Interpretable Semi-Supervised Swin Transformer for Brain Tumor Segmentation

Vignesh Raja, Sowmya Gonugunta

Department of Computer Science, Georgia State University

INTRODUCTION

Motivation

Accurate brain tumor segmentation in MRI images is critical for diagnosis and treatment planning. However, obtaining large, fully annotated datasets for training deep learning models is challenging and time-consuming. Furthermore, the lack of interpretability in many deep learning models hinders their clinical adoption

- **Brain tumor segmentation** (Necrotic Core, Edema, Enhancing Tumor) from MRI is vital for diagnosis/treatment.
- **CNN limitations**: Need large labeled datasets (costly!), struggle with long-range context.
- **Clinical Need**: Explainable AI (XAI) is crucial for trust and adoption. "Black box" models are insufficient.

DATA

Dataset: BraTS 2023 Adult Glioma (Part of BraTS 2025 challenge)

- **Source**: RSNA-ASNR-MICCAI Brain Tumor Segmentation (BraTS) Challenge - the standard benchmark.
- **Data**: Multi-modal MRI scans per patient (T1, T1Gd, T2, T2-FLAIR).
- **Annotations**: Expert-provided ground truth segmentations for: Necrotic and Non-Enhancing Tumor Core (NCR/NET), Peritumoral Edema (ED), Gadolinium-Enhancing Tumor (ET)
- **Usage**: Dataset partitioned into Labeled (40%) and Unlabeled (60%) sets to evaluate semi-supervised learning performance under data scarcity.

RESEARCH

METHODOLOGY

We propose a novel framework leveraging a Swin Transformer architecture in a semi-supervised learning setting with an emphasis on **interpretability.**

- **Swin Transformer**: Captures local & global context efficiently.
- **Semi-Supervised Learning (SSL)**: Leverages unlabeled data via Consistency Regularization for data efficiency.
- **Interpretability**: Attention Rollout visualizes model focus, enhancing trust.

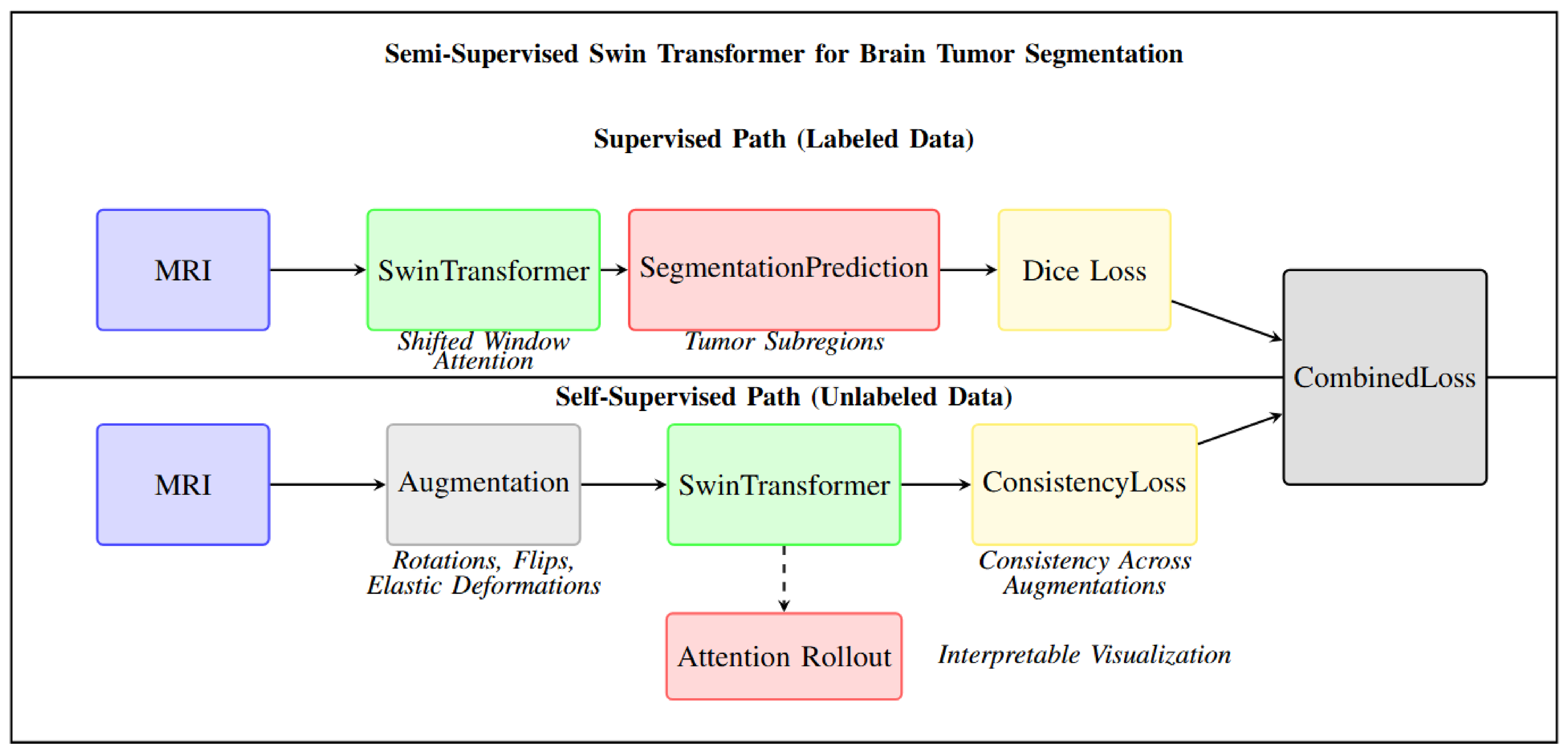

Figure 2. The proposed semi-supervised learning framework with Swin Transformer. The upper path processes labeled data with supervised Dice loss, while the lower path enforces consistency across augmented views of unlabeled data. The combined loss guides model optimization to leverage both data sources effectively. Attention Rollout provides interpretable visualization of model's focus areas.

RESULTS

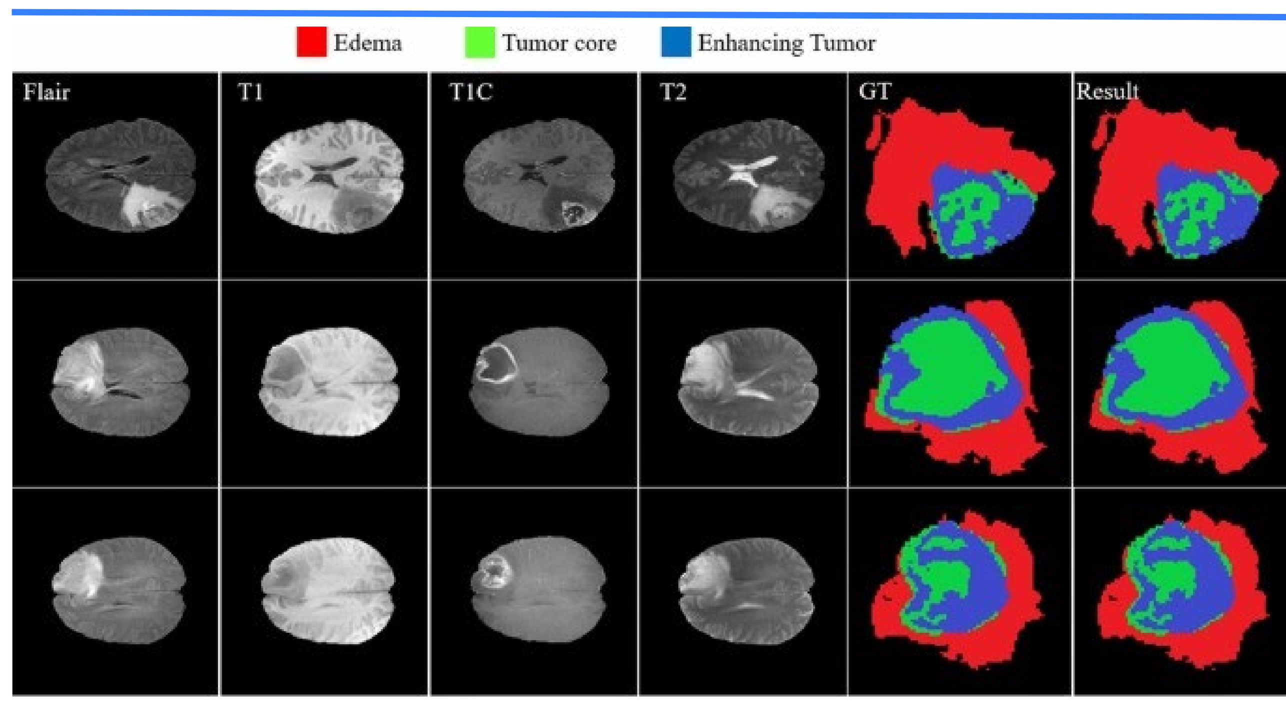

Figure 2: The results of brain tumor segmentation using the proposed strategy (the blue, green, and red colors are enhanced, core, and edema regions respectively).

Segmentation Examples & Interpretability: Improved segmentation accuracy, especially at boundaries. Attention Rollout highlights clinically relevant areas.

Ablation Insights

- SSL Boost: "Consistency Regularization improved Dice score".
- Interpretability: "Attention Rollout confirms model focuses on tumor regions".

CONCLUSIONS

CONCLUSION

We presented an interpretable, semi-supervised Swin Transformer for brain tumor segmentation. It achieves superior accuracy and data efficiency compared to baselines (U-Net) while providing crucial interpretability via Attention Rollout, making it promising for clinical application.

REFERENCES

[**BraTS Dataset**] U. Baid et al., "The Brain Tumor Segmentation (BraTS) Challenge 2023: Intracranial Glioma Radiogenomic Classification," arXiv preprint arXiv:2311.06901, 2023.

[**Swin Transformer**] Z. Liu et al., "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), pp. 10012-10022, 2021.

[**Consistency Regularization - FixMatch**] K. Sohn et al., "FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence," Adv. Neural Inf. Process. Syst. (NeurIPS), 2020.

[**Attention Rollout**] S. Abnar and W. Zuidema, "Quantifying Attention Flow in Transformers," Proc. 58th Annu. Meet. Assoc. Comput. Linguist. (ACL), pp. 4190-4197, 2020.

[**U-Net Baseline**] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Proc. Med. Image Comput. Comput. Assist. Interv. (MICCAI), pp. 234-241, 2015.

[**MONAI Framework**] M. J. Cardoso et al., "MONAI: An open-source framework for deep learning in healthcare," Medical Image Analysis (MedIA), vol. 79, p. 102448, 2022.

[**Data Augmentation**] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," J. Big Data, vol. 6, no. 1, p. 60, 2019.

[**Fig. 2**] Ranjbarzadeh, R., Bagherian Kasgari, A., Jafarzadeh Ghoushchi, S. et al. Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. Sci Rep 11, 10930 (2021).https://doi.org/10.1038/s41598-021-90428-8

CONTACT INFO

vraja@gsu.edu
sgonugunta1@gsu.edu