

On the Role of MRI Sequences in Cross-Dataset Generalization for Brain Tumor Segmentation

Henrique Zan Grande*, João G. Pitol*, Lucas B. Schuck*, Rafael V. Serenato*,
Rayson Laroca*, and Andre Gustavo Hochuli*

*Pontifical Catholic University of Paraná, Curitiba, Brazil

{joao.pitol,lucas.schuck,rafael.serenato}@pucpr.edu.br

{henrique.zgrande,rayson,aghochuli}@ppgia.pucpr.br

Abstract—Brain tumor segmentation in magnetic resonance imaging (MRI) is a critical task for diagnosis and treatment planning. Despite the success of deep learning architectures such as U-Net and its variants, performance degradation across datasets remains a major challenge, particularly under domain shift and limited annotated data. To address this issue, this study systematically evaluates how individual MRI sequences influence model robustness across two well-known datasets. A ResUNet-based framework is employed, where each modality is trained independently to isolate its effect under a controlled cross-dataset evaluation protocol with tumor size stratification, without target-domain training, or with limited domain adaptation. Results show that the T2f/FLAIR sequence achieves the best cross-dataset performance, with Dice scores exceeding 75%. It consistently outperforms other modalities across most tumor size ranges, while multi-sequence training further improves performance. Additionally, even limited target-domain adaptation yields rapid initial gains, reducing the need for extensive annotations and costly retraining. Our source code is publicly available at https://github.com/henrique-zan/brain_tumor_segmentation/.

I. INTRODUCTION

Automatic brain tumor segmentation in magnetic resonance imaging (MRI) is a prominent application of deep learning in medical computer vision [1]–[3]. This task plays a critical role in diagnosis, surgical planning, and clinical follow-up, supporting clinicians in analyzing imaging data. Advances in convolutional neural network architectures, especially U-Net [4] and its variants, have achieved strong performance on benchmark datasets and competitions such as Brain Tumor Segmentation (BraTS) [5], [6], establishing them as a standard approach for brain lesion segmentation.

Despite these advances, segmentation models often fail to generalize across datasets [7]. Differences in acquisition protocols, scanners, patient populations, and image characteristics create domain shifts that can substantially degrade performance on unseen data. This cross-dataset generalization problem is further aggravated by the limited availability of annotated medical images, which restricts model robustness and real-world applicability [7]. The choice of MRI sequence may be particularly important because T1n, T1c, T2w, and T2f/FLAIR provide complementary information about tissue contrast, edema, and tumor boundaries [8], [9].

Previous studies have primarily sought to maximize segmentation accuracy through architectural improvements or

multimodal fusion [8], [10]. However, fewer studies have systematically isolated the contribution of each MRI sequence under cross-dataset domain shift. A modality-specific analysis can reveal which sequences yield more transferable representations and can guide model design when computational resources or MRI sequences are limited. Moreover, evaluating adaptation with small amounts of labeled target-domain data can clarify the trade-off between annotation effort and performance improvement.

Based on this motivation, we investigate the following research questions:

- **(RQ1)** How does segmentation performance degrade across domains as a function of MRI modality?
- **(RQ2)** Which MRI sequences exhibit greater robustness in cross-dataset binary brain tumor segmentation?
- **(RQ3)** What is the relationship between target-domain annotation effort and performance improvement for different MRI modalities?

To address these questions, we conduct a systematic comparative analysis of cross-dataset generalization across MRI sequences for binary brain tumor segmentation (tumor vs. non-tumor). A standardized pipeline based on the widely used 2D ResUNet [10] is employed, where each modality is trained and evaluated independently under a controlled protocol [7] to isolate its contribution to robustness. Additionally, we apply fine-tuning to the best-performing modality to assess the impact of limited target-domain data on generalization, with emphasis on a controlled and analysis-driven evaluation setting.

This work makes four main contributions: (i) a controlled evaluation of modality-specific robustness under domain shift; (ii) a quantitative assessment of the generalization gap across modalities; (iii) the identification of the most robust MRI sequence for cross-dataset binary segmentation; and (iv) an analysis of data-efficient domain adaptation, relating annotation budget to performance gains.

The remainder of this paper is organized as follows. Section II reviews related work and highlights the research gap. Section III describes the problem, datasets, materials, and methodology. Section IV presents the experimental setup, evaluation metrics and discusses the results. Finally, Section V concludes the paper and outlines directions for future work.

II. LITERATURE REVIEW

The literature on brain tumor segmentation in MRI is largely dominated by convolutional neural network-based approaches, particularly U-Net [4] and its variants [1], [2]. Prior work has focused on architectural improvements and multimodal fusion strategies to enhance segmentation accuracy. Despite these advances, challenges related to domain shift and cross-dataset generalization remain insufficiently addressed.

U-Net has become a foundational architecture for brain tumor segmentation because its encoder-decoder structure and skip connections combine contextual representation with spatial detail [11]. However, performance can deteriorate in heterogeneous imaging settings, where variations in tumor morphology and image intensity hinder generalization [12].

Recent variants such as ResUNet, ResUNet+, and nnU-Net address some of these limitations through residual blocks, attention mechanisms, and automated configuration [10], [13], [14]. ResUNet combines the localization capabilities of U-Net with residual learning, which facilitates optimization and supports deeper feature extraction [13]. Mangayarkarasi et al. [10] reported an accuracy of 93.75% for ResUNet on the LGG MRI dataset. Metlek & Çetiner [14] reported that ResUNet+ improved the mean Dice score for whole-tumor segmentation by up to 5.58% over the compared methods. In parallel, nnU-Net has demonstrated strong and reproducible performance by automatically configuring architectural and preprocessing choices [13].

Multimodal fusion has also been extensively studied because T1, T1c, T2, and T2f/FLAIR provide complementary contrast and structural information. Studies by Yin et al. [8] and Xing et al. [9] show that combining modalities can improve segmentation accuracy and robustness over single-sequence inputs, with reported Dice scores above 90%. Three-dimensional architectures, including ResUNet variants with dual or mirrored encoders, can better exploit volumetric continuity and inter-slice consistency. However, they require more computation and may be more sensitive to class imbalance [9].

Although these methods achieve strong in-domain results, their robustness across heterogeneous datasets remains limited. Differences in acquisition protocols, scanners, and intensity distributions can induce substantial performance degradation on unseen domains [7], [14]. Multimodal inputs may reduce this sensitivity, but the individual contribution of each MRI sequence to cross-domain robustness has not been systematically quantified, especially under limited annotation budgets. This study addresses that gap through a controlled modality-specific evaluation and a data-efficient domain-adaptation analysis.

III. PROBLEM STATEMENT

Consider a model trained on one BraTS cohort and deployed on another using a single MRI sequence. The central problem is to determine how much segmentation performance is lost during this transfer, whether some sequences preserve accuracy better than others, and how efficiently a small labeled target subset can compensate for the shift.

To isolate these factors, we evaluate every modality with the same 2D ResUNet architecture, data preparation procedure, and optimization settings. The comparison covers in-domain evaluation, direct cross-dataset transfer, joint training on both datasets, and incremental target-domain fine-tuning.

The following section details the experimental design and methodological choices adopted to ensure a controlled and reproducible evaluation. We first describe the datasets and their characteristics, followed by the partitioning and preprocessing pipeline applied uniformly across all experiments. Next, the 2D ResUNet architecture and training configuration are presented, including loss functions and optimization strategy. Finally, we outline the evaluation.

A. Datasets

The *Brain Tumor Segmentation* (BraTS) challenge datasets from 2020 [5] and 2024 (Adult glioma segmentation task) [6] were selected due to their widespread adoption as standardized benchmarks and their complementary characteristics for evaluating model robustness. BraTS 2020 [5] comprises 369 multimodal MRI studies and serves as a well-established, controlled benchmark for method comparison. In contrast, BraTS 2024 [6] extends this scale to a substantially larger and more heterogeneous cohort, totaling approximately 1,350 cases, reflecting recent advances in data collection and acquisition diversity. Both datasets provide four MRI sequences (T1n, T1c, T2w, and T2f/FLAIR) along with expert-annotated segmentation masks for tumor subregions, as shown in Fig. 1.

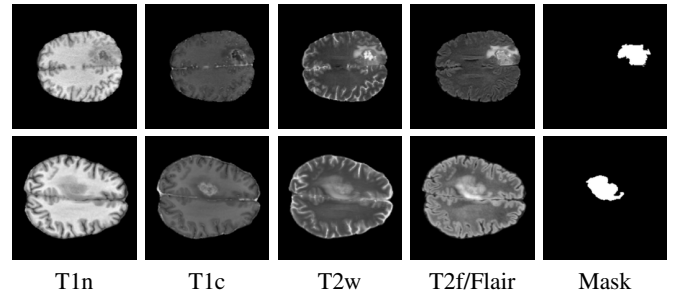


Fig. 1: Representative cases from BraTS 2020 (top) and BraTS 2024 (bottom) across the four MRI sequences and the corresponding binary mask.

Although the datasets follow standardized annotation and preprocessing practices, they differ in patient characteristics, tumor morphology, scanners, and acquisition protocols. For example, BraTS 2020 comprises pre-operative adult glioma scans from a limited set of contributing institutions, whereas BraTS 2024 is substantially larger and more heterogeneous, aggregating post-treatment acquisitions from a broader range of sites and scanners. These differences create a realistic domain shift while preserving a common segmentation task. The two datasets are therefore suitable for evaluating cross-dataset generalization and modality-specific robustness. Whenever available, we retained the original patient-level organization and naming conventions to support reproducibility.

Data preparation followed a patient-level holdout strategy, with 70% of the patients assigned to training, 15% to validation, and 15% to testing. This partitioning prevents data leakage by ensuring that slices from the same patient do not appear in multiple subsets. All images were resized to 240×240 pixels using the same preprocessing pipeline for both datasets. For each MRI modality, we extracted 2D slices containing tumor tissue to focus training and evaluation on informative regions and reduce computational cost. The original multiclass annotations were converted into binary masks representing tumor and non-tumor regions.

Data augmentation was applied only to the training set and included 90° rotations and horizontal and vertical flips. We deliberately excluded more aggressive intensity and geometric transformations because they may alter anatomical structures or create unrealistic tumor appearances [7], [15].

B. Architecture

The adopted architecture is a 2D ResUNet based on a ResNet-34 encoder [10] (see Fig. 2). ResUNet extends U-Net with residual learning to improve optimization, feature representation, and training stability [10], [13], [14]. The encoder uses a pretrained ResNet-34 backbone with channel widths of 64, 128, 256, and 512 to extract hierarchical features. Residual identity mappings facilitate gradient propagation, while the encoder progressively captures multi-scale contextual information. Skip connections transfer fine-grained spatial features from each encoder stage to the corresponding decoder stage.

The decoder contains four upsampling stages with channel progression $512 \rightarrow 256 \rightarrow 128 \rightarrow 64 \rightarrow 64$. Each stage combines transposed convolution, feature concatenation, and two 3×3 Conv-BatchNorm-ReLU blocks for feature fusion and spatial refinement. Dropout with $p = 0.2$ is applied within the convolutional blocks to reduce overfitting, particularly in limited-data and cross-dataset settings. The network receives a single-channel MRI slice and produces a binary segmentation map through a final 1×1 convolution.

Training Protocol and Hyperparameters. Models were trained independently for each MRI modality using a supervised objective that combines Dice loss and binary cross-entropy [7], [9]. Training was conducted for up to 80 epochs with a batch size of 32 and an initial learning rate of 1×10^{-3} . The learning rate was reduced adaptively according to validation performance. Early stopping with a patience of eight epochs was based on the validation loss. For each experiment, we retained the checkpoint with the lowest validation loss¹.

IV. EXPERIMENTS AND DISCUSSION

The experimental analysis comprises three parts using BraTS 2020 [5] and BraTS 2024 [6]. First, Section IV-A compares in-domain, cross-dataset, and multi-domain protocols to answer RQ1 and RQ2. Second, Section IV-B measures how increasing proportions of labeled target-domain data affect

fine-tuning performance, addressing RQ3. Third, Section IV-C examines tumor-size quartiles and representative segmentation cases. The size-stratified analysis accounts for the greater sensitivity of overlap-based metrics such as Dice to small absolute errors in small lesions.

Performance was measured using the Dice coefficient and Intersection over Union (IoU), two overlap-based metrics widely adopted in medical image segmentation [14], [16]. Both metrics quantify agreement between the predicted and reference masks. Metrics were computed over all tumor-containing test slices and averaged within each test set to enable consistent comparisons across modalities and experimental settings.

A. Generalization Evaluation (RQ1/RQ2)

To rigorously assess cross-domain robustness, we design evaluation protocols that progressively increase the distributional discrepancy between training and test data, enabling a controlled analysis of generalization behavior across MRI modalities.

Baseline. In the baseline setting, models are trained and evaluated within the same domain (BraTS 2020 or BraTS 2024 independently), providing an upper-bound reference under in-distribution conditions. As training, validation, and test sets share the same underlying data distribution, this scenario isolates the intrinsic segmentation capability of each modality without domain shift effects.

Cross-Dataset. To explicitly evaluate generalization under domain shift, models trained on one dataset are evaluated directly on the other without adaptation. Two configurations are considered: BraTS 2024 $\rightarrow$ 2020 and BraTS 2020 $\rightarrow$ 2024. This protocol introduces discrepancies in acquisition protocols, intensity distributions, and population characteristics, providing a realistic assessment of performance degradation across domains. It enables quantifying modality-specific robustness and identifying which MRI sequences are inherently more resilient to distributional shifts [7].

Multi-Domain Training. In this setting, training data from BraTS 2020 and BraTS 2024 are jointly used in balanced proportions (50/50), while preserving patient-level holdout splits to avoid data leakage. This protocol investigates whether exposure to heterogeneous data distributions improves generalization by promoting the learning of domain-invariant representations, and assesses the extent to which shared features across domains can mitigate performance degradation.

The results summarized in Table I provide key insights into cross-dataset generalization across MRI modalities. A consistent gap between in-domain and cross-domain performance is observed, confirming the sensitivity of segmentation models to domain shift (RQ1). While baseline results establish an upper bound, with FLAIR (T2F) and T2W achieving the highest Dice scores (e.g., 77.3–78.3 and 73.3–74.2), cross-dataset evaluation reveals substantial performance degradation, particularly for T2F (77.3 $\rightarrow$ 58.2) and T2W (73.3 $\rightarrow$ 62.3). Despite this, T2F maintains the highest cross-domain performance (up to 76.2),

¹To ensure full reproducibility, all models, trained weights, source code, and experimental artifacts are publicly available at https://github.com/henrique-zan/brain_tumor_segmentation/.

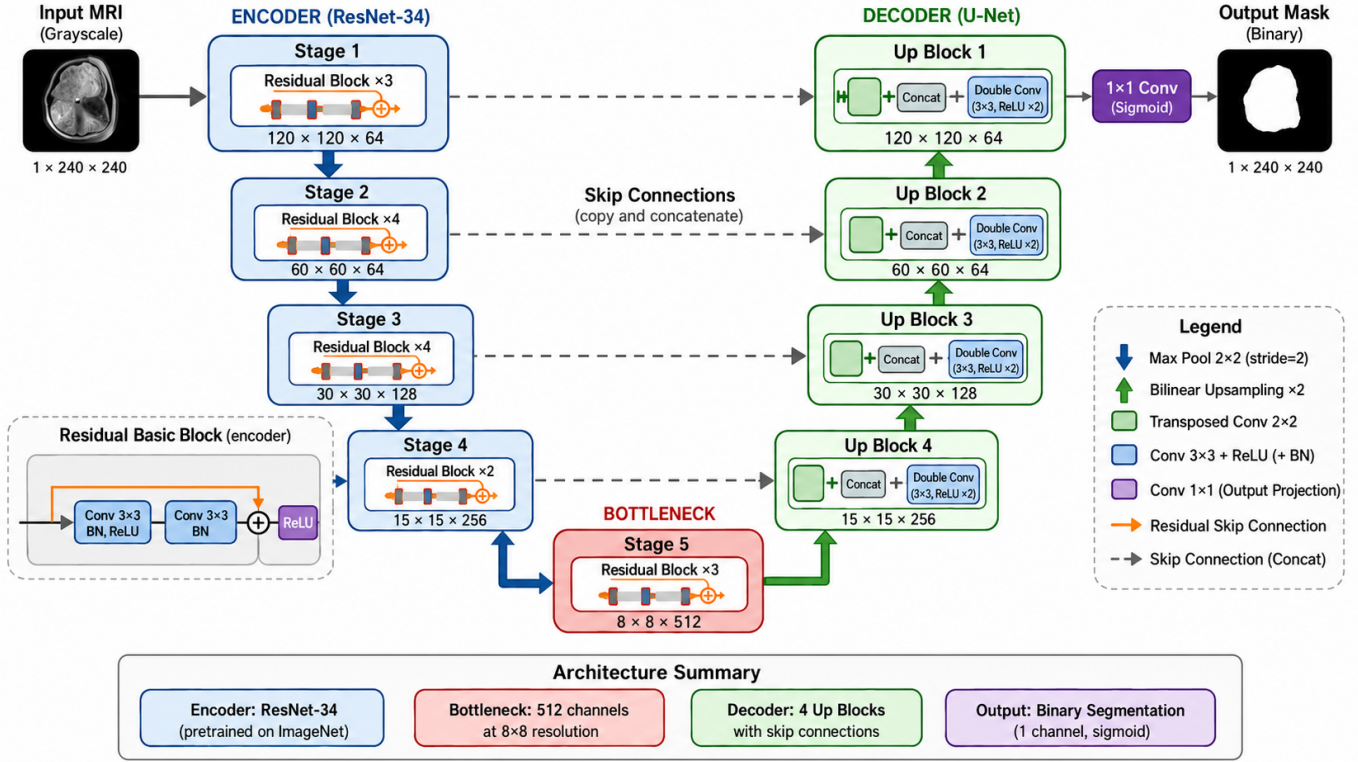


Fig. 2: Overview of the implemented 2D ResUNet architecture.

followed by T2W, whereas T1-based modalities show lower robustness (e.g., ≈ 49.7 Dice).

TABLE I: Segmentation performance (%) across evaluation protocols and MRI modalities.

Scenario	T1C		T1N		T2W		T2F	
	Dice	IoU	Dice	IoU	Dice	IoU	Dice	IoU
Baseline 2020	61.3	51.2	62.9	52.9	73.3	64.2	77.3	68.4
Baseline 2024	66.1	54.7	67.3	56.0	74.2	63.5	78.3	68.6
Cross-Dataset 20/24	49.7	38.8	49.7	39.1	62.3	51.4	58.2	46.9
Cross-Dataset 24/20	65.3	55.5	58.3	48.8	70.4	61.2	76.2	67.6
Multi-Domain 20/24	63.5	52.3	63.9	52.9	72.1	61.7	75.0	65.2

From a representation learning perspective, T2-based modalities enable more transferable features, while T1-based inputs induce more domain-dependent representations. The smaller drop (e.g., T2F: $78.3 \rightarrow 76.2$) suggests that training on BraTS 2024 yields more generalizable representations, driven by its larger scale and greater data diversity. Multi-domain training further mitigates degradation (e.g., T2F: 75.0 vs. $77.3/78.3$), indicating that exposure to heterogeneous data promotes domain-invariant features. Overall, modality choice plays a central role in model generalization, with FLAIR providing the most robust representations under domain shift (RQ2).

B. Domain Adaptation (RQ3)

This scenario evaluates domain adaptation under limited labeled target-domain data by fine-tuning models, initially trained on a source dataset, using incremental fractions (10% steps) of annotated samples from the best-performing modality (T2F/FLAIR). Results show that fine-tuning with only 10–20% of target-domain data already approaches the upper-bound performance, after which a plateau is observed, indicating diminishing returns. This directly answers RQ3, demonstrating that a small amount of target data is sufficient to substantially improve generalization while reducing annotation and computational costs. Notably, the largest gains are observed when adapting models trained on BraTS 2020 to BraTS 2024 (see Fig. 3), whereas the reverse direction (see Fig. 4) yields more modest improvements ($\approx 5\%$ Dice), consistent with the generalization analysis in Section IV-A, which suggests that models trained on BraTS 2024 already learn more transferable representations.

C. Qualitative Analysis

Table II shows segmentation performance stratified by tumor size quartile. Tumor size is defined as the number of foreground pixels in the ground-truth mask of each 2D slice. The slices are then divided into four equally populated quartiles, ranging from the smallest tumors (Q1) to the largest (Q4). Dice and IoU increase consistently from Q1 to Q4. This trend reflects both the greater sensitivity of overlap-based

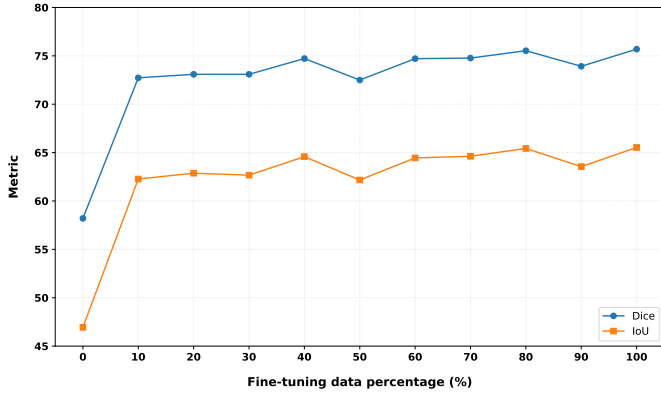


Fig. 3: Domain adaptation performance of a model trained on BraTS 2020 and progressively fine-tuned with increasing fractions of BraTS 2024 data.

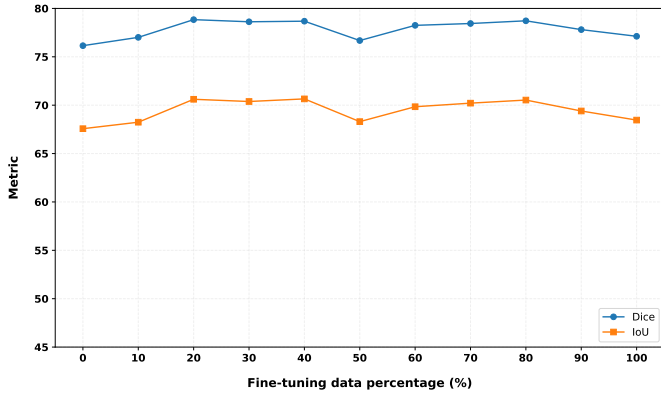


Fig. 4: Domain adaptation performance of a model trained on BraTS 2024 and progressively fine-tuned with increasing fractions of BraTS 2020 data.

metrics to segmentation errors in small structures and the richer morphological information available for larger tumors.

TABLE II: Segmentation performance (%) stratified by tumor size quartile. Values in parentheses indicate the mean tumor area within each quartile.

Quartile / Size	T1C		T1N		T2W		T2F	
	Dice	IoU	Dice	IoU	Dice	IoU	Dice	IoU
Q1 (171 px)	26.9	19.6	26.0	18.8	39.5	30.3	49.0	39.2
Q2 (729 px)	63.3	51.1	64.0	51.4	75.3	63.8	77.6	66.8
Q3 (1475 px)	75.8	64.5	76.6	65.0	84.0	74.1	84.0	75.0
Q4 (2957 px)	79.7	68.6	81.3	70.1	85.7	76.4	84.8	76.0

Modality performance also varies with size. For small tumors (Q1), T2f/FLAIR provides the best results, indicating higher sensitivity to subtle boundaries. For medium sizes (Q2–Q3), T2f/FLAIR and T2w achieve comparable performance, both outperforming T1-based modalities. For large tumors (Q4), T2w slightly surpasses T2f/FLAIR. Overall, T2f/FLAIR is preferable for small lesions, while T2f/FLAIR

and T2w offer similar performance for larger volumes, with a slight advantage for T2w in the upper quartile.

To complement Table II and better illustrate performance across tumor sizes, Fig. 5 presents segmentation masks for the case closest to the mean tumor size of each quartile.

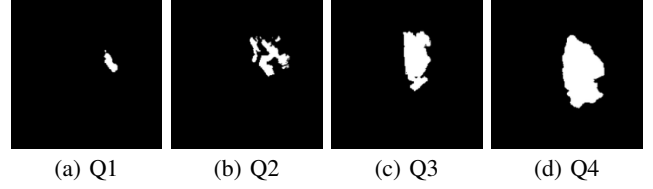


Fig. 5: Representative masks for each tumor-size quartile.

To contextualize the quantitative results, we present qualitative examples on the T2f/FLAIR sequence, including correct, intermediate, and failure cases (see Fig. 6). Each row shows the input image, the ground-truth overlay in green, and the predicted overlay in red.

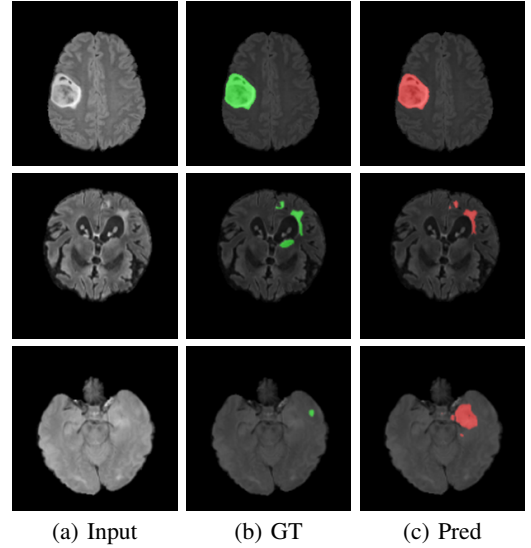


Fig. 6: Segmentation examples: high-accuracy case (top, Dice: 99.1%), intermediate case (middle, Dice: 70.1%), and failure case (bottom, Dice: 0.0%).

V. CONCLUSIONS

This work investigated the role of MRI modalities in cross-dataset generalization through a standardized ResUNet-based pipeline and controlled evaluation protocols. The results support four main conclusions. First, T2f/FLAIR consistently provides the best performance and strongest generalization, followed closely by T2w, while T1-based modalities exhibit lower robustness under domain shift (RQ1/RQ2). Second, multi-domain training reduces the generalization gap, improving transfer performance and bringing results closer to in-domain conditions. Third, tumor size significantly affects performance: T2f/FLAIR is more effective for small lesions, whereas T2w becomes comparable or slightly superior for

larger tumors. Fourth, domain adaptation via fine-tuning is highly data-efficient, with 10–20% of target-domain data already approaching upper-bound performance and limited gains thereafter (RQ3).

Overall, these findings demonstrate that model generalization is strongly influenced by modality choice, data diversity, and lesion characteristics, providing practical guidance for resource-constrained scenarios and modality selection in real-world applications.

The insights gained from this work motivate several future research directions, including investigating improved embedding spaces and adaptive architectures, extending the analysis to multimodal and 3D models for enhanced volumetric consistency, validating the findings on independent clinical datasets beyond BraTS, developing more effective strategies for small lesion detection, and assessing foundation models with respect to generalization, sensitivity, and computational efficiency.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)* – Finance Code 001, the *Financiadora de Estudos e Projetos (FINEP)* for the High-Performance Computing (HPC) infrastructure provided by the *Centro Integrado de Soluções em Inteligência Artificial (CISIA)* at the *Pontifícia Universidade Católica do Paraná (PUCPR)*, and the *Fundação Araucária*, in partnership with the *Secretaria da Ciência, Tecnologia e Ensino Superior do Estado do Paraná (SETI-PR)*, under Grant No. 653/2025 (FA/UNIVERSAL).

REFERENCES

- [1] M. Ali, S. O. Gilani, A. Waris, K. Zafar, and M. Jamil, “Brain tumour image segmentation using deep networks,” *IEEE Access*, vol. 8, pp. 153 589–153 598, 2020.
- [2] S. Solanki, U. P. Singh, S. S. Chouhan, and S. Jain, “Brain tumor detection and classification using intelligence techniques: An overview,” *IEEE Access*, vol. 11, pp. 12 870–12 886, 2023.
- [3] T. A. Soomro, L. Zheng, A. J. Afifi, A. Ali, S. Soomro, M. Yin, and J. Gao, “Image segmentation for mr brain tumor detection using machine learning: A review,” *IEEE Reviews in Biomedical Engineering*, vol. 16, pp. 70–90, 2023.
- [4] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241.
- [5] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R. T. Shinohara, C. Berger, S. H. Ha, M. Rozycki *et al.*, “Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge,” *arXiv preprint arXiv:1811.02629*, 2018.
- [6] M. C. de Verdier, R. Saluja, L. Gagnon, D. LaBella, U. Baid, N. H. Tahon, M. Foltyn-Dumitru, J. Zhang, M. Alafif, S. Baig *et al.*, “The 2024 brain tumor segmentation (brats) challenge: Glioma segmentation on post-treatment mri,” *arXiv preprint arXiv:2405.18368*, 2024.
- [7] L. Zhang, X. Wang, D. Yang, T. Sanford, S. Harmon, B. Turkbey, B. J. Wood, H. Roth, A. Myronenko, D. Xu, and Z. Xu, “Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 7, pp. 2531–2540, 2020.
- [8] S. Yin, J. Chang, Y. Wang, J. Zhu, S. Wang, and W. Lu, “Tumor segmentation in brain mr images using res-unet with multimodal fusion,” in *2025 Asia-Europe Conference on Cybersecurity, Internet of Things and Soft Computing (CITSC)*, 2025, pp. 28–32.
- [9] Q. Xing, Z. Li, Y. Jing, and X. Chen, “A 3d dual encoder mirror difference resu-net for multimodal brain tumor segmentation,” *IEEE Access*, vol. 13, pp. 1621–1635, 2025.
- [10] S. Mangayarkarasi, M. M. M. R. Asha, C. Kishore, and A. V., “Brain tumor segmentation using res-unet,” in *International Conference on Smart Electronics and Communication Systems (ISENSE)*, 2024.
- [11] T. Yang and J. Song, “An automatic brain tumor image segmentation method based on the u-net,” in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, 2018, pp. 1600–1604.
- [12] A. Goswami and M. Dixit, “An analysis of image segmentation methods for brain tumour detection on mri images,” in *2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT)*, 2020, pp. 318–322.
- [13] L. Huang, A. Miron, K. Hone, and Y. Li, “Segmenting medical images: From unet to res-unet and nnunet,” in *IEEE International Symposium on Computer-Based Medical Systems (CBMS)*, 2024, pp. 483–489.
- [14] S. Metlek and H. Çetiner, “Resunet+: A new convolutional and attention block-based approach for brain tumor segmentation,” *IEEE Access*, vol. 11, pp. 69 884–69 902, 2023.
- [15] B. A. Krinski, D. V. Ruiz, R. Laroca, and E. Todt, “DACov: A deeper analysis of data augmentation on the computed tomography segmentation problem,” *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 11, pp. 1724–1741, 2023.
- [16] S. Rajendran, S. K. Rajagopal, T. Thanarajan, K. Shankar, S. Kumar, N. M. Alsubaie, M. K. Ishak, and S. M. Mostafa, “Automated segmentation of brain tumor mri images using deep learning,” *IEEE Access*, vol. 11, pp. 64 758–64 768, 2023.