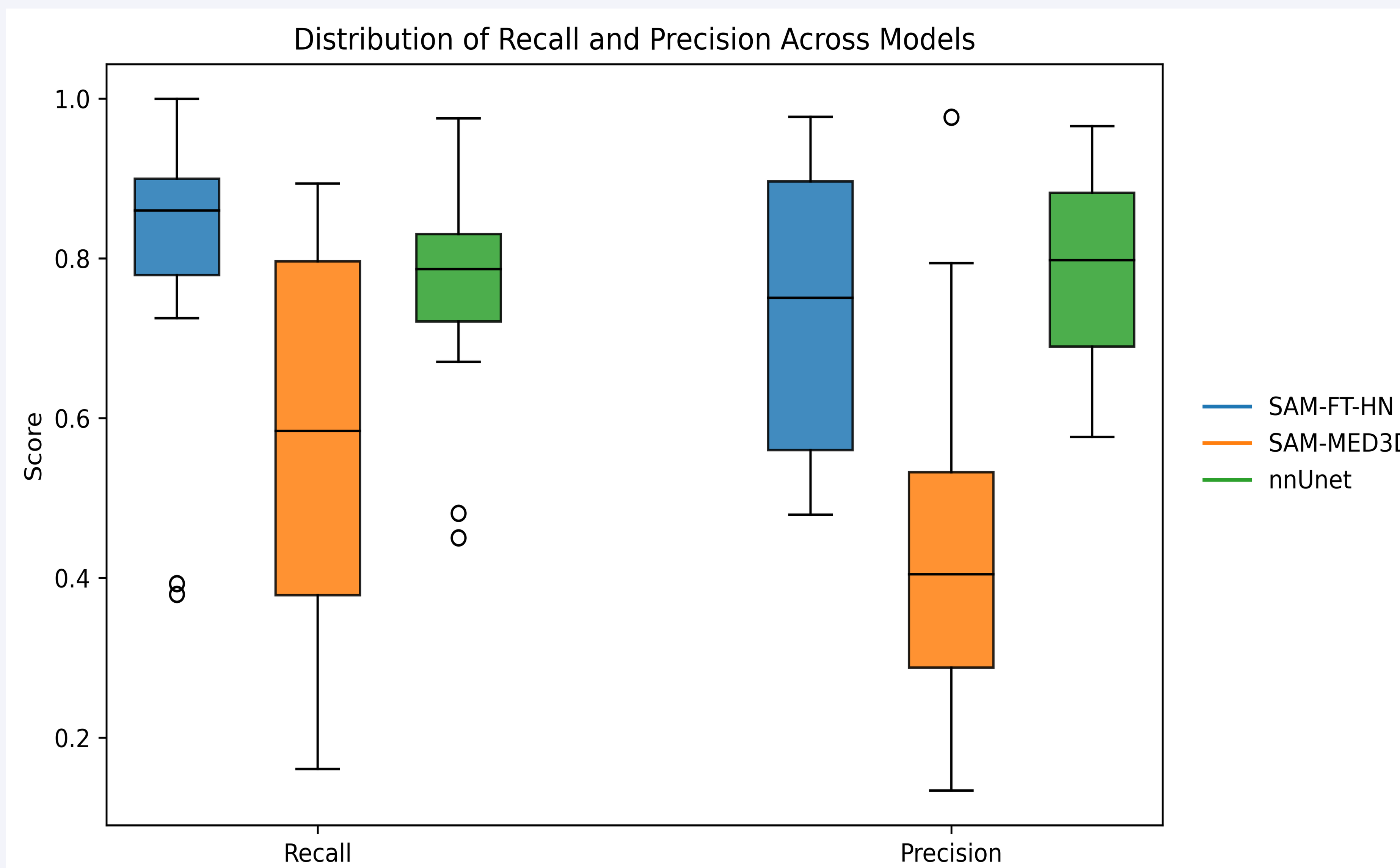

Figure 1: Distributions of recall and precision for SAM-FT-HN, SAM-MED3D, and nnUNet.

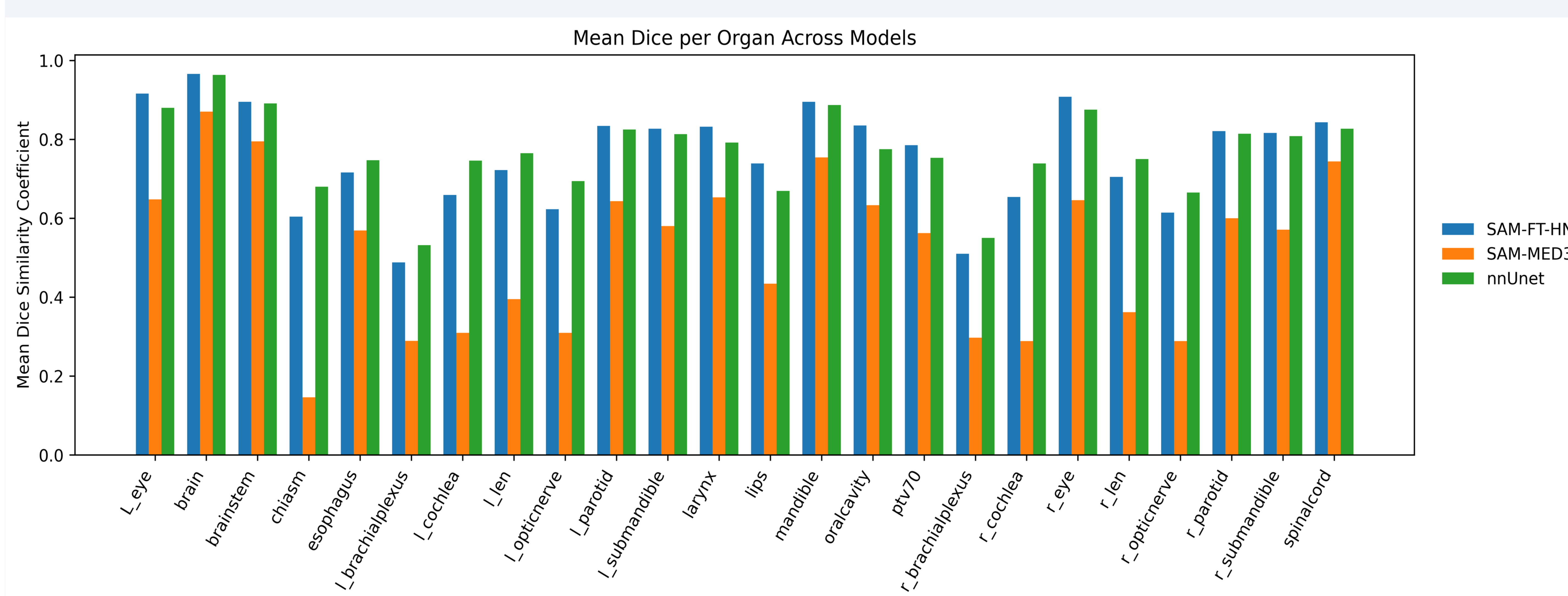

Figure 2: Comparison of organ-level mean Dice scores across segmentation models.

RESULTS

- For large OARs, SAM-FT-HN achieved Dice performance comparable to nnUNet (0.802 vs. 0.788, $p = 0.0020$) and significantly outperformed SAM-Med3D (0.605, $p < 10^{-4}$).
- For small OARs, SAM-FT-HN showed a substantial Dice improvement over SAM-Med3D (ΔDice = +0.355, $p = 0.0156$) while maintaining performance close to nnUNet (ΔDice = −0.066).
- Across all OARs, SAM-FT-HN demonstrated significantly higher recall than both SAM-Med3D ($p = 0.0011$) and nnUNet ($p = 0.0315$), with no significant difference in precision compared to nnUNet ($p = 0.114$).
- SAM-FT-HN significantly reduced the average time required for manual contouring, enhancing efficiency in clinical workflows.

CONCLUSION

- Fine-tuning a foundation model with anatomically informed prompting and adaptive thresholding substantially improves small-organ segmentation, addressing a key limitation of the original model.
- SAM-FT-HN achieves competitive segmentation accuracy while maintaining faster inference compared with conventional deep learning models.
- The model provides a robust, practical and clinically promising solution for real time head and neck radiotherapy contouring and adaptive treatment workflows.

REFERENCES

1. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. (2023). Segment Anything. arXiv:2304.02643.
2. Wang, H., Guo, S., Ye, J., Deng, Z., Cheng, J., Li, T., Chen, J., Su, Y., Huang, Z., Shen, Y., Fu, B., Zhang, S., He, J., & Qiao, Y. (2023). SAM-Med3D: Towards general-purpose segmentation models for volumetric medical images. arXiv:2310.15161
3. sensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods, 18(2), 203–211.

CONTACT INFO

Yuqing Xia
UT Health Science Center at San Antonio
University of Texas at San Antonio

Email: xiay1@livemail.uthscsa.edu