Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 20 No. 6 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v20i06.46979

PAPER

Advancing Brain Tumor Segmentation in MRI Scans: Hybrid Attention-Residual UNET with Transformer Blocks

Sobha Xavier P(✉),
Sathish P K, Raju G

Department of Computer Science, Christ (Deemed to be University), Bengaluru, Karnataka, India

sobha.xavier@res.christuniversity.in

ABSTRACT

Accurate segmentation of brain tumors is vital for effective treatment planning, disease diagnosis, and monitoring treatment outcomes. Post-surgical monitoring, particularly for recurring tumors, relies on MRI scans, presenting challenges in segmenting small residual tumors due to surgical artifacts. This emphasizes the need for a robust model with superior feature extraction capabilities for precise segmentation in both pre- and post-operative scenarios. The study introduces the Hybrid Attention-Residual UNET with Transformer Blocks (HART-UNet), enhancing the U-Net architecture with a spatial self-attention module, deep residual connections, and RESNET50 weights. Trained on BRATS'20 and validated on Kaggle LGG and BTC_postop datasets, HART-UNet outperforms established models (UNET, Attention UNET, UNET++, and RESNET 50), achieving Dice Coefficients of 0.96, 0.97, and 0.88, respectively. These results underscore the model's superior segmentation performance, marking a significant advancement in brain tumor analysis across pre- and post-operative MRI scans.

KEYWORDS

attention UNET, post-operative MRI, residual tumors, RESNET-50, UNET, UNET++

1 INTRODUCTION

Accurate segmentation of brain tumors is crucial for physicians to identify the tumor's location, size, and shape, guiding surgical planning and minimizing damage to healthy tissue. However, in the post-operative recurrent tumor segmentation, accuracy is challenging due to variations in image appearance, low contrast, and the intricate structure of the brain after surgery. Deep learning methods, particularly convolutional neural networks (CNNs) such as the RESNET and UNET architectures, have shown comparatively better results in addressing these challenges [1,2,3]. Research efforts have focused on improving the precision and efficiency of UNET-based models for brain tumor segmentation in MRI scans [4,5,6,7]. Some studies propose novel models, like an enhanced UNET with RESNET variants [8] as backbone, combining the strengths of UNET and RESNET for accurate segmentation. Other models, such as Attention UNET

Xavier, P.S., Sathish, P.K., Raju, G. (2024). Advancing Brain Tumor Segmentation in MRI Scans: Hybrid Attention-Residual UNET with Transformer Blocks. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(6), pp. 103–115. https://doi.org/10.3991/ijoe.v20i06.46979

Article submitted 2023-11-24. Revision uploaded 2024-01-19. Final acceptance 2024-01-19.

[5] and deep residual UNET [8], aim to enhance segmentation accuracy. While the original RESNET or UNET models may not be optimal for brain tumor segmentation in post operative or other challenging context, modifications have been suggested. For example, Zhang et al. [9] introduced a "dual-pathway" RESNET model, and Zhao et al. [10] incorporated an attention mechanism into the RESNET model to improve segmentation accuracy. Studies emphasize the significance of enhanced architectures over state-of-the-art methods for achieving further improved segmentation results, accelerating convergence, and ensuring rich hierarchical feature extraction.

This research aims to develop an efficient model capable of accurately segmenting brain tumors in both pre-operative and post-operative brain MRI. Post-operative brain MRI presents challenges such as a resection cavity, surgical artifacts, and very small tumor sizes. The availability of post-operative datasets containing recurrent tumors and their corresponding ground truth is limited, making it challenging to effectively train models with such datasets. To address this limitation, the proposed solution is to design a model with enhanced deep feature extraction capabilities. The proposed Hybrid Attention-Residual UNET with Transformer Blocks (HART-UNet) integrates a spatial attention module, deep residual connections, and RESNET50 weights at initialization. The model is designed to extract detailed spatial information from the input MRI dataset in both pre- and post-operative contexts. The attention mechanism enhances the model's feature capture, and residual connections are maintained throughout the network to facilitate information flow and mitigate the vanishing gradient problem. The proposed model is trained on the BRATS'20 [11,12,13] pre-operative dataset and validated on the Kaggle LGG preoperative dataset and BTC_postop [14] post-operative dataset, showing superior performance compared to state-of-the-art methods. The graphical abstract of the entire workflow is depicted in Figure 1.

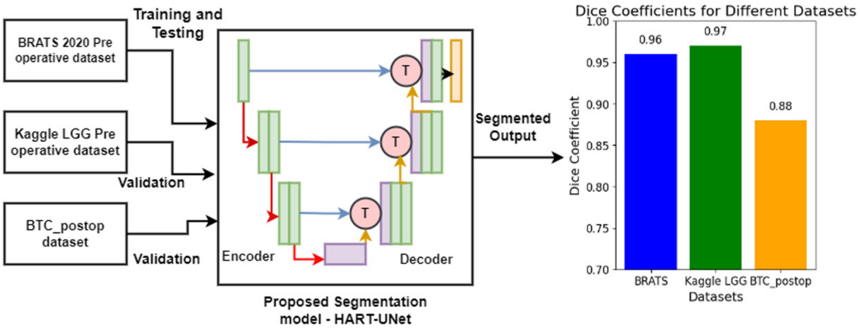

Fig. 1. Workflow overview

The following are the main contributions of this study to brain tumor analysis in pre and post-operative MRI:

1. This work introduces HART-UNet, a novel model that integrates the UNET segmentation architecture with a transformer block positioned between the encoder and decoder architectures. Additionally, it incorporates residual connections to address the vanishing gradient problem.
2. The transformer block includes a spatial self-attention module, empowering the model with the capability to selectively attend to specific spatial regions within an image.
3. The proposed model is initialized with pretrained weights from the RESNET 50 model, enhancing its feature extraction capability.
4. The performance of the proposed model is compared with existing state-of-the-art segmentation models, demonstrating impressive results for both pre-operative and post-operative datasets.

The paper is organized as follows: Section 2 presents related work, and Section 3 explains the dataset and methods. Section 4 presents the results, and Section 5 concludes the paper. The results demonstrate that the proposed HART-UNet model enhances brain MRI analysis both pre-surgery for treatment planning and post-surgery for follow-up treatment.

2 LITERATURE STUDY

Over the past few years, there has been widespread adoption of deep learning-based methods in medical image analysis. Among these methods, convolutional neural networks (CNNs) have emerged as highly successful in numerous medical image segmentation tasks [5]. CNN-based segmentation techniques can automatically learn hierarchical representations of medical images, enabling them to effectively capture essential image features [7,15]. This section discusses state-of-the-art techniques for medical image segmentation, including their applications, advantages, and disadvantages.

The U-Net architecture is a favored choice for medical image segmentation, particularly in the field of brain MRI segmentation. Olaf Ronneberger et al. introduced the U-Net, a convolutional network, in their publication titled "U-Net: Convolutional Networks for Biomedical Image Segmentation" [5]. It is renowned for its capacity to handle small and typically shaped objects, making it an excellent choice for activities involving medical imaging. It consists of a contracting path and an expansive path, allowing it to learn both low-level and high-level image features [6,15]. U-Net has fewer parameters compared to other architectures, making it more computationally efficient, crucial for real-time applications. U-Net can be fine-tuned for a particular task and has been pre-trained on a huge dataset, saving time and computational resources. However, U-Net has limitations in dealing with semantic information, leading to potential overfitting and poor segmentation performance [5].

Numerous variants and adaptations of the U-Net design have been proposed since the original publication to enhance its performance in medical imaging tasks. Alterations include dense connections, attention processes, and various normalization strategies. In "Attention U-Net: Learning Where to Look for the Pancreas" [4], Oktay et al. proposed an attention mechanism to the U-Net design. Milletari et al. proposed V-Net, an extension of U-Net, for volumetric medical image segmentation in their article titled "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation" [16,17]. In a separate publication, Zhou et al. introduced U-Net++ as another variant in their paper titled "U-Net++: A Nested U-Net Architecture for Medical Image Segmentation" [18]. U-Net++ improves upon the original U-Net design by incorporating nested U-Net blocks, facilitating more efficient feature reuse and yielding improved performance. X. Zhang et al.'s work presented RESNET50, another architecture widely applied in medical imaging [2]. By integrating residual connections, the RESNET architecture seeks to solve the issue of vanishing gradients in deep networks.

Table 1 provides a comparison of different popular deep learning models for image segmentation, including U-Net, U-Net++, Attention U-Net, RESNET-50, V-Net, DenseNet, and PSNet. These models have significantly advanced the field of image segmentation by offering distinct architectures and techniques to achieve precise and accurate delineation of objects and regions of interest in images. RESNET-50 introduced skip connections to tackle the challenges of training deep networks, while U-Net's U-shaped design enabled effective feature extraction and localization. V-Net expanded upon U-Net by extending it to 3D for volumetric segmentation, and DenseNet improved information flow and feature reuse. Notably, DeepMedic specializes in medical imaging, DenseNet emphasizes feature reuse, and PSNet incorporates attention mechanisms to enhance segmentation accuracy. Together, these models have revolutionized image segmentation, greatly improving precision and advancing the field as a whole.

Table 1. A comparative analysis of various well-known deep learning models for medical image segmentation

Model	Key Features	Pros	Cons
U-Net	Symmetric encoder-decoder architecture, skip connections	Good localization, handles small datasets well	Limited receptive field
U-Net++	Nested architecture with multiple encoder-decoder paths, dense skip connections	Improved localization, better performance on large datasets	More complex, computationally expensive
Attention U-Net	Adds an attention mechanism to the U-Net architecture	Better handling of class imbalance, more accurate segmentation	May overfit to training data
RESNET-50	Deep residual learning architecture with skip connections	Good performance on large datasets, improved handling of vanishing gradients	May require more training data, may overfit to training data
V-Net	3D extension of U-Net, uses convolutions in all dimensions	Better handling of 3D data, improved localization	Computationally expensive
DenseNet	Dense blocks with dense skip connections	Good performance on large datasets, improved feature reuse	More complex, computationally expensive
PSPNet[19]	The Pyramid Scene Parsing Network (PSPNet) leverages pyramid pooling modules to gather contextual information at multiple scales.	Incorporates multi-scale contextual information. Achieves high accuracy in segmentation tasks.	Requires more computational resources than traditional UNET architecture.

Jieneng Chen et al. [20] introduced TransUNet, a compelling fusion of Transformers and U-Net, offering a robust solution for medical image segmentation. In this model, the Transformer plays a dual role by encoding tokenized image patches from a convolutional neural network (CNN) feature map. This input sequence facilitates the extraction of global contexts. Meanwhile, the decoder part of TransUNet performs up-sampling of the encoded features. These features are subsequently merged with the high-resolution CNN feature maps, contributing to precise localization during the segmentation process.

Based on the insights gained from the above-mentioned studies, we propose HART-UNet, an improved UNET model. It includes a transformer unit composed of a spatial self-attention module and residual connections, added between down-sampling layers. The entire model incorporates pre-trained RESNET-50 initial weights. The proposed model is trained using publicly available Brats 2020 pre-operative brain MRI data, achieving a Dice coefficient value of 0.96. Subsequently, we evaluated the model on another pre-operative dataset, Kaggle LGG, where the Dice coefficient reached 0.97. For the BTC_postop dataset containing post-operative brain MRIs, the model achieved a Dice coefficient of 0.88%. These results demonstrate that our proposed model excels in terms of accuracy, Dice score, and IOU, showcasing a more precise and reliable segmentation of brain tumors.

3 MATERIALS AND METHODS

3.1 Dataset description

We trained our proposed model using the BraTS2020 pre-operative dataset, validated its performance on the Kaggle LGG preoperative dataset, and assessed its capabilities on the BTC_post-operative dataset. A comprehensive evaluation was conducted, comparing its performance with four other state-of-the-art models.

The corresponding image count and distributions for the BraTS2020, Kaggle LGG datasets, and BTC_post-operative dataset, are presented in Table 2.

Table 2. Image count and distributions for BraTS2020, Kaggle LGG, and BTC_post-operative datasets

Dataset	Total Images	Training	Testing	Validation
BRATS 2020	1700	1300	200	200
Kaggle LGG	1373	–	–	100
BTC_postop	19	–	–	19

The BraTS2020 challenge dataset comprises MRI scans from 365 patients, totaling 1,700 images diagnosed with glioma brain tumors. The dataset includes ground truth annotations for tumor segmentation. For training our model, we used 1700 images, reserving the remaining 200 for testing purposes and 200 for validation. To validate the accuracy of the proposed model, we employed the Kaggle LGG Segmentation dataset, which incorporates FLAIR abnormality masks and MR images of the brain from 1373 patients with lower-grade glioma. This dataset is sourced from The Cancer Genome Atlas (TCGA), and it includes FLAIR sequencing, as well as genomic cluster data provided by The Cancer Imaging Archive. Patient information and tumor genomic groupings are available in the data.csv file [21,22].

In addition to these datasets, we utilized the BTC_postop, a publicly available post-operative dataset comprising 7 glioma patients and 12 meningioma patients, accompanied by corresponding tumor masks. This specific dataset was employed to assess the efficiency of our model in post-operative recurrent tumor segmentation.

3.2 Proposed segmentation model – HART-UNet

This study introduces a Hybrid Attention-Residual UNET with Transformer Blocks (HART-UNet), an enhanced UNET model with a RESNET-50 backbone. The enhanced U-Net model comprises a multi-layered UNET architecture with a sequence of convolutional and max-pooling layers, specifically consisting of four sets with filter sizes of 32, 64, 128, and 256. Residual connections are incorporated within these layers to facilitate the flow of information and mitigate the vanishing gradient problem, enhancing the model's depth. The incorporation of a combined spatial and self-attention module, strategically placed between the encoder and decoder architectures, empowers the model with the capability to selectively attend to particular spatial regions within an image. The model further integrates skip connections to enable shortcut paths that preserve high-level semantic information, promoting more effective feature reuse. Following the convolutional layers, flatten layers with dimensions of 256, 512, and 256 are employed, allowing the model to capture spatial hierarchies and flatten the extracted features. Up-sampling layers with appropriate dimensions are then utilized to restore spatial resolution. The entire model is initialized with ResNet-50 weights [23,24], leveraging pre-trained features from ImageNet for improved performance. The final output layer produces the segmented output, representing the predicted regions of interest in the input images. This comprehensive architecture is designed to enhance the U-Net model's ability to accurately segment brain tumors in medical images, leveraging the power of residual connections, attention mechanisms, and pre-trained weights from ResNet-50.

The Transformer block operates on a sequence of vectors or embeddings, derived from the previous layer. In the Multi-Head Self-Attention module, linear transformations yield Query, Key and Value vectors [25]. Scaled Dot-Product Attention calculates attention scores, normalized with SoftMax to obtain attention weights. The weighted sum of Value vectors captures diverse information. Residual Connection and Layer Normalization are applied before and after the attention mechanism. The subsequent Feedforward Neural Network, with non-linear activations like ReLU, processes the attention mechanism's output for feature transformation. Another Residual Connection and Layer Normalization finalize the block, yielding a sequence of vectors post self-attention and feedforward processing. This modular design adeptly captures intricate dependencies in input sequences. Overall, the architecture synergizes the robust feature extraction capabilities of the U-Net architecture, transformer unit, and ResNet-50 backbone with the finesse to capture intricate details facilitated by skip connections. The incorporation of attention mechanisms amplifies the model's capacity to concentrate on informative regions, thereby enhancing the accuracy of segmentation results. Figure 2a represents the detailed architecture of the proposed segmentation model, and Figure 2b illustrates the detailed working of the transformer unit.

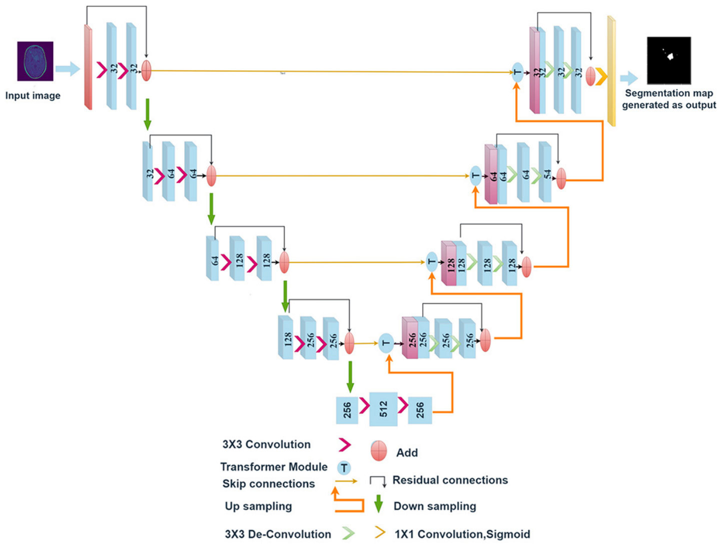

Fig. 2a. The architecture of the proposed model

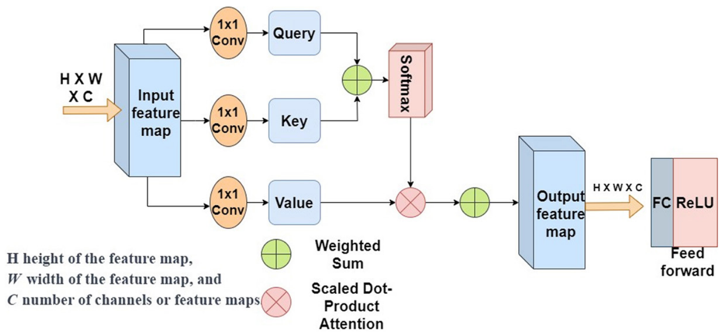

Fig. 2b. The architecture of the transformer module

4 EXPERIMENTAL SETUP AND RESULTS

The experimental setup involved using the BRATS 2020 datasets for training and testing. The dataset consisted of 1,700 samples in NIfTI format. Data augmentation technique, such as sharpening augmentation, is applied to enhance the input images.

The 'cv2.filter2D' function with a specific kernel (3 × 3) is used to apply sharpening, which helps enhance the edges and details in the images. The proposed model and other state of the art models were implemented using TensorFlow and Keras in the Colab Pro environment, utilizing a high-performance GPU for accelerated computations. Figure 3a to 3d illustrates the segmentation output of different models. Figure 3e illustrates the qualitative results displaying the segmentation outputs generated by the proposed model.

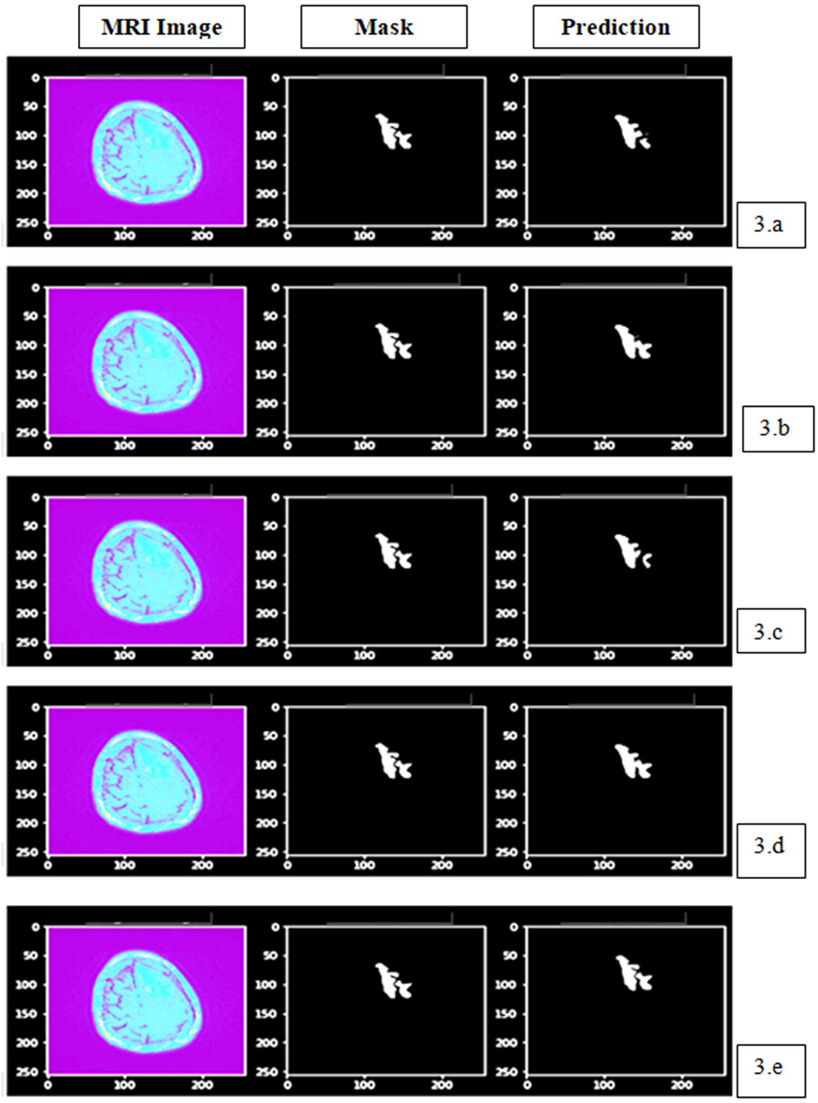

Fig. 3. Segmentation output of different models, 3a UNET model, 3b UNET++, 3c Attention Unet, 3d RESNET50 and 3e Proposed model

To assess the performance of the segmentation model, we employed the well-known performance measure, the Dice score. It assesses the similarity between the predicted segmentation (model output) and the ground truth segmentation (actual labels). The Dice score is calculated using a specific formula in equation 1.

$$\text{Dice} = \frac{2(x \cap y)}{x + y} \tag{1}$$

Where, X is the set of pixels in the predicted segmentation mask. Y is the set of pixels in the ground truth segmentation mask. $x \cap y$ represents the number of pixels that are common to both X and Y. X and Y are the total number of pixels in X and Y, respectively. In the context of segmentation models, a higher Dice score indicates better segmentation accuracy. The charts (Figure 4), shows loss, dice coefficient and IOU [26,27,28].

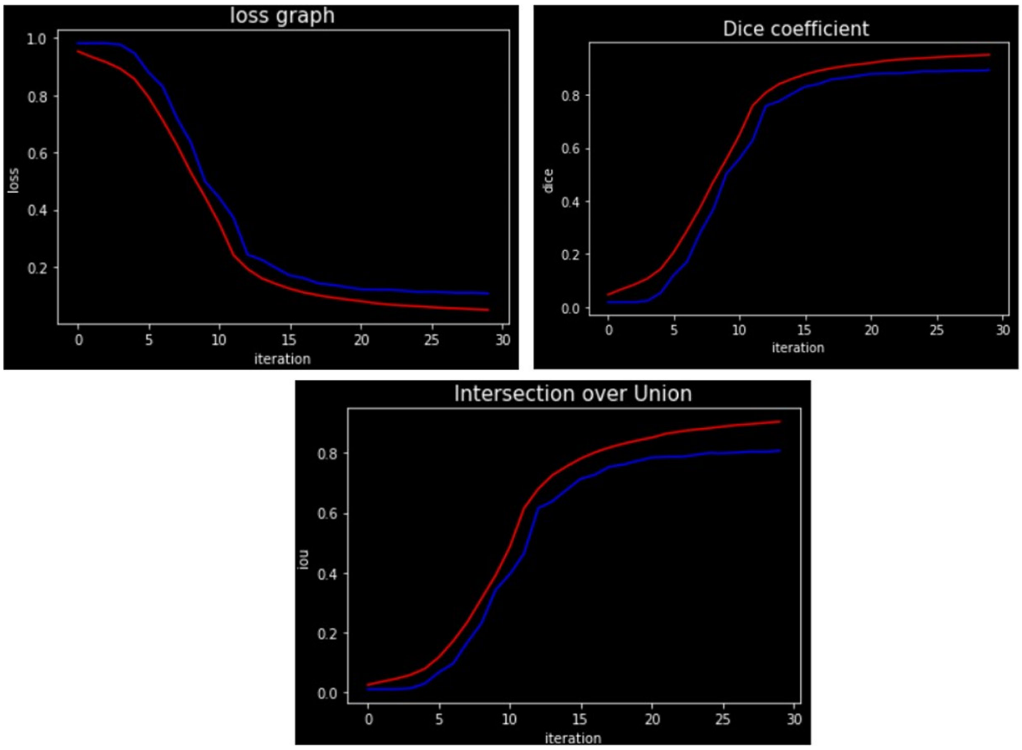

Fig. 4. Accuracy and loss graph of proposed model. (Using BRATS'20 dataset)

Upon evaluating the segmentation output of various networks, including UNET, UNET++, RESNET50, Attention UNET, and the proposed model, it can be observed that the proposed UNET model with a RESNET backbone and attention mechanism achieves the most accurate segmentation results. A comparison of the evaluation metrics for the proposed method against other state-of-the-art models is presented in Table 3. Among the evaluated models, the proposed model achieved the highest performance with a low loss of 0.03 and a high IOU of 0.92 on the BraTS'20 dataset, as well as a Dice Coefficient of 0.96 on the Kaggle LGG dataset. Table 4 and Figure 5 present the performance comparison of the proposed model with state-of-the-art models in both pre- and postoperative MRI datasets.

Table 3. Performance comparison with state-of-the-art models (using BRATS 2020 pre-operative dataset)

Model Name	Epochs	Loss	IOU	Dice Coefficient
UNET	30	0.1	0.67	0.8
UNET++	30	0.06	0.82	0.89
Attention UNET	30	0.17	0.62	0.87
RESNET50	30	0.06	0.81	0.78
Proposed Model	**30**	**0.03**	**0.92**	**0.96**

Table 4. Performance comparison with state-of-the-art models (using pre- and post-operative datasets)

Model Name	Dice Coefficient (BRATS 2020 Dataset)	Dice Coefficient (Kaggle LGG Dataset)	Dice Coefficient (BTC_Postop Dataset)
UNET	0.8	0.82	0.67
UNET++	0.89	0.84	0.72
Attention UNET	0.87	0.78	0.6
RESNET50	0.78	0.84	0.65
Proposed Model	**0.96**	**0.97**	**0.88**

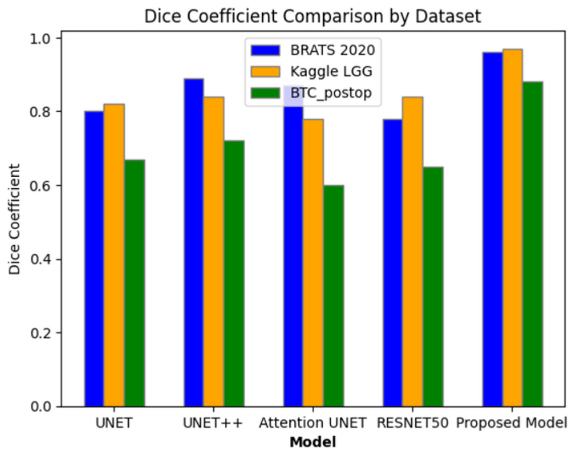

Fig. 5. Proposed model vs. state-of-the-art models in pre- and post-operative MRI datasets

5 DISCUSSION AND FUTURE DIRECTION

This study introduces HART-UNet, integrating UNET segmentation with a transformer block and residual connections for enhanced brain tumor analysis in pre- and post-operative MRI. The transformer block, featuring a spatial self-attention module, improves spatial region focus. The model, initialized with pretrained RESNET 50 weights, enhances feature extraction. Despite being trained on pre-operative data, our experiment explores the model's efficacy in post-operative scenarios, achieving a 0.97 Dice score on the publicly available BTC_postop dataset. The results demonstrate the model's ability to accurately capture intricate details of brain structures, specifically when other models struggle to precisely depict the contours of the mask. By incorporating the attention mechanism, the model effectively assigns higher weights to relevant regions of the input image, while diminishing the influence of less informative areas. This selective highlighting of features enables the model to focus on the critical regions that contribute significantly to accurate segmentation.

While the proposed model demonstrates promising results in pre-operative brain tumor segmentation, it has limitations when dealing with severe anatomical variations or rare pathologies in the post-operative scenario. Future work may include assessing the model's impact by initializing it with weights from other well-known backbone architectures, like DenseNet or EfficientNet, to gauge its impact on segmentation accuracy. Additionally, the application of efficient augmentation methods [29] to the available post-operative dataset, coupled with fine-tuning the pretrained segmentation model using the augmented post-operative dataset, could effectively address the current model's limitations and significantly enhance its capabilities in the post-operative recurrent tumor scenario.

6 CONCLUSION

In conclusion, this study introduces the Hybrid Attention-Residual UNET with Transformer Blocks (HART-UNet), an innovative deep learning model for brain tumor segmentation. By incorporating a spatial self-attention module and deep residual connections into the U-Net architecture, and initializing with RESNET50 weights, the model excels in capturing intricate details from input data. Training and validation on diverse datasets, including BRATS'20, Kaggle LGG, and BTC_postop, showcase the model's superior performance. Achieving a Dice Coefficient of 0.96 on BRATS'20, 0.97 on Kaggle LGG, and 0.88 on BTC_postop, HART-UNet outperforms established models. The study contributes an efficient solution for precise brain tumor segmentation in both pre-operative and post-operative MRI scans, laying the groundwork for advancements in medical image analysis.

7 DATA AVAILABILITY

1. Kaggle LGG: https://www.kaggle.com/datasets/mateuszbuda/lgg-mri-segmentation.
2. BRATS'20: https://www.kaggle.com/datasets/awsaf49/brats20-dataset-training-validation.
3. BTC_postop: https://openneuro.org/datasets/ds002080/versions/4.0.0/derivatives.

8 STATEMENTS & DECLARATIONS

8.1 Funding

The authors affirm that they did not receive any financial assistance, grants, or support while preparing this manuscript.

8.2 Competing interests

The authors have no relevant financial or non-financial interests to disclose.

8.3 Author contributions

Each of the authors contributed to the conception and design of the study. Sobha Xavier P handled material preparation, data collection, and analysis. Additionally, Dr. Sathish P K and Dr. Raju G were responsible for material preparation and analysis. The final manuscript was approved by all authors.

8.4 Ethics approval

We have duly confirmed that ethical approval is not required for this study.

8.5 Consent to participate

As there were no individual participants involved in this study, no individual participant consent was necessary.

8.6 Consent to publish

This study does not involve any related information regarding consent.

9 REFERENCES

[1] M. Havaei *et al.*, "Brain tumor segmentation with deep neural networks," *Medical Image Analysis,* vol. 35, pp. 18–31, 2017. https://doi.org/10.1016/j.media.2016.05.004

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90

[3] Xavier, P. Sobha, G. Raju, and S. U. Asawthy, "Pre and post operative brain tumor segmentation and classification for prolonged survival," in *International Conference on Soft Computing and Pattern Recognition*, Cham: Springer International Publishing, 2021, pp. 608–616. https://doi.org/10.1007/978-3-030-96302-6_57

[4] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, and B. Glocker, "Attention u-net: Learning where to look for the pancreas," 2018. arXiv preprint arXiv:1804. 03999. https://doi.org/10.48550/arXiv.1804.03999

[5] L. Sun, S. Zhang, H. Chen, and L. Luo, "Brain tumor segmentation and survival prediction using multimodal MRI scans with deep learning," *Frontiers in Neuroscience,* vol. 13, p. 810, 2019. https://doi.org/10.3389/fnins.2019.00810

[6] K. Kamnitsas *et al.*, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis,* vol. 36, pp. 61–78, 2017. https://doi.org/10.1016/j.media.2016.10.004

[7] 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: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18, pp. 234–241. Springer International Publishing, 2015. https://doi.org/10.1007/978-3-319-24574-4_28

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1026–1034. https://doi.org/10.1109/ICCV.2015.123

[9] Y. Zhang and Z. L. Wang, "Joint adversarial learning for domain adaptation in semantic segmentation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 4, 2020, pp. 6877–6884. https://doi.org/10.1609/aaai.v34i04.6169

[10] H. Zhao , Dongyoon, Jiwhan Kim, and Junmo Kim, "Deep pyramidal residual networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5927–5935. https://doi.org/10.1109/TMI.2019.2963101

[11] B. Menze *et al.*, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Transactions on Medical Imaging,* vol. 34, no. 10, pp. 1993–2024, 2014. https://doi.org/10.1109/TMI.2014.2377694

[12] S. Bakas *et al.*, "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features," *Nature Scientific Data*, vol. 4, 2017. https://doi.org/10.1038/sdata.2017.117

[13] Bakas, Spyridon, Mauricio Reyes, Andras Jakab, Stefan Bauer, Markus Rempfler, Alessandro Crimi, Russell Takeshi Shinohara *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).

[14] H. Aerts and D. Marinazzo, "BTC_postop," *OpenNeuro*, 2022. https://doi.org/10.18112/openneuro.ds002080.v4.0.0

[15] A. Tursynova, B. Omarov, A. Sakhipov, and N. Tukenova, "Brain stroke lesion segmentation using computed tomography images based on modified U-Net model with ResNet blocks," *International Journal of Online & Biomedical Engineering,* vol. 18, no. 13, 2022. https://doi.org/10.3991/ijoe.v18i13.32881

[16] F. Milletarì, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D vision (3DV)*, IEEE, 2016, pp. 565–571. https://doi.org/10.1109/3DV.2016.79

[17] X. Guan *et al.*, "3D AGSE-VNet: An automatic brain tumor MRI data segmentation framework," *BMC Medical Imaging,* vol. 22, no. 1, pp. 1–18, 2022. https://doi.org/10.1186/s12880-021-00728-8

[18] Z. Zhou, M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4, pp. 3–11. Springer International Publishing, 2018. https://doi.org/10.1007/978-3-030-00889-5_1

[19] R. Yousef *et al.*, "U-Net-Based models towards optimal MR brain image segmentation," *Diagnostics,* vol. 13, no. 9, p. 1624, 2023. https://doi.org/10.3390/diagnostics13091624

[20] Chen, Jieneng, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan L. Yuille, and Yuyin Zhou, "Transunet: Transformers make strong encoders for medical image segmentation," arXiv preprint arXiv:2102.04306 (2021).

[21] Rajkumar and R. Priya, "Brain tumor segmentation using a hybrid approach for medical images," in *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, pp. 1443–1446, 2020. https://doi.org/10.1109/ICACCS48705.2020.9287658

[22] M. A. Naser and M. J. Deen, "Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images," *Comput Biol Med*, vol. 121, p. 103758, 2020. Epub 2020 Apr 22. PMID: 32568668. https://doi.org/10.1016/j.compbiomed.2020.103758

[23] Jimenez-Sanchez, "ResNet-50 pre-trained weights in ImageNet for TensorFlow," *Zenodo*, 2022. https://doi.org/10.5281/zenodo.1208042

[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90

[25] H. Chen, D. Jiang, and H. Sahli, "Transformer encoder with multi-modal multi-head attention for continuous affect recognition," *IEEE Transactions on Multimedia*, 2020. https://doi.org/10.1109/TMM.2020.3037496

[26] R. R. Shamir, Y. Duchin, J. Kim, G. Sapiro, and N. Harel, "Continuous dice coefficient: A method for evaluating probabilistic segmentations," arXiv preprint arXiv:1906.11031 (2019). https://doi.org/10.48550/arXiv.1906.11031

[27] F. J. D. Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, and D. González-Ortega, "A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network," *In Healthcare*, MDPI, vol. 9, no. 2, p. 153, 2021. https://doi.org/10.3390/healthcare9020153

[28] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks," in *Medical Image Understanding and Analysis: 21st Annual Conference*, MIUA 2017, Edinburgh, UK, July 11–13, 2017, Proceedings 21, pp. 506–517. Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-60964-5_44

[29] S. M. Mangj, P. H. Hussan, and W. M. R. Shakir, "Efficient deep learning approach for detection of brain tumor disease," *International Journal of Online & Biomedical Engineering*, vol. 19, no. 6, 2023. https://doi.org/10.3991/ijoe.v19i06.40277

10 AUTHORS

Sobha Xavier P Research Scholar, in School of Engineering and Technology, Christ Deemed to be University, Bangalore Campus. Assistant Professor in the Department of Computer Science and Engineering at Jyothi Engineering College. Her primary research interests include Image Processing, Computer Vision, Data Science, Machine Learning, and Deep Learning (E-mail: sobha.xavier@res.christuniversity.in).

Sathish P K Associate Professor in the Department of Computer Science and Engineering at the School of Engineering and Technology, Christ Deemed to be University, Bangalore Campus. He holds a Ph.D. degree from Visvesvaraya Technological University, completed in 2019. He also has an AMIE certification and a Master's degree (MTech). His expertise lies in the fields of Computer Science, Artificial Intelligence, Image Processing, Computer Vision, and Pattern Recognition.

Raju G Professor in the Department of Computer Science and Engineering at the School of Engineering and Technology, Christ Deemed to be University, Bangalore Campus. He has a diverse educational background, with a Master's degree in Physics, a Master's degree in Computer Applications, and a Doctoral degree in Computer Science from the University of Kerala. Additionally, he completed his M. Tech in Computer Science and IT from Manonmaniam Sundaranar University, Tirunelveli. His primary research interests include Image Processing, Computer Vision, Data Science, Machine Learning, and Deep Learning. He has published over 120 research articles and has supervised 24 Ph.D. scholars.