Integrated U-Net 3+ with Decomposed Multi head Self-Attention Block and Filter Response Normalization for Brain Lesion Localization and Tracking using 3D MRI

*M.*Vinitha[1]*, *K.* Vinod Kumar[2], *Chandu* Delhipolice[3], *Vaddamani* Jagadeeswara Reddy[4], and *K.* Pavan Kumar[2]

[1]Department of Computer Science and Engineering (AIML), St.Peter's Engineering College, Chennai, India
[2]Department of Computer Science and Engineering, Sanskrithi School of Engineering College, Puttaparthi, Sri Sathyasai District, India
[3]Department of Computer Science and Engineering, Koneru lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Vijayawada, India
[4]Department of Computer Science and Engineering (AI&ML), Annamachrya Institute of Technology and Sciences (Autonomous), Kadapa, India

Abstract. In modern era, a fully automated system that uses 3D Magnetic Resonance Imaging (MRI) scans to segment brain tumors is very important for accurate diagnosis and better treatment planning. However, existing model face some challenges due to its morphology, variability and high computational cost. To overcome this problem, this paper proposed U-Net 3+ with Decomposed and Multi head Self-Attention Block with Filter Response Normalization (UNet3+DMSABFRN). The Decom-UNet3+ integrates decomposed convolution with a spatial attention mechanism and employs asymmetric convolutions and depth-wise separable convolutions to enhance directional awareness and improve inter channel feature extraction. In the MSAB layer, several self-attention heads are combined to capture complex relationships, and convolutions applied to adjust size of the output feature maps. After that, a Filter Response Normalization (FRN) layer is used for reducing the effect of batch size on the network. All MRI sequences are experimented with BraTS2020 dataset and by using Adam optimizer with a learning rate of 0.001 on T1 MRI sequence, the model achieved 99.90% accuracy and a Dice Similarity Coefficient (DSC) of 98%, showing that UNet3+DMSABFRN performs better than existing model like U-Net.

1 Introduction

In recent years, brain tumor is a diseases category that creates a severe threat to human life and health characterized by high mortality and recurrence rates [1]. Artificial intelligence (AI) has growth quickly due to its multiple applications across diverse fields and particularly useful in medical imaging domain, where the adoption of computer vision and computer-aided diagnosis (CAD) systems [2]. Therefore, precise tumor segmentation is crucial for accurate diagnosis, effective treatment planning, and tracking cancer progression [3]. The presence of tumor heterogeneity, unclear boundaries and differences in imaging methods make this task very challenging [4]. Recent progress in using 3D MRI scans for brain tumor detection has greatly improved diagnostic accuracy, mainly because of the growing use of deep learning techniques [5].

At present, many brain tumor segmentation studies focus on improving segmentation accuracy but often overlook the issues of too many parameters and the high computational cost of complex algorithms [6]. Brain tumors are generally divided into two types such as primary and secondary. The primary brain tumor starts from brain cells and secondary brain tumor develop from abnormal cells in other areas such as arteries, the cerebral membrane, or even from outside the brain [7]. In the human skull, tumors appear as abnormal growths. To examine the brain, noninvasive techniques like electroencephalography [8]. In recent decades, CNN have worked successfully in medical images [9]. Most explainability methods use saliency-based XAI approaches because medical imaging is mainly linked with visual tasks [10].

In this research, aims to overcome existing challenges, by implementing U-Net 3+ with Decomposed and Multi head Self-Attention Block with Filter Response Normalization (UNet3+DMSABFRN) for brain tumor segmentation.

The main contribution of this research as given as follows,

- The Decom-UNet3+ integrates decomposed convolution with a spatial attention mechanism and employs asymmetric convolutions and depth-wise separable convolutions to enhance directional awareness and improve inter channel feature extraction.
- In the MSAB layer, several self-attention heads are combined to capture complex relationships,

---

* Corresponding author: hereisvinitha@gmail.com

and convolutions applied to adjust size of the output feature maps.

- After that, a Filter Response Normalization (FRN) layer is used for reducing batch size in network.
- The proposed UNet3+DMSABFRNused publicly available dataset like BraTS2020, which help to segment the tumor in brain.

The remaining of this research is structured as follows: section 2 denotes related work, section 3 present proposed methodology and section 4, 5 represent result, discussion and conclusion.

2 Related work

Sidratul Montaha et al. [11] introduced fully automated method using 2D U-Net architecture was applied to separate tumor areas from healthy tissue. The model was tested on all MRI sequences to identify which sequence gives the best performance. This 2D U-Net model was trained on brain MRI dataset like BraTS2020 for each of the four MRL sequences, and the results were compared with the ground truth to determine which sequence provides the most accurate segmentation. However, the current approach faces some challenges like the performance in diverse clinical scenarios, due to a conventional U-Net model, the system does not fully capture the variability across different types of brain tumors.

Evans Kipkoech Rutoh et al. [12] presented an advanced U-Net variant integrating Attention Based Inception blocks (ABI Net) was introduced to segment brain tumors in 3D multimodal MRI images. The attention module used to concentrate on important areas of the input features and handle the challenges associated with brain tumor segmentation such as class imbalance, small object detection and noisy images. The inception module allows the model to extract features at various levels and be able to accurately segment tumors of different sizes and shapes, making it a reliable tool for assisting radiologists in clinical practice. However, the ABI-Net architecture faces challenges due to its limited interpretability, reliance on a single dataset and lack of advanced post-processing techniques.

Nisar Ahmad and Yao-Tien Chen [13] employed Eff-SAM use of 3D framework, was used by adapts SAM for medical applications with the Parameter Efficient Fine-Tuning (PEFT) mechanism. This framework adds adapters with PEFT to adjust SAM's encoder making the model more effectively, work better in medical imaging while keeping it computationally efficient. In addition, strong data preprocessing steps were applied to improve the model's ability to generalize across different datasets. However, the segmentation models lack the flexibility and scalability needed to address a wide range of complex medical imaging tasks across different organs, due to contextual understanding, prompt encoder adaptability, and computational efficiency.

Meshari D. Alanazi et al. [14] explored an Artificial Intelligence (AI) framework that combines Convolutional Long Short-Term Memory (ConvLSTM) layers with Densely Connected Convolutions (DCC). It helps the system learn detailed image features and track changes over time across different MRI slices. The M-BDCU-Net model provides a practical and robust solution to this challenge, making it a valuable tool for the medical imaging and scientific communities. However, despite the model's promising efficiency and robustness, its current scope may still be limited in addressing the full spectrum of clinical and research demands. Due to the absence of extended configurations and broader validation across diverse medical scenarios. Yu Wang et al. [15] presented Multi-View Coupled Cross-Modal Attention Network (MSegNet), an innovative transformer to help separate parts of medical images more accurately. It combines cross-modal attention mechanisms with multi-view design to fully utilize the spatial and depth information in multimodal MRI data. This allows MSegNet to effectively capture detailed intermodal relationships and model long-range dependencies. However, the MSegNet model still faces challenges in adapting to diverse medical imaging applications and tumor types, due to its limited ability to process unique anatomical regions, optimize 3D MRI segmentation for brain tumors and generalize across different medical devices.

3 Proposed Methodology

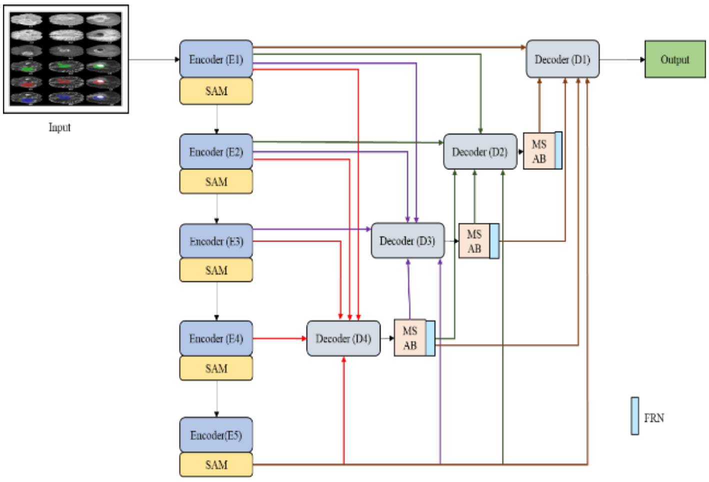

Fig. 1. Overall architecture of proposed U-Net 3+ with decomposed multi head self-attention block and filter response normalization.

This research proposed UNet3+DMSABFRN for brain tumor segmentation, first data are collected from dataset. Next, pre-processing is applied such as normalization, rescaling and transformer, finally the proposed UNet3+DMSABFRN is used to segment the tumor in brain and the overall architecture presented in Figure 1.

3.1 Dataset Acquisition

All experiments are conducted using publicly available dataset like BraTS2020 and detailed description of the dataset is given below:

3.1.1 BraTS2020 dataset

The BraTS2020 dataset containing 369 patient cases, with 293 cases for High Grade Glioma (HGG) and 76 cases for Low Grade Glioma (LGG), which validation test contains 125 patient cases. This dataset contains mpMRI scans from patients with glioblastoma and other parts of brain tumors and composed of four different MRI types such as FLAIR, T1, T1ce, T2. The segmentation masks in the BraTS 2020 dataset annotations contain four labels such as Edema (Ed), Non-Enhancing Tumor (NET), Background and Enhancing Tumor (ET). the training dataset is further class-wise split into validation set and test set. 80% of total data of the training split is utilized for training, 10% of the data is taken for validation purposes 10% of the data is used for the testing of the proposed model, and the dataset split was stratified [16].

3.2 Data pre-processing

The 3D MRI scans is often difficult and computationally intensive, so pre-processing techniques are required to enhance the performance of proposed method.

3.2.1 Normalization

To standardize all numerical features to the same scale and min-max normalization is applied. This technique used to adjust feature values within the range of 0 to 1.

3.2.2 Rescaling

The dataset is resized to 128×128×1 voxel by using only the middle slice instead of all 155 slices from the original 3D brain MRI image are 240 × 240 × 155 pixels.

3.2.3 Transformation

Several transformations were applied to the input images, including random rotations translations, shearing, zooming, and horizontal flipping. These transformations were carefully chosen to simulate variations in image acquisition, such as different angles, patient positioning and magnification levels, while maintaining spatial coherence across different MRI modalities.

3.3 Proposed U-Net 3+ with Decomposed Multi head Self-Attention Block and Filter Response Normalization (UNet3+DMSABFRN)

The proposed UNet3+DMSABFRN enhances brain tumor segmentation by combining U-Net3+ decompose and multi head self-attention block with filter response normalization. The U-Net 3+ backbone aggregates features from all encoder and decoder stages, allowing the model to capture both fine tumor boundaries and broader contextual information across varying tumor sizes. The Decomposed Multi-Head Self-Attention Block improves long-range dependency modeling while reducing the computational cost of standard self-attention, enabling the network to focus more effectively on tumor-relevant regions regardless of their position and shape. Filter Response Normalization replaces traditional batch normalization, providing stable convergence and robust performance even with small medical datasets and limited batch sizes, which is common in brain MRI analysis. Together, these components improve the accuracy of tumor boundary delineation, enhance detection of small and irregular tumor regions, and provide more reliable segmentation maps for clinical decision-making.

3.3.1 Decomposed Convolution Module

The encoder of Decom-UNet3+ is a decomposed convolution module, designed to enhance computational efficiency and feature extraction capabilities for retinal vessel segmentation. Within this module, multi-path asymmetric convolutions independently capture multi-channel spatial information along the height and width dimensions, improving directional sensitivity. These are combined with depth-wise separable convolutions, which extract local features in both width and height dimensions for each channel separately, ensuring independent computations within the 2D space. This design significantly reduces computational complexity while improving model efficiency. The asymmetric convolution extends the receptive field at a low computational cost through successive 1-D convolution operations, enhancing structural feature representation. This improves the fine-grained segmentation of retinal vessels while minimizing redundant computations and optimizing feature extraction. Finally, pointwise convolutions facilitate cross-channel interactions, optimizing feature distribution and performing channel selection, ensuring high-quality feature representations for multi-scale decoding. By employing a sequential combination of 3 × 1 and 1 ×3 asymmetric convolutions, a 3 × 3 receptive field can be constructed, while two consecutive applications of this combination can further expand the receptive field to 5 × 5. This approach enables encoders to capture multi-scale feature information. The depth-wise separable convolution introduced in the decomposed convolution module strengthens feature associations along the width and height dimensions, improving segmentation of thick vessels and preventing the formation of voids within vessel regions.

The algorithms for the multi-scale convolution with directional awareness in the decomposed convolution module are presented Equations (1)-(3),

$$F_{multi}(\mathrm{X}) = Concat(F_{3\times3}(\mathrm{X}), F_{5\times5}(\mathrm{X}), F_{depthwise}(\mathrm{X})) \quad (1)$$

$$F_{5\times5}(\mathrm{X}) = F_{3\times3}(F_{3\times3}(\mathrm{X})) \quad (2)$$

$$F_{3\times3}(\mathrm{X}) = \sigma(W_{1\times3} * W_{3\times1} * \mathrm{X}) \quad (3)$$

Equation (1) denotes the multi-scale modeling capability of the decomposed convolution module, while (3) employs asymmetric 1×3 and 3×1. 1-D convolutions, the receptive field is maintained at 3 × 3 while reducing computational complexity. This design also increases the model depth, enabling higher-level feature representations and enhancing both feature expressiveness and generalization ability.

3.3.2 Filter Response Normalization (FRN)

In U-Net 3+, batch normalization (BN) is commonly used to normalize convolution outputs, but its performance heavily depends on the batch size N; with small batches, the results are poor, and although BN performs better with larger batches, it still has limitations. Filter Response Normalization (FRN) addresses this issue by removing the dependency on batch size and often surpasses BN when NN is large. Unlike BN, FRN does not subtract the mean value instead, it replaces variance with the mean of the squared norm v^2, followed by scaling and shifting with a small constant to prevent division errors. Since FRN avoids mean subtraction, its outputs may deviate from zero, and when combined with ReLU activation, this can lead to many values collapsing to 0 or 1, which may hinder effective training and reduce model performance. τ denotes learnable and use threshold ReLU to eliminate the bias phenomenon, namely TLU, as shown in Equation (4),

$$Z = \max(y, \tau) = Relu(y - \tau) + \tau \quad (4)$$

3.3.3 Multi head Self-Attention Block

Segmentation requires highly accurate pixel-level predictions, where both detailed features and long-range dependencies are important to clear up ambiguities in local predictions. For large SAR images, understanding the inherent connections between pixels improves classification accuracy, especially for small regions. However, CNNs struggle to capture these global contextual relationships because their receptive fields are limited to local neighborhoods. To address this, a Multi-Scale Attention Block (MSAB) is added in the later stage of encoder long-range dependencies in SAR images are captured. In MSAB, the Multi-Head Self-Attention (MHSA) layer understands complex patterns by combining the outputs of multiple attention heads and a 1×1 convolution is applied to adjust the output feature map dimensions. Let x represent input feature and y denotes output feature map and mathematical Equations (5)-(7) is given below:

$$y = concat(y_1, y_2\ , \ldots\ldots, y_n) \quad (5)$$

$$y_i = softmax\left(\frac{Q.K^T}{\sqrt{d_i}}\right).V \quad (6)$$

$$Q = W_q x, K = W_k x, V = W_v x \quad (7)$$

Where, y_i denotes output feature map and $d_{i=}\, {}^{d}/_{n}$ denotes dimension of head. Three matrices Q, K and V are created, which correspond to queries, keys and values. In Equation (6), the attention weights for each self-attention head are first calculated using the dot product of Q and K. These attention maps are normalized using dimension di then passed through a SoftMax function to generate attention scores, which helps the model understand the global information. Finally, the output y_i is produced by combining the values in V based on these scores.

4 Result and discussion

This section discusses the details related to the evaluation metrics used to examine the proposed UNet3+DMSABFRN model's performance. Furthermore, the results obtained through the proposed method with detailed discussion are presented in below:

4.1 Experiments

All experiments are conducted on Google Colab Pro (GCP) server equipped with a 35 GB v100 GPU and 25 GB High RAM. Various experiments were performed on the proposed method to determine the optimal combination of hyperparameters.

4.2 Evaluation metrics

This section discusses about performance of the proposed UNet3+DMSABFRN is evaluated using metrics such as Accuracy, Dice score, specificity and sensitivity. These metrics are employed to measure and assess the accuracy and effectiveness of the segmentation methods. This metrics are discussed below:

4.2.1 Accuracy

Accuracy is used to prediction of accurate number is divided by the total number of predictions is used to measure the model's accuracy is presented in Equation (8),

$$Accuracy = \frac{TrueNegative + TruePositive}{TrueNegative + FalsePositive + TruePositive + FalseNegative} \quad (8)$$

4.2.2 Dice score

$$\boldsymbol{Dice\ score} = \left(\frac{2*(\boldsymbol{Intersection\ between\ predicted\ and\ ground\ truth\ masks})}{\boldsymbol{Size\ of\ predicted\ mask+Size\ of\ graound\ truth\ mask}}\right) \tag{9}$$

The dice score or F1-score, is a common metric to evaluating the performance of brain tumor segmentation, which range from 0 to 1 shown in Equation (9).

4.2.3 Sensitivity

The sensitivity calculated the ratio of true positive cases that are accurately identified by the segmentation method. It measures the ability of the segmentation approach to identify the tumor region in the brain correctly presented in Equation (10),

$$\boldsymbol{Sensitivity} = \left(\frac{\boldsymbol{True\ positive}}{\boldsymbol{True\ positive+False\ negative}}\right) \tag{10}$$

4.2.4 Specificity

Specificity computes the ratio of true negative cases that are accurately identified by the segmentation method. It measures the ability of the segmentation technique to identify the healthy regions in the brain correctly presented in Equation (11),

$$\boldsymbol{Specificity} = \left(\frac{\boldsymbol{True\ negative}}{\boldsymbol{True\ negative+False\ positive}}\right) \tag{11}$$

4.3 Performance analysis

Table 1. Performance analysis for proposed method using BraTS2020 dataset.

Models	**Accuracy (%)**	**Dice score (%)**			**Specificity (%)**			**Sensitivity (%)**		
		WT	**TC**	**ET**	**WT**	**TC**	**ET**	**WT**	**TC**	**ET**
U-Net	98.54	94.65	95.85	94.87	98.65	98.56	98.46	97.34	92.86	96.85
V-Net	97.93	95.73	93.76	95.73	97.87	97.25	95.73	93.74	97.45	97.39
TransBTS	98.35	97.37	89.76	97.42	97.36	98.35	96.90	95.43	96.99	89.99
Mobile Net	99.34	96.54	96.87	95.25	96.46	99.24	96.45	98.42	94.75	95.79
Proposed UNet3+DMSABFRN	99.90	97.90	98.67	97.97	99.80	99.50	98.78	98.99	97.86	97.89

This section evaluates the performance of proposed UNet3+DMSABFRN by comparing against existing networks such as U-Net, V-Net, TransBTS and Mobile Net. This comparison aims to demonstrates performance of the proposed UNet3+DMSABFRN in terms of Dice score, sensitivity and specificity. In experiment, the BraTS2020 dataset compared with existing methods, which results are present in Table 1.

4.4 Computational complexity

Table 2. Comparison of inference time and computational efficiency for financial fraud detection.

Models	**Parameters (in millions)**	**Memory size**	**Inference time (in seconds)**
U-Net	9.6	0.40	1200
V-Net	8.54	0.39	1400
TransBTS	7.54	0.68	900
Mobile Net	7.9	0.54	1000
Proposed UNet3+DMSABFRN	6.7	0.30	700

This section used to evaluate the complexity of UNet3+DMSABFRN, which higher compared to existing methods such as U-Net, V-Net, TransBTS and Mobile Net. The proposed UNet3+DMSABFRN demonstrated 700 seconds, memory size also low and parameters. Table 2 represent the comparison of parameters, memory size and inference time across existing methods.

4.5 Statistical Analysis

This section validates statistical, p-value (<0.05) were computed using Python's scipy.stats.ttest_rel function. Training and evaluation were performed five times (one for each fold), ensuring statistical robustness and fair comparison. Notably, the p-values for WT, TC and ET segments all are less than 0.05. Consequently, at the 0.05 significance level, the segmentation dice scores of UNet3+DMSABFRN were shown to be significantly better than those of U-Net. This statistical analysis offers valuable insights into the ability of proposed model to perform effectively on new, previously unseen data.

4.6 Ablation study

In this research, UNet3+DMSABFRN model is proposed to handle and automate the process of segmentation. The proposed UNet3+DMSABFRN get superior results by comparing to existing methods. First, the baseline U-Net3+, second U-Net3+ incorporate with decomposed and U-Net3+ incorporate with FRN. The output of this discussion is presented in Table 3.

Table 3. Ablation study for proposed method for BraTS2020 dataset.

Models	Accuracy (%)	Dice score (%)			Specificity (%)			Sensitivity (%)		
		WT	TC	ET	WT	TC	ET	WT	TC	ET
Baseline U-Net3+	90.93	95.90	96.53	93.76	94.64	97.40	93.67	94.32	94.56	93.64
Decom with U-Net3+	93.54	96.26	97.34	94.45	96.57	95.89	95.78	96.56	96.84	94.56
U-Net3+ with FRN	97.37	95.43	95.67	96.34	98.63	98.56	97.99	97.42	96.99	96.99
Proposed UNet3+DMSABFRN	99.90	97.90	98.67	97.97	99.80	99.50	98.78	98.99	97.86	97.89

4.7 Comparative analysis

The proposed UNet3+DMSABFRN achieved better results while compared to state-of-arts such as U-Net, ABI-Net, Eff-SAM, M-BDCU-NET and MSegNet. The performance is evaluated using BraTS2020 dataset, the proposed UNet3+DMSABFRN achieved accuracy of 99.90% and Dice score in WT for 97.90%, TC for 98.67% and ET for 97.97. Table 4 present the comparison of the proposed method with existing models.

Table 4. Comparative analysis for proposed method using BraTS2020 dataset.

Models	Accuracy (%)	Dice score (%)			Specificity (%)			Sensitivity (%)		
		WT	TC	ET	WT	TC	ET	WT	TC	ET
U-Net [11]	99.41	93			99.68			NA		
ABI-Net [12]	NA	87.8	85.1	83.5	NA			NA		
Eff-SAM [13]	Na	88.4	85.3	81.8	99.4	98.6	97.7	97.1	96.2	95.5
M-BDCU-NET [14]	98	82			NA			98		
MSegNet [15]	NA	96.28	97.66	95.81	91.15			90.50		
Proposed UNet3+DMSABFRN	99.90	97.90	98.67	97.97	99.80	99.50	98.78	98.99	97.86	97.89

5 Discussion

In this research, proposed U-Net 3+ with Decomposed and Multi head Self-Attention Block with Filter Response Normalization (UNet3+DMSABFRN). The Decom-UNet3+ integrates decomposed convolution with a spatial attention mechanism and employs asymmetric convolutions and depth-wise separable convolutions to enhance directional awareness and improve inter channel feature extraction. In the MSAB layer, several self-attention heads are combined to capture complex relationships and convolutions applied, which used to adjust the size of the output feature maps. After that, a Filter Response Normalization (FRN) layer is used to reduce effect of batch size on the network. Experiments on BraTS2020 dataset demonstrates that the proposed UNet3+DMSABFRN has fewer parameters and superior segmentation performance. The experiments were only tested using the p-value, which demonstrated a statistically significant difference between the proposed and the existing methods. The experimental results show that the 6.7 of parameters, 0.30 is memory size and 700 for inference time; dice score for WT, TC and ET are 97.90%, 98.67% and 97.97%.

6 Conclusion

This research presents UNet3+DMSABFRN, a new deep learning framework designed for segmentation of tumor in 3D MRI brains. This proposed model combines decom, MSAB layer and filter response normalization in U-Net 3+ framework. This section discusses about performance of the proposed UNet3+DMSABFRN is evaluated using metrics such as Accuracy, Dice score, specificity and sensitivity. The proposed model achieved a dice score for WT, TC and ET are 97.90%, 98.67% and 97.97%. The proposed UNet3+DMSABFRN trained on BraTS 2020 dataset by using four MRI sequences such as FLAIR, T1, T1ce, T2 to find, which one give best segmentation results compared to the ground truth 3D MRI images. The proposed UNet3+DMSABFRN get superior results by comparing to existing methods. First, the baseline U-Net3+, second U-Net3+ incorporate with decomposed and U-Net3+ incorporate with FRN. In future, UNet3+DMSABFRN, model's adaptability will be improved across different MRI modalities and multi-institutional datasets for better robustness. Further optimization for efficiency and multi-class tumor tracking could strengthen its practical applications.

References

1. X. Liu, J. Xia, W. Gu, T. Yao, J. Shen, D. Tang, MISA-Net: A multi-scale feature interaction network for brain tumor segmentation. IRBM **46**,

100891 (2025). https://doi.org/10.1016/j.irbm.2025.100891

2. C. Aumente-Maestro, M. Müller, B. Remeseiro, J. Díez, M. Reyes, AUDIT: An open-source Python library for AI model evaluation with use cases in MRI brain tumor segmentation. Comput. Methods Programs Biomed. **271**, 108991 (2025). https://doi.org/10.1016/j.cmpb.2025.108991
3. S. Kundu, S. Dutta, J. Mukhopadhyay, N. Chakravorty, FFLUNet: Feature fused lightweight UNet for brain tumor segmentation. Comput. Biol. Med. **194**, 110460 (2025). https://doi.org/10.1016/j.compbiomed.2025.110460
4. H. Hedibi, M. Beladgham, A. Bouida, A combined attention mechanism for brain tumor segmentation of lower-grade glioma in magnetic resonance images. Comput. Biol. Med. **193**, 110380 (2025). https://doi.org/10.1016/j.compbiomed.2025.110380
5. Q. Guo, Y. Wang, Y. Zhang, H. Qi, Y. Hu, Y. Jiang, Hyper-BTS: Brain tumor segmentation based on hypergraph guidance. Pattern Recognit. **169**, 111926 (2026). https://doi.org/10.1016/j.patcog.2025.111926
6. L. Zhang, Y. Li, Y. Liang, T. Liu, W. Zhang, J. Sun, A lightweight segmentation model based on dilated multi-scale residual attention U-Net for brain tumor segmentation. Eng. Appl. Artif. Intell. **160**, 111925 (2025). https://doi.org/10.1016/j.engappai.2025.111925
7. E. Güvenç, M. Ersoy, G. Çetin, Deep learning-based automated segmentation of brain tumors using synthetic MR images generated with DCGAN. Comput. Stand. Interfaces **96**, 104054 (2025). https://doi.org/10.1016/j.csi.2025.104054
8. M. Obayya, A. Alshuhail, K. Mahmood, M.H. Alanazi, M. Alqahtani, N.O. Aljehane, H. Almansour, M.A. Al-Hagery, A novel U-net model for brain tumor segmentation from MRI images. Alex. Eng. J. **126**, 220-230 (2025). https://doi.org/10.1016/j.aej.2025.04.051
9. H. Luo, D. Zhou, Y. Cheng, S. Wang, MPEDA-Net: A lightweight brain tumor segmentation network using multi-perspective extraction and dense attention. Biomed. Signal Process. Control **91**, 106054 (2024). https://doi.org/10.1016/j.bspc.2024.106054
10. S.B. Shaheema, N.B. Muppalaneni, Explainability-based panoptic brain tumor segmentation using a hybrid PA-NET with GCNN-ResNet50. Biomed. Signal Process. Control **94**, 106334 (2024). https://doi.org/10.1016/j.bspc.2024.106334
11. S. Montaha, S. Azam, A.K.M. Rakibul Haque Rafid, M.Z. Hasan, A. Karim, Brain tumor segmentation from 3D MRI scans using U-Net. SN Comput. Sci. **4**, 386 (2023). https://doi.org/10.1007/s42979-023-01854-6
12. E.K. Rutoh, Q. ZhiGuang, J.C. Bore-Norton, N. Bahadar, ABI-Net: Attention-based inception U-Net for brain tumor segmentation from multimodal MRI images. IEEE Access. **13**, 134898-134916 (2025). https://doi.org/10.1109/ACCESS.2025.3585926
13. N. Ahmad, Y.T. Chen, Eff-SAM: SAM-based efficient method for brain tumor segmentation in multimodal 3D MRI scans. IEEE Access. **13**, 90063-90087 (2025). https://doi.org/10.1109/ACCESS.2025.3571464
14. M.D. Alanazi, A. Maraoui, I. Werda, A.B. Atitallah, T.M. Alanazi, M. Albekairi, A. Sahbani, A. Yousef, Magnetic resonance imaging brain segmentation using bi-directional convolutional long short-term memory U-Net with densely connected convolutions. IEEE Access. **13**, 75847-75860 (2025). https://doi.org/10.1109/ACCESS.2025.3562132
15. Y. Wang, J. Xu, Y. Guan, F. Ahmad, T. Mahmood, A. Rehman, MSegNet: A multi-view coupled cross-modal attention model for enhanced MRI brain tumor segmentation. Int. J. Comput. Intell. Syst. **18**, 63 (2025). https://doi.org/10.1007/s44196-025-00787-7
16. The BraTS2020 dataset link: https://www.kaggle.com/datasets/awsaf49/brats2020-training-data/code.m