Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 14
TPCEE 2026

Research on Multi sequence Magnetic Resonance Imaging for Brain Tumor Classification and Fine Boundary Segmentation

Moses Lesa, Yong Liu *

Nanjing University of Science and Technology, Nanjing, China

* Corresponding Author Email: liuy1602@njust.edu.cn

Abstract. This paper proposes a unified framework for joint classification and boundary fine segmentation to address the issues of separation and insufficient information complementarity in brain tumor classification and segmentation tasks in multi sequence magnetic resonance imaging. This method first designs a multimodal adaptive fusion module, utilizing channel attention mechanism to dynamically allocate weights to the four sequences T1, T1CE, T2, and FLAIR, achieving effective integration at the feature level; Secondly, a dual task joint learning architecture is constructed, and classification and segmentation branches are set up on the basis of a shared encoder. Gradient equilibrium collaborative optimization is achieved through a homoscedastic uncertainty weighted loss function; Finally, a boundary aware refinement module is introduced in the segmentation decoder, which constrains the accuracy of tumor contours using sub differentiable Hausdorff distance loss. Experiments on the BraTS 2020 dataset showed that this method achieved Dice coefficients of 90.67%, 83.92%, and 79.81% for the whole tumor (WT), tumor core (TC), and enhanced tumor (ET) regions, respectively; the average HD95 was reduced to 5.12 mm; and the classification accuracy reached 93.6%. with an AUC of 0.957. All metrics outperformed single-task and serial pipeline baselines, validating the effectiveness of multi-task collaboration and boundary constraints.

Keywords: Brain tumor; Multi sequence MRI; Boundary refinement.

1. Introduction

Brain tumor is one of the most common and threatening malignant diseases in the central nervous system, among which glioma has the highest incidence rate and the worst prognosis among brain tumors because of its highly heterogeneous and invasive growth characteristics [1]. Therefore, early and accurate classification and boundary localization of brain tumors are of great clinical significance for developing personalized treatment plans and improving patient prognosis.

Magnetic resonance imaging (MRI) has become the preferred imaging method for clinical diagnosis and preoperative evaluation of brain tumors due to its advantages of non radiation, high soft tissue resolution, and multi parameter imaging [2]. In routine clinical practice, a combination of MRI sequences is typically used for imaging, including T1-weighted (T1) images, T1-weighted with contrast enhancement (T1ce) images, T2-weighted (T2) images, and fluid attenuated inversion recovery (FLAIR) images. Different sequences, starting from different physical mechanisms, can reflect information such as the anatomical structure of tumors, the degree of blood-brain barrier damage, the extent of edema, and the area of necrosis. For example, T1ce sequence can clearly display the enhanced core area of the tumor, while FLAIR sequence has high sensitivity to peritumoral edema [3]. However, how to effectively integrate the aforementioned multiple sequence information to achieve precise classification of tumor types and fine segmentation of tumor boundaries simultaneously remains an important challenge in the current field of medical imaging analysis.

In recent years, the rapid development of deep learning technology has provided powerful tools for automatic analysis of brain tumors. Researchers have proposed binary or multi classification methods based on convolutional neural networks (CNN), which automatically distinguish between low-grade and high-grade gliomas by extracting high-dimensional features from MRI images [4-5]. Encoder–decoder architectures, as exemplified by U-Net and its variants, have made significant progress and

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

are capable of performing pixel-level segmentation of the whole tumor (WT), tumor core (TC), and enhancing tumor (ET) regions [6–7]. In addition, Transformer based methods have also been introduced to capture global dependencies over longer distances [8]. However, existing research mostly treats classification and segmentation as two independent tasks, lacking collaborative optimization under a unified framework. The classification task focuses on global semantic judgment, while the segmentation task relies on local fine features, and there is a natural information complementarity between the two.

In response to the above issues, this paper proposes a joint classification and fine boundary segmentation framework for brain tumors based on multi sequence MRI. Firstly, a multimodal adaptive fusion module was designed to dynamically allocate weights to different MRI sequences through channel attention mechanism, achieving effective integration at the feature level; Secondly, a dual task joint learning architecture was constructed, which sets up classification and segmentation branches separately based on the shared feature extraction backbone, and achieves collaborative optimization through a multi task loss function; Thirdly, a boundary aware refinement module is introduced in the segmentation decoder to constrain the accuracy of tumor contours using a boundary loss function, significantly improving the quality of segmentation boundaries. This article conducted systematic experiments on the internationally recognized BraTS 2020 dataset, and verified the effectiveness of the proposed method in terms of classification accuracy and segmentation accuracy through comparisons with multiple baseline methods and comprehensive ablation experiments.

2. Research methodology

This method uses BraTS 2020 as the experimental benchmark to construct a three-stage architecture consisting of shared encoder, dual task decoding, and boundary refinement (as shown in Figure 1). BraTS 2020 provides four co registered sequences, T1, T1CE, T2, and FLAIR, including 369 training data and 125 validation data, labeled to cover enhanced tumors, tumor cores, and WT domains.

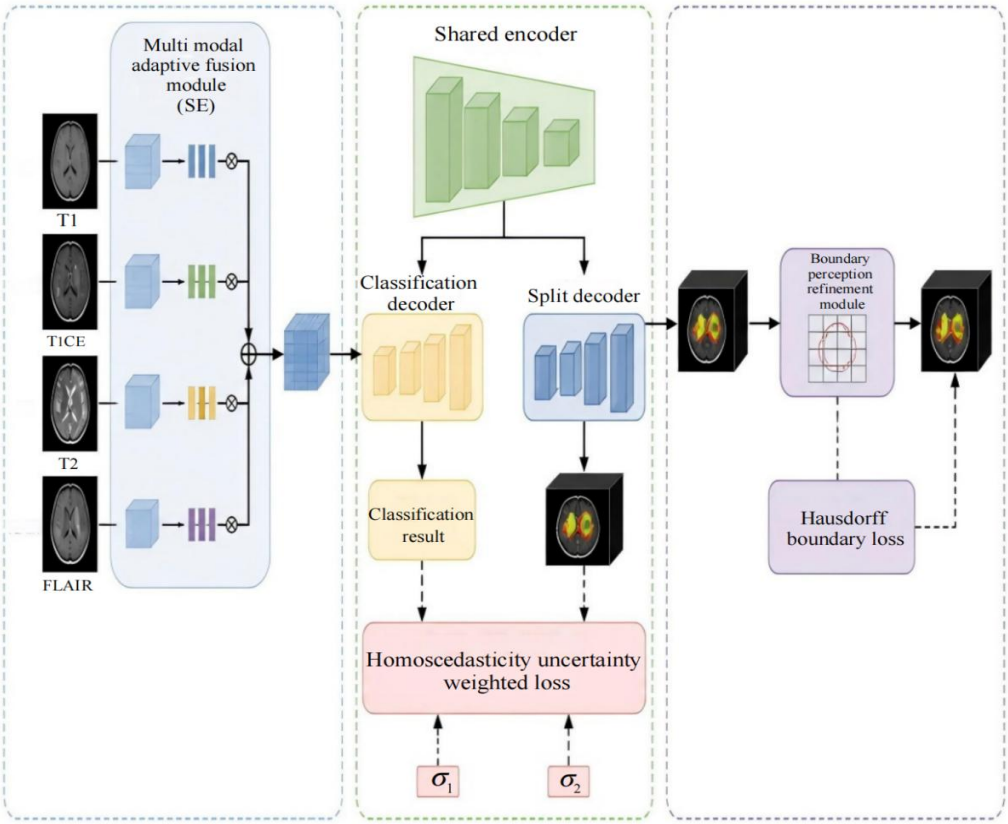

Figure 1. Joint classification and fine boundary segmentation framework for brain tumors based on multi sequence MRI

2.1. Multi modal adaptive fusion module

The four co registered sequences T1, T1CE, T2, and FLAIR exhibit complementary contrast characteristics in different subregions of the tumor. To avoid modal bias caused by simple concatenation or equal-weighted fusion, a Squeeze-and-Excitation (SE) mechanism is introduced at each scale of the shared encoder to perform dynamic reweighting based on modal-level features.

Let the feature maps of four modalities at a certain scale be concatenated channel by channel to obtain $\mathbf{U}\in\mathrm{R}^{C\times H\times W}$, where $C=4C_m$ and C_m is the number of channels in a single modality at that scale. The SE module first performs a squeezing operation and performs global average pooling on each channel:

$$z_c=\frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W}\mathbf{U}_c(i,j),\ c=1,\ldots,C \tag{1}$$

z_c represents the global response strength of the c-th channel, and H and W are spatial dimensions. Subsequently, the triggering operation learns the nonlinear dependencies between channels through two fully connected layers:

$$\mathbf{s}=\sigma(\mathbf{W}_2\,\delta(\mathbf{W}_1\mathbf{z})) \tag{2}$$

Among them, $\mathbf{z}=[z_1,\ldots,z_C]^{\top}$; $\mathbf{W}_1\in\mathrm{R}^{(C/r)\times C}$ is the dimensionality reduction weight, $\mathbf{W}_2\in\mathrm{R}^{C\times(C/r)}$ is the dimensionality enhancement weight, and the compression ratio r controls the bottleneck width (taking $r=16$ is a common configuration); $\delta(\cdot)$ is ReLU, $\sigma(\cdot)$ is Sigmoid, and the output $\mathbf{s}\in(0,1)^C$ is the normalized channel attention weight. The final fusion feature is channel by channel recalibration:

$$\widetilde{\mathbf{U}}_c=s_c\mathbf{U}_c \tag{3}$$

Since the spliced channels are arranged in segments according to the mode, the weight of each channel corresponding to the same mode in $\mathbf{s}$ jointly determines the relative contribution of the mode in the fusion, so that the network can adaptively emphasize the sequence with more information on different scales, without relying on manually specifying the mode priority.

2.2. Dual task joint learning architecture

The multi-scale features $\{\widetilde{\mathbf{U}}^{(l)}\}_{l=1}^{L}$ output by the shared encoder are simultaneously fed into two lightweight decoupling heads, the segmentation decoder and the classification head. The segmentation decoder restores spatial resolution through step-by-step upsampling and skip connections, and outputs a voxel by voxel class probability map $\hat{\mathbf{Y}}^{seg}$; The classification head applies fully connected and Softmax to the global pooling features at the end of the encoder, and outputs the probability distribution $\hat{\mathbf{y}}^{cls}$ of tumor type/level.

The two tasks share representations, but the loss scales are naturally different. The segmentation loss is aggregated on hundreds of thousands of voxels, while the classification loss is only calculated on the category dimension. Directly adding equal weights would overwhelm the classification gradient [9]. To this end, homoscedasticity uncertainty weighting is used to construct the total loss:

$$\mathrm{L}_{total}=\frac{1}{2\sigma_{seg}^2}\mathrm{L}_{seg}+\frac{1}{2\sigma_{cls}^2}\mathrm{L}_{cls}+\log\sigma_{seg}+\log\sigma_{cls} \tag{4}$$

Among them, L_{seg} is the segmentation loss, and L_{cls} is the classification cross entropy; σ_{seg}^2 and σ_{cls}^2 are learnable task noise variances that are automatically adjusted during training to obtain smaller effective weights for high uncertainty tasks, while the $\log\sigma$ term prevents unbounded increase in noise variance. This design eliminates the process of manually searching for task weights, allowing the two objectives to achieve balanced collaboration at the gradient level on a shared encoder.

2.3. Boundary perception refinement module

The standard Dice loss is sensitive to volume overlap, but has a weaker penalty for boundary shift, and the fuzzy boundary caused by glioma invasive growth is precisely the main source of segmentation error. At the final stage of the segmentation decoder, boundary aware refinement is introduced to predict a tumor boundary probability map $\widehat{\mathbf{B}}$ from the end features of the decoder, and a sub differentiable Hausdorff distance loss is applied for boundary constraints [10].

Assuming the predicted boundary point set is P, the true boundary point set is G, and the Hausdorff boundary loss is defined as the bidirectional average minimum distance:

$$L_{HD}=\frac{1}{|P|}\sum_{p_i\in P} d_{\min}(p_i,G)+\frac{1}{|G|}\sum_{g_j\in G} d_{\min}(g_j,P) \tag{5}$$

Where $d_{\min}(p_i,G)=\min_{g_j\in G}\|p_i-g_j\|_2$, represents the Euclidean distance from the predicted boundary point p_i to the nearest truth boundary point; $d_{\min}(g_j,P)$ symmetrically penalizes the truth boundary points of missed segmentation. Unlike the original Hausdorff distance, which takes the maximum and minimum distance, this form performs average aggregation on each boundary point, resulting in denser gradients and easier backpropagation; Simultaneously, bidirectional summation penalizes both false positive and false negative boundaries.

The final loss of the boundary branch is:

$$L_{bd}=L_{HD}+\lambda_{BCE} L_{BCE}(\widehat{\mathbf{B}},\mathbf{B}^*) \tag{6}$$

L_{BCE} provides pixel by pixel boundary classification supervision, while $\mathbf{B}^*$ is obtained by extracting morphological contours from the true value segmentation mask; Take a smaller value of λ_{BCE} to let distance constraints dominate the direction of boundary optimization. The boundary loss L_{bd} is merged into L_{seg}, and together with the classification loss, it is optimized end-to-end through L_{total}.

3. Experimental analysis

The experiment was conducted on the BraTS 2020 training set (369 cases) for five fold cross validation, and the validation set (125 cases) was used for final evaluation. The Dice similarity coefficients and HD95 (mm) of the WT, TC, and ET subregions are reported on the segmentation end, while the accuracy and AUC of IDH mutation state prediction are reported on the classification end.

Table 1 compares the method proposed in this article with two representative baselines. Using nnU Net as the strong CNN segmentation baseline and TransBTS as the Transformer segmentation baseline; The classification task is compared with a single task classification network and a simple "segmentation classification" serial pipeline. All segmentation indicators are averaged on the validation fold of five fold cross validation.

Table 1. Performance comparison with baseline method on BraTS 2020 validation set

Method	WT Dice	TC Dice	ET Dice	Average HD95 (mm)	Classification Acc. (%)	Classification AUC
nnU-Net	88.95	85.06	—	—	—	—
TransBTS	90.98	82.85	78.69	—	—	—
Single task segmentation (encoder+UNet decoding in this article)	89.42	82.31	77.56	5.84	—	—
Single task classification (encoder+classification header in this article)	—	—	—	—	91.3	0.934
Serial pipeline (segmentation before classification)	89.58	82.47	77.92	5.79	91.8	0.938
This article (Joint Framework)	90.67	83.92	79.81	5.12	93.6	0.957

Compared with single task segmentation, the joint framework improved Dice by 1.25 and 1.61 percentage points on WT and TC, respectively, and reduced average HD95 by 0.72 mm; The classification accuracy has improved by 2.3 percentage points compared to single task classification. This indicates that the joint optimization on the shared encoder did not damage the segmentation representation due to task competition, but instead enhanced the encoder's ability to discriminate the overall nature of tumors through global semantic constraints of classification targets. Compared with serial pipeline, the advantages of joint framework are more clear. In the serial scheme, segmentation errors are directly transmitted to the classification head, resulting in limited classification improvement; Joint training enables the classification gradient to participate in representation shaping during the feature extraction stage, and the two tasks are mutually regularized on a shared encoder. Compared with TransBTS, the WT Dice in this paper is slightly lower, but the TC and ET Dice are both higher, indicating that our method has a relative advantage in subregions with smaller volumes and more complex boundaries.

The ablation experiment used shared encoder, standard UNet decoder, and equal weight multi task loss as baselines, gradually introducing SE fusion module, uncertainty weighting, and boundary refinement branch. The results are shown in Table 2. The table reports both classification accuracy and average HD95 to observe the independent and joint effects of each module on the two tasks.

Table 2. Module ablation experiment results (BraTS 2020 validation set, five fold mean)

Configuration	SE fusion	Uncertainty weighting	Boundary refinement	WT Dice	TC Dice	ET Dice	Average HD95	Classification Acc.
B0 (baseline)	✗	✗	✗	88.71	80.84	75.93	6.31	91.2
B1	✓	✗	✗	89.36	81.72	76.88	5.97	91.9
B2	✗	✓	✗	88.94	81.15	76.34	6.08	92.8
B3	✗	✗	✓	89.05	82.47	78.61	5.39	91.1
Complete model	✓	✓	✓	90.67	83.92	79.81	5.12	93.6

Three independent action paths can be read from Table 2. SE fusion brings consistent Dice gain across all three subregions, with ET having the highest gain, indicating that channel attention has a more pronounced modal selection effect on enhancing tumors, which have the highest contrast on T1CE but sparse information on other sequences. The improvement of classification accuracy by uncertainty weighting is significantly greater than that of segmentation. Under equal weight loss, the classification gradient is suppressed by segmentation loss. After introducing learnable noise variance, the classification task obtains effective weights that match it. The effect of boundary refinement is highly concentrated on boundary indicators and ET subregions, with an average HD95 decrease of 0.92 mm and ET Dice increase of 2.68 percentage points, while WT Dice only increases by 0.34. ET is the category with the smallest volume and the most irregular boundary among the three subregions, and its segmentation quality is most sensitive to boundary constraints. This asymmetric gain directly verifies the target of Hausdorff boundary loss.

The complete model is not simply the sum of the gains of each individual module. WT Dice further improved by 1.31 and 1.62 percentage points compared to B1 and B3, respectively, and the classification accuracy improved by 1.7 and 0.8 percentage points compared to B1 and B2, respectively. This indicates that the modal reweighting features provided by SE fusion improve the input quality of the boundary refinement module, while uncertainty weighting enables the classification head to more fully utilize these features, and there is a positive coupling between the three.

Figure 2 further illustrates the boundary error heatmap of typical cases. When refining without boundaries, the high error areas are distributed in a circular pattern along the outer contour of ET, concentrated in the transition zone between the tumor and surrounding edema; After introducing boundary refinement, the high error loop of the same case significantly narrows, and the overall color shifts towards blue. Improvement is not evenly distributed across the entire tumor surface, but rather concentrated in the invasive boundary segment.

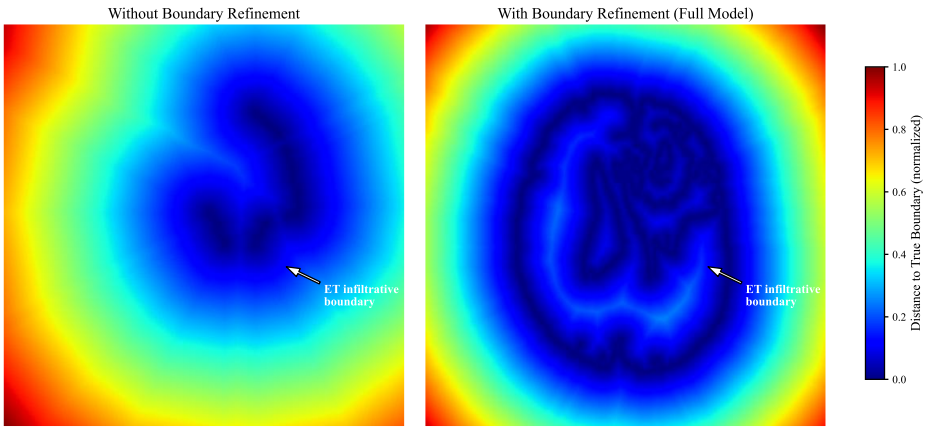

Figure 2. Typical case boundary error heatmap

To verify whether SE fusion truly learns data-driven modal selection rather than degenerating into equal weight averaging, Figure 3 presents the average SE weights obtained from four MRI sequences at three scales of the encoder in the form of a bar chart. T1CE has the highest weight at shallow and mid-level scales, FLAIR's weight increases and approaches T1CE at deeper scales, while the weights of T1 and T2 remain relatively stable at lower levels. The shallow and mid-level features mainly carry enhanced local contrast information of the tumor, and T1CE has the strongest discriminative power at this scale; The aggregation of deep features leads to a larger receptive field, and FLAIR's advantage in characterizing edema areas and the overall extent of tumors begins to play a role. This result indicates that the SE module has formed a dynamic modal selection in the scale dimension, dominated by T1CE locally and transitioning to FLAIR globally, rather than simply assigning fixed high weights to a sequence.

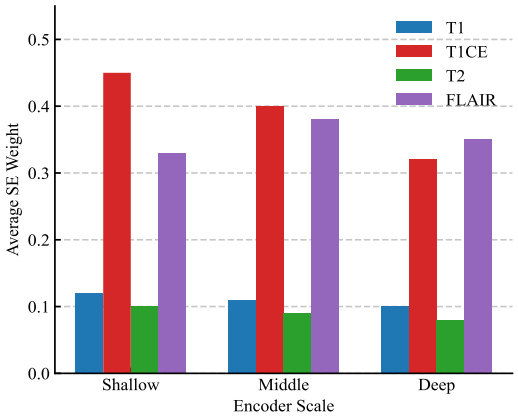

Figure 3. The average SE attention weight of four MRI sequences at each encoder scale

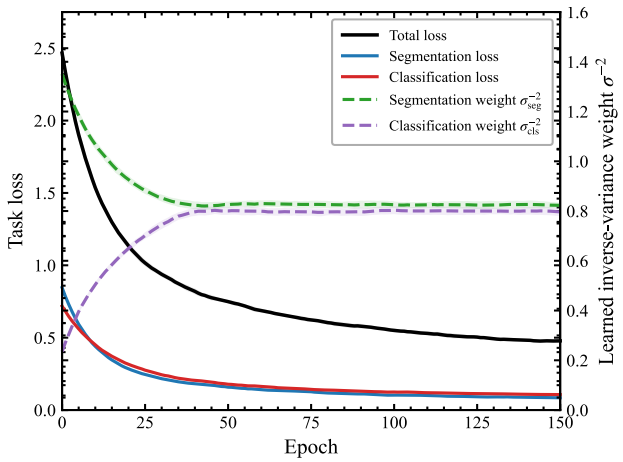

Figure 4. The variation curve of training loss and learnable task weights with epoch

Figure 4 shows the total loss and training curves for each task component. The segmentation loss rapidly decreases in the first 30 epochs and then enters slow convergence, while the classification loss tends to stabilize after about 50 epochs; The uncertainty weights σ_{seg}^{-2} and σ_{cls}^{-2} stabilized at similar levels after undergoing adjustments before and during training, indicating that the uncertainty weighting mechanism completed the automatic calibration of effective weights between tasks within a limited number of training steps, and there was no degradation pattern where the weight of one task continued to increase while the weight of another task was completely suppressed.

4. Conclusion

Comparative experiments show that the joint framework outperforms single task and serial schemes in both classification and segmentation. The ablation experiment decomposed the gains of Dice and HD95 into three sources: SE fusion mode selection, uncertainty weighted gradient equalization, and boundary refinement contour constraint. The ET subregion was used as the stress test point to verify the target of each module. Visual analysis provides mechanistic evidence from two perspectives: the spatial distribution of boundary errors and the scale evolution of SE weights.

References

[1] Chen S, Zhong Z, Chen Y, et al. Prediction of lymphovascular invasion in invasive breast cancer via intratumoral and peritumoral multiparametric magnetic resonance imaging machine learning-based radiomics with Shapley additive explanations interpretability analysis. [J]. Quantitative imaging in medicine and surgery, 2025, 15(9): 7833-7846. DOI: 10.21037/QIMS-2024-2685.

[2] Wenbin B, Yangwenyi L, Jincui Y, et al. The diagnostic value of biparametric magnetic resonance imaging vesical imaging-reporting and data system combined with tumor contact length for assessing muscle invasion in bladder cancer lesions scored vesical imaging-reporting and data system 3: A single-center retrospective analysis [J]. Indian Journal of Urology, 2026, 42(3): 225-231. DOI:10.4103/IJU.IJU_128_26.

[3] Hamza A, Damaševičius R. Deep Learning for Brain Tumor Segmentation and Classification: A Systematic Review of Methods and Trends [J]. Computers, Materials & Continua, 2025,86(1):1-41. DOI:10.32604/CMC.2025.069721.

[4] Bindemann M, Schanze T. Log-Polar Image Mapping for CNN-independent Rotation and Scale Invariant Brain Tumor Classification [J]. Current Directions in Biomedical Engineering, 2026, 12(1): 41-44. DOI:10.1515/CDBME-2026-0111.

[5] Nefediev A N, Staroverov E N, Davydov V R. Application of Convolutional Neural Networks to Improve Tumor Segmentation in Prostate MRI [J]. Bulletin of the Russian Academy of Sciences: Physics, 2026, 90(9): 1531-1538. DOI:10.1134/S1062873826701492.

[6] Tlahig S, Bindemann M, Schanze T. 2.5D Multimodal MRI Brain Tumor Segmentation: Beyond 2D Deep Learning Approaches [J]. Current Directions in Biomedical Engineering, 2026, 12(1): 580-583. DOI: 10.1515/CDBME-2026-0247.

[7] Kang H ,Elham E ,Wang Z , et al.Machine learning pipeline for automated segmentation and classification of complicated cystic renal masses on MRI: comparison with radiologists' assessments [J]. Japanese Journal of Radiology, 2025, 44(3): 1-12. DOI: 10.1007/S11604-025-01912-4.

[8] Jin S, Fayaz M, Dang M L, et al. SwinHCAD: A Robust Multi-Modality Segmentation Model for Brain Tumors Using Transformer and Channel-Wise Attention [J]. Computers, Materials & Continua, 2025, 86(1): 1-23. DOI:10.32604/CMC.2025.070667.

[9] Pacal I. Modern Vision Transformer Models for Accurate and Explainable Brain Tumor MRI Classification [J]. OBM Neurobiology, 2026, 10(3): 1-22. DOI: 10.21926/OBM.NEUROBIOL.2603352.

[10] Salam A, Abrar M, Anwer W R, et al. Channel-Attention DenseNet with Dilated Convolutions for MRI Brain Tumor Classification [J]. Computer Modeling in Engineering & Sciences, 2025, 145(2): 2457-2479. DOI: 10.32604/CMES.2025.072765.