3D MRI SEGMENTATION USING RESU-NET AND U-NETR ARCHITECTURE FOR THE DETECTION OF PEDIATRICS BRAIN TUMOR

Yash Vora, Phuong Nguyen

University of Iowa

ABSTRACT

Pediatric brain tumor segmentation in 3D magnetic resonance imaging (MRI) is critical for accurate diagnosis and treatment planning. This study explores the performance of two advanced architectures, ResU-NET and UNETR, for segmenting tumor regions across multi-modal MRI data. The dataset, sourced from the BraTS Pediatric Tumor Segmentation Challenge, includes modalities such as T1c and T2-FLAIR, along with labels for enhancing tumor, non-enhancing tumor, cystic components, and edema. ResU-NET employs deep residual connections to efficiently capture spatial details, while UNETR integrates Vision Transformers for enhanced contextual understanding. Data preprocessing and augmentation, including elastic deformations and affine transformations, were employed to increase training diversity. Evaluation results demonstrate robust segmentation performance for dominant classes but highlight challenges with subtle and overlapping features like edema and cystic components. The findings underscore the need for further advancements in architecture and training strategies to improve segmentation of underrepresented classes. Our study contributes to developing efficient 3D segmentation pipelines, paving the way for better pediatric brain tumor management.

Index Terms — ResU-NET, UNETR, image segmentation, pediatrics brain tumor

1. INTRODUCTION

Magnetic Resonance Imaging (MRI) plays an important role in early diagnosis and treatment of pediatric brain tumors, offering detailed visualization of tumor regions. Interpretation of those MRI images requires segmentation process where regions of image were identified. However, segmenting tumor structures from 3D MRI data is highly challenging due to noise, subtle intensity variations [1], and overlapping features [2]. Recently, Convolutional neural networks (CNNs) have been applied and show the improvement on image segmentation especially UNET [3]. UNET, which is based on an encoder-decoder architecture, uses skip connections to combine the high-level semantic feature maps from the decoder and corresponding low-level

detailed feature maps from the encoder. Other efforts to modify the traditional UNET model by incorporating other architectures such as transformer [4], residual blocks [5], or recurrent feedback loops [6] to address problems of UNET as well as deep learning model such as vanishing gradient during training, handling imbalance multiple classes, accuracy improvement by focusing on important features.

Pediatric brain tumors, although less common than those in adults, are the leading cause of cancer-related deaths in children. Aggressive tumors like astrocytoma, diffuse midline glioma (DMG), and diffuse intrinsic pontine glioma (DIPG) pose significant challenges due to their poor prognosis. DMGs, which grow in the central nervous system (CNS) and include glial cells like astrocytes, have a median survival time of only 9–12 months [7]. These tumors are particularly resistant to chemotherapy, partly due to the blood-brain barrier. Similarly, DIPG, which originates in the brainstem's pons, has no known cure, and most affected children survive only about six months without treatment [8]. Accurate brain tumor segmentation from magnetic resonance (MR) images is crucial for effective diagnosis and treatment planning. However, manual segmentation is time- consuming and inconsistent, and existing automated models, often developed for adult cases, do not adapt well to the unique characteristics of pediatric brain tumors, such as their smaller size and differing growth patterns [9].

In this study, we explore two state-of-the-art architectures for pediatric brain tumor segmentation: ResU-NET [10] and UNETR [11]. ResU-NET, an extension of the traditional UNET, leverages residual connections to improve feature extraction and gradient flow, effectively capturing fine-grained spatial details [10]. On the other hand, UNETR introduces Vision Transformers (ViT) into the segmentation pipeline, utilizing a transformer-based encoder to process 3D patches of input images. This allows UNETR to model long-range dependencies and global contextual information, addressing key limitations of convolutional networks [11].

This study aims to evaluate and compare the segmentation performance of ResU-NET and UNETR, highlighting their respective strengths and limitations. We hypothesize that UNETR, with its transformer-based architecture, will

demonstrate superior segmentation accuracy for underrepresented classes compared to ResU-NET. By analyzing their ability to segment underrepresented classes and handle overlapping features, we propose future directions for improving segmentation accuracy and clinical applicability. Our findings contribute to the development of efficient 3D segmentation pipelines for pediatric brain tumor management.

3. DATASET

The dataset, derived from the BraTS Dataset (Brain Tumor Segmentation dataset) , is a widely used benchmark in the field of medical imaging for the segmentation of brain tumors from MRI scans. It is part of the BraTS Challenge, which aims to advance the development of algorithms for accurate brain tumor segmentation. It includes data from multiple institutions and is part of the BraTS Challenge, which has been held annually to improve segmentation algorithms [12]. This project is a part of Pediatric Tumor Segmentation Challenge (https://www.synapse.org/Synapse:syn58894466), consists of multi-modal MRI scans (T1, T1c, T2, and FLAIR) with annotations for enhancing tumor (ET), non-enhancing tumor (NET), cystic components (CC), edema (ED), and background [13].

Due to resource constraints, we focused our analysis on T1c and FLAIR, which are highly relevant for detecting enhancing tumors and edema. The dataset includes 260 3D images with a resolution of 240×240×155, split into 160 training, 40 validation, and 60 test samples. To increase the effective dataset size, data augmentation techniques such as elastic deformations and affine transformations were applied.

To manage computational resources and emphasize local features, images were cropped into patches of 128×128×128 for training. Both ResU-NET and UNETR were trained using the Adam optimizer with default hyperparameters. Dice score, Focal Tversky Loss, and Cross-Entropy Loss were employed as evaluation metrics to assess model performance across classes. The label distribution observed in the training dataset is as follows:

- **Label 0 (Background):** 1,773,727,227
- **Label 1 (Enhancing Tumor - ET):** 1,469,057
- **Label 2 (Non-Enhancing Tumor - NET):** 8,336,922
- **Label 3 (Cystic Components - CC):** 642,449
- **Label 4 (Edema - ED):** 1,424,345

4. METHODOLOGY

This study investigates the implementation and evaluation of ResU-NET and UNETR architectures for the segmentation of pediatric brain tumors using multi-modal MRI data. The methodology is designed to address the challenges of 3D segmentation, focusing on the exclusive use of T1c and FLAIR modalities for their complementary representation of tumor structures.

4.1. Data Preprocessing

The dataset was obtained from the BraTS Pediatric Tumor Segmentation Challenge 2024. Each MRI volume was padded from its original dimensions of $240{\times}240{\times}155$ to $260{\times}260{\times}160$ using reflect padding to ensure consistent input dimensions. Intensity normalization was applied to standardize the inputs across all samples, achieving zero mean and unit standard deviation for each modality. The process of data preprocessing was represented in Fig.1 below:
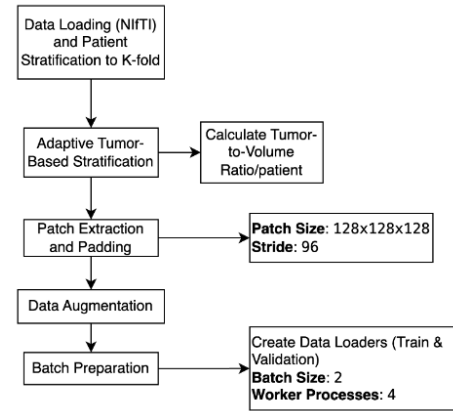

Fig. 1: Data Preprocessing

Patch extraction was performed to divide the MRI volumes into cubic patches of $128{\times}128{\times}128$. Tumor-focused sampling was employed using an adaptive binning approach based on the data distribution of tumor ratios. Quantile-based binning was applied to ensure a balanced representation of tumor structures, with the number of bins dynamically adjusted to account for sample size. Tumor-focused sampling was prioritized with a probability of 0.7, targeting regions containing higher tumor ratios. Additionally, random cropping of non-tumor regions ensured adequate representation of healthy tissue, enhancing model generalization. A stratified sampling strategy, informed by tumor ratio bins, was used to create balanced folds, maintaining diversity in training and testing splits.

4.2. Model Architecture

Two deep learning architectures were implemented to address the segmentation task:

- **ResU-NET**: An advanced version of the UNET architecture, incorporating residual connections to improve gradient flow and enhance feature extraction. Its encoder-decoder structure, paired with skip connections, preserves spatial details during the reconstruction phase. Fig 2 illustrates the architecture of ResU-Net, as described in [10].
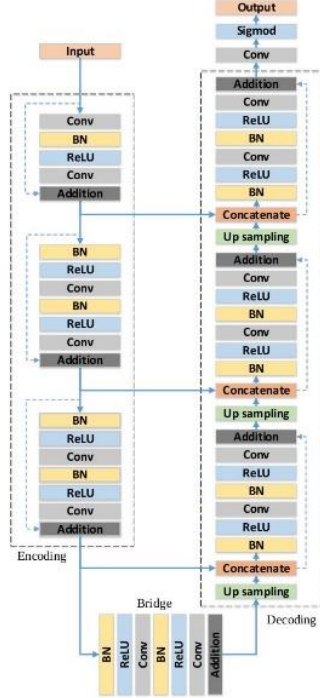

Fig. 2: The architecture of the proposed deep ResU-NET [10].

- **UNETR**: A Vision Transformer-based model that employs a transformer encoder to capture long-range dependencies in 3D data. By processing 3D patches, the encoder models global context using self-attention mechanisms. The CNN-based decoder reconstructs the segmentation map while leveraging skip connections to ensure spatial precision. Fig 3 represents the UNETR architecture that is based on the paper of [11]
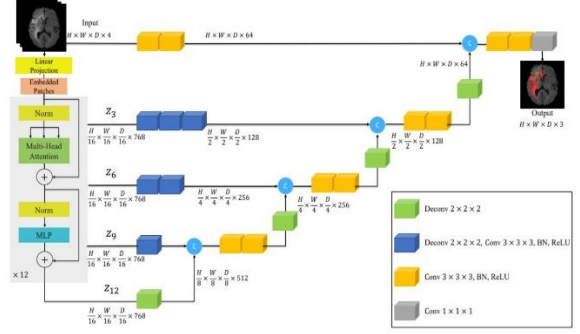

Fig. 3: UNETR Architecture [11]

4.3. Data Augmentation

To enhance model robustness and reduce overfitting, extensive data augmentation techniques were applied during training. Spatial augmentations included random flips and rotations across all three axes. Affine transformations introduced minor scaling (*±1%*) and translations (*±5* pixels in the x and y directions, and *±2* pixels along z). Elastic deformations, with an alpha of 20 and sigma of 4, introduced smooth spatial variability. Intensity augmentations, such as Gaussian noise and scaling, simulated clinical imaging variations.

4.4. Training Protocol

A patch-based training approach was employed, where *128×128×128* patches served as input to the models. The training process utilized a combined loss function comprising Focal Tversky Loss and Cross-Entropy Loss to optimize segmentation performance.

The Tversky Loss [14] is an extension of the Dice Loss, specifically designed to address class imbalance by weighing false positives (FP) and false negatives (FN). The Tversky index is defined as:

$$\text{Tversky Index} = \frac{|P \cap G|}{|P \cap G| + \alpha \cdot |FP| + \beta \cdot |FN|}$$

where P is the predicted mask, G is the ground truth mask, and α, $\beta \in [0,1]$ are parameters that control the trade-off between false positives and false negatives. In this study, α was set to 0.3 and β to 0.7 to reflect the higher importance of reducing false negatives compared to false positives in the context of medical imaging.

This choice prioritizes sensitivity over specificity, as false negatives (missed tumor regions) are clinically more critical

than false positives (regions incorrectly identified as tumors). Assigning a larger weight to β ensures the model penalizes false negatives more heavily, encouraging it to capture all tumor regions, even at the expense of slightly higher false positives. This trade-off aligns with the clinical need for comprehensive tumor detection in segmentation tasks.

The Tversky Loss is computed as:

$$\text{Tversky Loss} = 1 - \text{Tversky Index}$$

To further enhance the model's ability to focus on hard-to-classify regions, the Focal Tversky Loss [15] modifies the Tversky Loss by adding an exponent γ which emphasizes challenging examples. The Focal Tversky Loss is defined as:

$$\text{Focal Tversky Loss} = (1 - \text{Tversky Index})^{\gamma}$$

In this study, γ was set to 2, intensifying the penalization of misclassified regions. A higher value of γ increases the focus on harder examples, such as smaller or subtle tumor regions, which are often underrepresented in segmentation tasks. This adjustment ensures that the model is particularly sensitive to regions where predictions are more uncertain, improving overall segmentation performance for difficult classes.

4.5. Inference and post-processing

During inference, a sliding window approach with overlapping patches ($128 \times 128 \times 128$, stride $64 \times 64 \times 64$) was used to generate predictions across the entire MRI volume. Post-processing involved filtering out small, segmented regions (< 100 voxels) to remove noise and applying morphological operations, such as dilation and erosion, to refine segmentation boundaries.

5. EXPERIMENTAL SETUP

The experiments were conducted using the PyTorch and TorchIO libraries on an NVIDIA A30 GPU provided by the University of Iowa's IDAS platform. The following settings were employed for training and evaluation:

5.1. Hardware and Software

- **Hardware**: NVIDIA A30 GPU with sufficient memory for 3D medical image processing.
- **Software**: PyTorch (https://pytorch.org/) for deep learning, TorchIO (https://torchio.readthedocs.io/) for

data augmentation, and MONAI (https://monai.io/) for preprocessing and medical imaging utilities.

5.2. Training Configuration

- **Input patch size**: *$128 \times 128 \times 128$*.
- **Batch size**: 2 (limited by GPU memory).
- **Optimizer**: Adam optimizer with a learning rate of *1×10^{-4}* and gradient clipping.
- **Learning rate scheduler**: Cosine annealing with a 2% warm-up phase.
- **Loss function**: Combined Focal Tversky Loss and Cross-Entropy Loss.
- **Cross-validation**: Stratified 5-fold cross-validation to balance tumor and non-tumor regions.

5.3 Evaluation Metrics

The performance of ResU-NET and UNETR was assessed using two key metrics, chosen to provide complementary insights into segmentation accuracy:

5.4.1. Dice coefficient (Dice score)

The Dice Coefficient evaluates the overlap between predicted and ground truth segmentation masks. It is particularly effective for measuring segmentation accuracy in medical imaging, where class imbalance is common. The metric is defined as:

$$\text{Dice Coefficient} = \frac{2 \cdot |P \cap G|}{|P| + |G|}$$

where P and G denote the predicted and ground truth masks, respectively.

5.4.2. Intersection over Union (IoU)

IoU quantifies the ratio of the intersection to the union of the predicted and ground truth regions. It complements the Dice Coefficient by penalizing false positives and false negatives more heavily, offering a stricter measure of segmentation accuracy. It is defined as:

$$\text{IoU} = \frac{|P \cap G|}{|P \cup G|}$$

These metrics provide a comprehensive evaluation of model performance, highlighting their strengths and limitations in handling the complex task of pediatric brain tumor segmentation.

6. RESULTS

The performance of the different methods is summarized in Table 1, which includes metrics such as cross-entropy loss, Dice score, and Focal Tversky loss for both the training and validation datasets. Example segmentation results for these methods are presented in Figure 4. The cross-entropy loss for both ResU-NET and UNETR is less than 0.2, indicating that both models achieve reasonably good performance. The Dice scores for the two models range from 0.442 to 0.512, highlighting room for improvement in segmentation quality. The Focal Tversky loss ranges from 0.409 to 0.571, reflecting the models' ability to address class imbalance to some extent.

UNETR outperforms ResU-NET in Dice score on both the training and validation sets, indicating better segmentation quality and closer alignment with ground truth. Moreover, UNETR shows lower standard deviation in Dice scores, demonstrating more stable performance across samples compared to ResU-NET. Interestingly, ResU-NET achieves better performance than UNETR in Focal Tversky loss across both datasets, suggesting it may handle challenging segmentation tasks or regions with imbalanced classes more effectively. However, UNETR demonstrates lower cross-entropy loss than ResU-NET, indicating that it minimizes overall classification error more efficiently. Our random prediction results in Fig 5 (a) of ResU-NET and Fig 5 (b) for UNETR showing our models predict fairly well that the predictions are closely match the ground truths.

Table 1: Comparison of the performance between ResU-NET and UNETR

Models	Metrics	Training	Validation
ResU-NET	Cross Entropy Loss	0.149 ± 0.022	0.145 ± 0.028
	Focal Tversky Loss	0.548 ± 0.011	0.571 ± 0.016
	Dice score	0.452 ± 0.009	0.462 ± 0.019
UNETR	Cross Entropy Loss	0.098 ± 0.008	0.114 ± 0.035
	Focal Tversky Loss	0.409 ± 0.009	0.420 ± 0.019
	Dice score	0.512 ± 0.004	0.504 ± 0.008

We evaluated our models on the test sets, addressing the impact of class imbalance by calculating the Intersection over Union (IoU) and accuracy for each class, and constructing precision-recall curves as shown in Fig 6 and Fig 7. In Fig 6, Class 0, representing the background, achieves the highest IoU (nearly 1) for both ResU-NET (a) and UNETR (b). This result is expected, as the background dominates the dataset and is relatively straightforward to predict.

For Class 1 (Enhancing Tumor, ET) and Class 2 (Non-Enhancing Tumor, NET), the IoU scores are significantly lower. ResU-NET slightly outperforms UNETR for ET, while UNETR shows a marginal advantage for NET, suggesting subtle differences in their ability to segment these regions. Classes 3 (Cystic Core, CC) and 4 (Edema, ED) exhibits extremely low IoU values for both models, reflecting the significant challenges associated with detecting rare and subtle structures like CC and ED.

In terms of accuracy in Fig 7, background again exhibits near-perfect scores for both models, consistent with its dominant presence in the dataset. NET achieves relatively high accuracy, indicating that it is easier to segment compared to other tumor subtypes. However, ET shows notably lower accuracy, highlighting the models' difficulty in segmenting this region. CC and ED continue to display minimal accuracy scores, underscoring the difficulty posed by their infrequency and variability in the dataset.

These results reveal the strengths and limitations of both models in handling class imbalance, with robust performance for dominant classes but significant challenges in accurately segmenting rare and complex tumor subregions.
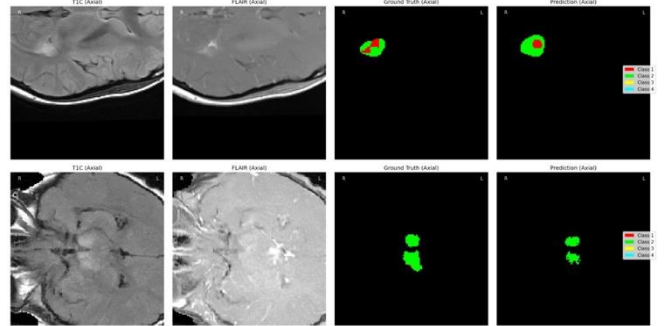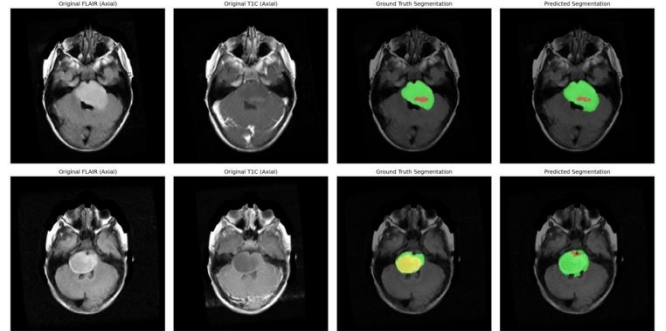

Fig. 5.a Prediction results using ResU-NET

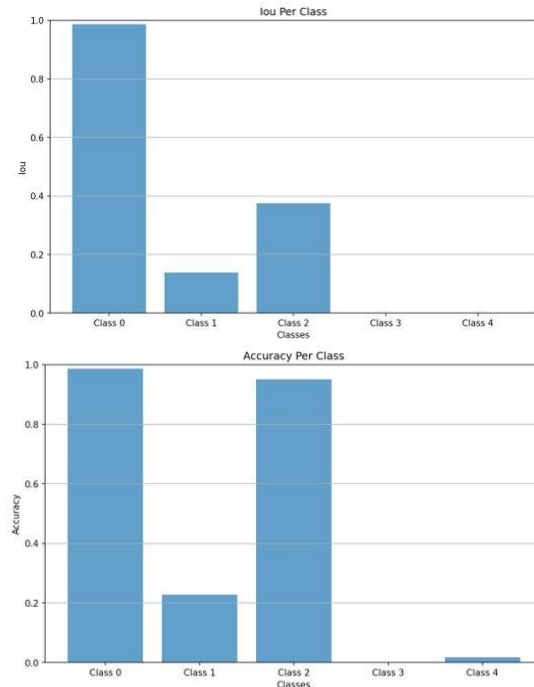**Fig. 5.b** Prediction results using UNETR

Fig. 6.a Bar plot of Iou and Accuracy of ResU-NET
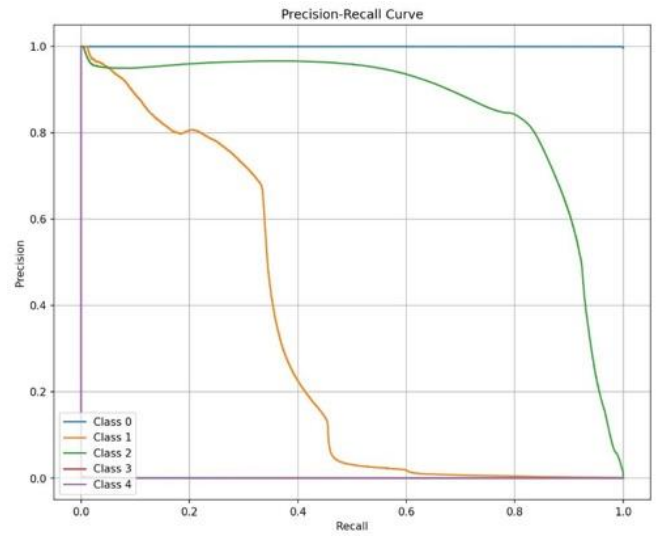

Fig. 6.b Bar plot of Iou and Accuracy of UNETR

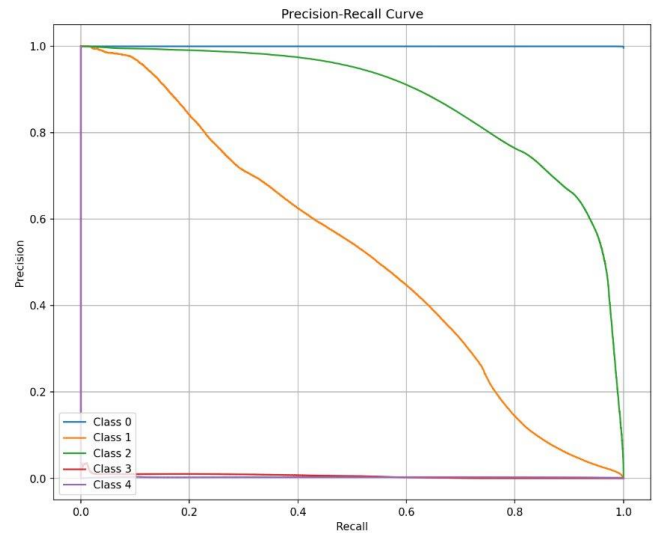**Fig. 7.a** Precision and recall curve of ResU-NET

Fig. 7.b Precision and recall curve of UNETR

For precision and recall curve, the background curve remains in the top-right corner, indicating near-perfect precision and recall for both ResU-NET and UNETR. This is consistent with its dominant presence in the dataset. For ET, both models demonstrate a steep drop in the curve, indicating poor balance between precision and recall. ResU-NET slightly outperforms UNETR for this class, but both struggle with false positives when recall increases. The NET's curve achieves a better balance between precision and recall, with UNETR slightly outperforming ResU-NET in terms of stability and curve shape.

Both models perform poorly on CC and ED, with near-zero precision and recall. These results underscore the difficulty in identifying rare tumor subtypes, likely due to their infrequent appearance and lack of clear distinguishing features.

7. DISCUSSION

7.1 Models results discussion

The results indicate that both ResU-NET and UNETR face significant challenges in segmenting smaller, less frequent classes like CC and ED (Fig 5.b). While background and NET are handled well by both models, the performance disparity for other classes highlights several key issues. First is the class imbalance, the dataset's imbalance, with background dominating, heavily influences the models' ability to generalize across all classes. This is evident in the IoU and accuracy metrics, as well as the precision-recall curves for CC and ED. In addition, ResU-NET demonstrates better IoU for smaller classes like ET but struggles with precision-recall balance for NET. UNETR exhibits more consistent performance across IoU, accuracy, and precision-recall metrics, particularly for NET. However, it does not significantly outperform ResU-NET for challenging classes.

The poor performance on rare class (CC and ED) for both models suggest the need for additional strategies, such as class-balancing techniques, weighted loss functions, or data augmentation targeting rare structures. The observed performance aligns with the earlier analysis of loss functions, where ResU-NET performed better in Focal Tversky Loss, likely aiding its IoU for smaller classes. However, UNETR's overall better Dice score may explain its balanced precision-recall curves for NET.

7.2 Comparison with other state-of-art models

When evaluating the performance of ResU-NET and UNETR in the context of other state-of-the-art models, such as MedNeXt [16] and nnU-Net [17], significant variations in segmentation quality across different model architectures emerge. ResU-NET achieves relatively low Dice scores across all tumor subregions, with negligible segmentation capabilities for CC (0.00) and ED (0.0001). These results highlight the model's limited ability to capture complex tumor features, likely due to its simpler architecture and absence of advanced feature extraction mechanisms.

UNETR, while outperforming ResU-NET for ET and NET segmentation with Dice scores of 0.28 and 0.716, respectively, similarly struggles with CC and ED (0.00). This suggests that while UNETR is better at capturing detailed structural features compared to ResU-NET, it lacks the robustness required for rare or challenging tumor regions.

Notably, nnU-Net [17] significantly outperforms both ResU-NET and UNETR across all tumor subregions, achieving Dice scores exceeding 0.8 for NET, CC, and ED. nnU-Net adopts a hierarchical segmentation approach where the model predicts WT and then identifies ET, CC, and ED, while categorizing the remaining WT region as NET. This aligns with radiologists' manual segmentation practices, focusing on discernible traits for certain regions and inferring the remainder as NET. This approach effectively mitigates challenges related to class imbalance, particularly for underrepresented subregions.

On the other hand, MedNeXt [16] does not directly classify all six subregions. Instead, it predicts only ET, TC, and WT, leaving the remaining subregions—NET, CC, and ED—to be inferred during testing by the BraTS organizers. This division of labor reflects this year's BraTS-PED task design, where models are evaluated on their ability to predict the primary regions, and the secondary regions are determined programmatically.\MedNeXt achieves the highest Dice score for ED (0.967), reflecting its superior ability to segment diffuse tumor regions. It also performs competitively on ET (0.62) and CC (0.66), positioning it as one of the most balanced models for segmentation tasks. In contrast, models like UNETR and ResU-NET directly predicted NET alongside the other subregions, which likely contributed to their poor performance on ED and CC due to severe class imbalance and the dominance of NET. This direct prediction approach results in diluted performance for other subregions when NET occupies a disproportionate amount of the tumor region.

Given these challenges, ResU-NET demonstrated some potential for capturing fine-grained details in challenging regions such as ET (Dice score of 0.512). However, its overall segmentation quality, particularly for rare structures, falls short of advanced models like nnU-Net and MedNeXt. This analysis underscores the importance of strategic label prediction frameworks, such as hierarchical segmentation, and highlights the trade-offs between direct segmentation of all subregions and task-specific inference methods.

7.3 Limitation

The study faced several limitations that impacted segmentation performance, particularly for the ED and CC. The subtle intensity variations and overlapping patterns between tumor subregions often led to misclassification of ED and CC, highlighting the difficulty of distinguishing these regions from other tumor types. Additionally, the models demonstrated a focus imbalance, with strong predictions for NET regions reducing sensitivity to the nuanced spatial and intensity features of ED and CC. Resource constraints also pose challenges. For example, GPU constraints limited batch size to 2, reducing model generalization for rare subregions like CC and ED. This limitation likely reduced the ability of the models to generalize well for ED and CC segmentation. These results indicate the need for advanced techniques, such

as hybrid architectures, adaptive loss functions, and class-specific augmentations, to better capture the distinct spatial-intensity features of these challenging subregions. Future work should address these challenges by leveraging more computational resources and incorporating innovative approaches to improve segmentation performance for ED and CC.

8. CONCLUSION

The model demonstrates robust segmentation performance for dominant classes but struggles with subtle and overlapping features, such as edema (ED) and cystic components (CC). Addressing these limitations through improved architectures, advanced techniques, and enhanced computational resources could significantly improve segmentation performance. Our future work will focus on exploring innovative strategies to develop a more accurate and balanced model capable of effectively handling imbalanced classes.

9. ACKNOWLEDGMENTS

14. REFERENCES

[1] Yu, Ying, Chunping Wang, Qiang Fu, Renke Kou, Fuyu Huang, Boxiong Yang, Tingting Yang, and Mingliang Gao. "Techniques and challenges of image segmentation: A review." *Electronics* 12, no. 5 (2023): 1199.

[2] Haja, Asmaa, Stefania Radu, and Lambert Schomaker. "A Comparison of Different U-Net Models for Segmentation of Overlapping Organoids." In *Proceedings of the 2022 9th International Conference on Biomedical and Bioinformatics Engineering*, pp. 1-10. 2022.

[3] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pp. 234-241. Springer International Publishing, 2015.

[4] Vaswani, A.: 'Attention is all you need. ', Advances in Neural Information Processing Systems, 2017.

[5] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.

[6] Liang, Ming, and Xiaolin Hu. "Recurrent convolutional neural network for object recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3367-3375. 2015.

[7] Warren, K.E.: 'Diffuse intrinsic pontine glioma: poised for progress.', Frontiers in Oncology, 2012, 2, pp. 205

[8] Al Sharie, Sarah, Dima Abu Laban, and Maysa Al-Hussaini. "Decoding diffuse midline gliomas: a comprehensive review of pathogenesis, diagnosis and treatment." *Cancers* 15, no. 19 (2023): 4869.

[9] Kharaji, Mona, Hossein Abbasi, Yasin Orouskhani, Mostafa Shomalzadeh, Foad Kazemi, and Maysam Orouskhani. "Brain Tumor Segmentation with Advanced nnU-Net: Pediatrics and Adults Tumors." *Neuroscience Informatics* (2024): 100156.

[10] Zhang, Zhengxin, Qingjie Liu, and Yunhong Wang. "Road extraction by deep residual u-net." *IEEE Geoscience and Remote Sensing Letters* 15, no. 5 (2018): 749-753.

[11] Hatamizadeh, Ali, Yucheng Tang, Vishwesh Nath, Dong Yang, Andriy Myronenko, Bennett Landman, Holger R. Roth, and Daguang Xu. "Unetr: Transformers for 3d medical image segmentation." In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp. 574-584. 2022.

[12] Nourel, H., Tahon., Nader, Ashraf., Ahmed, W., Moawad., Anastasia, Janas., Ujjwal, Baid., Rachit, Saluja., Yury, Velichko., Divya, Ramakrishnan., Kiril, Krantchev., Jeffrey, D., Rudie., Spyridon, Bakas., Mariam, Aboian.: '1. Dsai-05 the brain tumor segmentation (brats-mets) challenge 2023: brain metastasis segmentation on pre-treatment mri.', Neuro-oncology advances, 2024

[13] Kazerooni, A.F., Khalili, N., Gandhi, D., Liu, X., Jiang, Z., Anwar, S. M., ... & Linguraru, M. G. : 'The Brain Tumor Segmentation in Pediatrics (BraTS-PEDs) Challenge: Focus on Pediatrics (CBTN-CONNECT-DIPGR-ASNR-MICCAI BraTS-PEDs)', arXiv preprint arXiv:2404.15009., 2024.

[14] Salehi, S.S.M., Erdogmus, D., & Gholipour, A.: 'Tversky loss function for image segmentation using 3D fully convolutional deep networks. ', International workshop on machine learning in medical imaging 2017, September, pp. pp. 379-387, 2017.

[15] Abraham, Nabila, and Naimul Mefraz Khan. "A novel focal tversky loss function with improved attention u-net for lesion segmentation." *2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019)*. IEEE, 2019.

[16] Hashmi, S., Lugo, J., Elsayed, A., Saggurthi, D., Elseiagy, M., Nurkamal, A., Walia, J., Maani, F. A., & Yaqub, M. : 'Optimizing brain tumor segmentation with MedNeXt: BraTS 2024 SSA and pediatrics. ', arXiv, 2024.

[17] Bengtsson, M., Keles, E., Durak, G., Anwar, S., Velichko, Y. S., Linguraru, M. G., Waanders, A. J., & Bagci, U: 'A new logic for pediatric brain tumor segmentation.', arXiv, 2024.