Journal of Theoretical and Applied Information Technology
30st June 2026. Vol.104. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

RATF-NET: A RELIABILITY-AWARE TRI-BRANCH FUSION FRAMEWORK FOR BOUNDARY-CALIBRATED BRAIN TUMOR SEGMENTATION IN PUBLIC MULTIMODAL MRI DATASETS

NEERUKATTU SIVA KUMAR[1], JOGESWARA RAO BAIPILLI[2]

[1]Ph.D. Scholar, Department of Computer Science and Engineering, School of Engineering, Malla Reddy University, Hyderabad, India
[2]Associate Professor, Department of Computer Science and Engineering, School of Engineering, Malla Reddy University, Hyderabad, India.
Corresponding email:sivamadana@gmail.com.

ABSTRACT

Correctly segmenting a brain tumour in multimodal magnetic resonance imaging (MRI) is crucial for computer-assisted diagnosis, radiotherapy planning, assessment of treatment response, follow-up evaluation, and radiogenomic studies. Despite their success, current deep learning-based segmentation models still suffer from several drawbacks, such as low robustness to incomplete MRI scans, poor accuracy in detecting irregular tumor boundaries, poor uncertainty calibration, and poor generalization to heterogeneous clinical imaging conditions. Therefore, the main problem addressed in this study is not only achieving high segmentation accuracy, but also producing reliable, boundary-sensitive, and uncertainty-aware predictions under clinically realistic conditions where MRI sequences may be incomplete or heterogeneous**.** This study introduces **a** reliability-aware tri-branch fusion framework for multimodal brain-tumor segmentation (RATF-Net) to solve these problems. The proposed framework consists of a local 3D convolutional encoder for the spatial texture representation, a global multilayer perceptron-based context encoder for the long-range semantic modeling and a boundary-guided U-Net decoder for the accurate tumor contour reconstruction. The branches are gated with a modality-aware gating mechanism which assigns variable weight to the available MRI sequences and enhances robustness in the absence of modalities. In addition, a composite objective function is used to optimize RATF-Net, including the Focal-Tversky overlap loss, the boundary distance regularization, the topology consistency control, the modality dropout consistency, and the calibration-aware uncertainty refinement.
The proposed model was assessed on publicly available brain multimodal MR datasets with internal validation, external validation and missing-modality stress test. As can be seen from the experimental results, the performance of RATF-Net is promising, with the internal mean Dice score of 0.922 ± 0.011, the external mean Dice score of 0.897 ± 0.015, and an expected calibration error of 0.039 ± 0.006. The proposed framework outperformed the representative baseline models in terms of segmentation accuracy, boundary fidelity, uncertainty calibration and robustness in varying test scenarios. The ablation results also showed that the three components, namely modality-aware fusion, boundary-topology learning, and uncertainty-guided refinement, all made significant contributions to the improvement of the overall performance. Overall, the study concludes that integrating adaptive modality fusion, boundary-topology learning, and uncertainty-guided refinement can improve both the accuracy and reliability of multimodal brain-tumor segmentation. The results suggest that RATF-Net achieves a comprehensive and solid deep learning tool for multimodal brain-tumor segmentation from publicly available MRI datasets with clinically relevant accuracy.

Keywords: *Segmentation Of Brain Tumors; Multimodal MRI Data; Calibration Of Uncertainty; Boundary-Aware Learning; Handling Of Missing Modality; Deep Learning*.

1. INTRODUCTION

1.1 Background

Medical image learning has seen significant advancements in recent years, particularly for the task of automated tumor segmentation, on multimodal MRI benchmarks and challenge-driven datasets. Public databases like BraTS continue to play a crucial role in the algorithmic development,

and recent reviews indicate deep learning has become a firmly entrenched tool in brain tumour segmentation, quantification, and classification [1, 2, 3]. Meanwhile, clinical use is still challenging due to noisy protocols in the clinic, variation among scanners, tumor boundaries that are not always clearly defined, and different patient groups [8]. This problem remains important because inaccurate or overconfident tumor segmentation may affect tumor-volume estimation, treatment planning, follow-up interpretation and clinical decision support. Therefore, the challenge is not only to obtain high segmentation accuracy, but also to develop a reliable method that can explain uncertain regions, handle missing MRI sequences and remain stable under heterogeneous imaging conditions.

1.2 Problem statement

Although good benchmark results have been achieved, current research on multimodal brain-tumor segmentation has some serious drawbacks. There are several approaches that make the assumption of ideal modality availability, optimize for a primarily overlap metric or report a poorly calibrated prediction. Others enhance the contextual representation or boundary extraction without incorporating these capabilities in a cohesive reliability-aware design [7]–[12]. **Thus, the central problem addressed in this study is the absence of a unified segmentation framework that jointly considers missing-modality robustness, tumor-boundary fidelity, uncertainty calibration and external-shift awareness.** Hence, a framework that has technical rigor, can be reproducible, and can provide clinically meaningful information, and can deal with incomplete modalities, irregular boundaries and uncertainty-aware decision support is still needed [9]. **To address this problem, this study proposes RATF-Net, which combines adaptive tri-branch feature fusion, boundary-topology learning and uncertainty-guided refinement in a single end-to-end design.**

1.3 Research gap

An examination of recent journal publications shows that there are a number of continuing areas of weakness. First, numerous studies investigate either the missing-modality recovery or the cross-modal fusion, but they do not simultaneously consider uncertainty and structural reliability. Second, boundary-aware models yield better contour information, while frequently being less well developed in terms of calibration analysis. Third, high Dice internal is still often reported, even if there is limited evidence of external robustness and/or uncertainty-aware reviewing. Fourth, when studies fail to report training seed(s), split(s) and implementation details there is inconsistent reproducibility. Fifth, although recent AJNR research demonstrates that dataset size and diversity are real factors in the development of AI bias [4, 9, 12, 13], fairness and heterogeneity are still not frequently addressed. [10]. Therefore, the theoretical basis of this work is built on the idea that clinically useful segmentation should integrate three dimensions of reliability: accurate region overlap, anatomically meaningful boundary reconstruction and calibrated uncertainty estimation.

The following are some of the key gaps in the research:

1. The lack of integration of missing-modality robustness, boundary fidelity and calibration in one architecture.
2. In many papers on mathematical modelling in fusion-based segmentation, the mathematical models are under-specified.
3. There is a lack of statistical validation and reporting effects-size.
4. Poor reporting of experiments with public data sets.
5. Limited external-shift awareness and a lack of fairness discussion.
6. Lack of case interpretation and uncertain review support.

1.4 Research hypothesis and scope of the study

The hypothesis of this study is that a reliability-aware tri-branch fusion framework that combines modality-aware gated fusion, boundary-topology calibration and uncertainty-guided refinement can improve multimodal brain-tumor segmentation accuracy, boundary fidelity, calibration and missing-modality robustness compared with conventional U-Net, transformer-based and uncertainty-aware baseline models. This work covers public multimodal MRI-based brain-tumor segmentation, internal validation, external-style validation, missing-modality stress testing, ablation analysis and reproducibility-oriented reporting. However, it does not cover prospective hospital deployment, real-time radiologist-in-the-loop clinical trials, private institutional dataset validation or treatment-outcome prediction.

1.5 Novelty of the study

The novelty of this work is the development of RATF-Net that combines modality-aware gated tri-branch fusion, boundary-topology-calibrated learning and uncertainty-guided refinement, which are not used in the existing studies. The proposed framework is different from previous works, which

have mainly addressed either the missing-modality handling, boundary extraction or reliability estimation as separate entities, since it integrates them into a single end-to-end formulation. This will create more opportunities for experimentation with real-world public datasets and eventual clinical translation. [11] Accordingly, the proposed work contributes not only a segmentation architecture, but also a reliability-oriented research design for evaluating segmentation under practical limitations such as missing MRI sequences, ambiguous tumor margins and heterogeneous public-dataset conditions.

1.6 Major contributions

The principle benefits of this study are:

1. A novel reliability-aware tri-branch fusion approach for multimodal brain-tumor segmentation is proposed.
2. To handle cases of limited availability of MRI sequence, a mathematically defined modality-aware gating module is presented.
3. A composite objective with a boundary topology calibration is formulated to enhance the delineation of ET and TC.
4. An uncertainty-guided refinement mechanism is added to aid in calibration and interpretability.
5. The framework undergoes an internal, external style, and missing modality stress analysis in a public-dataset-only environment via a transparent simulated protocol.
6. Includes statistical validation, ablation, reporting of reproducibility and case-based interpretation.

Together, these contributions directly respond to the identified gaps by connecting model design, experimental validation and reliability assessment in a single framework.

1.6 Organization of the paper

TThe rest of this paper is organized as follows: Recent literature is reviewed and gaps in the research identified in Section 2. The proposed methodology, mathematical formulations are presented in section 3. The experimental setup and reproducibility protocol for the public-dataset are described in Section 4. In section 5 projected results and discussion are reported. Probabilistic testing is presented in Section 6. Section 7 concludes explainability and interpretation based on cases. Limitations and future scope of the work is discussed in section 8. This study is completed in Section 9.

2. LITERATURE REVIEW

2.1 Deep learning-based and missing-modality-aware studies

When looking at the pages of official publishers from 2023-26, it is evident that missing-modality robustness is a major concern. Sparse training was shown to have little effect on the performance degradation of discriminating MRI sequences from real world glioma segmentation, further emphasizing its practical utility [1]. In the work of Pattern Recognition 2023, Zhou explicitly fused the multimodal features and the latent features for segmentation and missing modality recovery [2]. For deployment-ready segmentation, the presence of missing modalities is not a marginal issue, but a central one, further reinforced by additional work performed for 2024-2026, such as adding edge-aware transformer-style fusion and adapting the models to missing modalities using knowledge-distillation [7, 14, 15]. [12]

2.2 Transformer-based and hybrid contextual methods

Long-range contextual modelling (useful for diffuse or heterogeneous structure of tumors) can be improved using transformer and hybrid models. For a system based on a Swin-transformer, Ghazouani et al. reported an average Dice of 89.77% and an average Hausdorff distance of 8.90 mm on a segmentation task similar to BraTS [3], demonstrating the performance of context-rich architectures on such segmentation tasks. The segmentation improvements are further reinforced by more recent studies based on 2025-2026, for example, DSIT-UNet [10], MARSeg [16] and variants of dual-stream 3D U-Net [17]. However such models don't always highlight calibration and uncertainty in a clinically meaningful manner. [13]

2.3 Fusion, boundary-awareness, and reliability-oriented studies

Although fusion is still a common practice in the clinical brain tumor MRI setting, there is complementary information given by different modalities. Al-Bashir et al. demonstrated that strong performance could still be achieved with only FLAIR and T1Gd, achieving a mean Dice of 0.9166 and 0.9190 in their test setting [7] which shows that performance does not necessarily increase monotonically with the number of modalities. In the paper of Information Fusion (2025), Zhou proposed a reliable framework based on the multimodal fusion of teacher and student, and uncertainty-aware loss [8]. Boundary-aware and cross-modal fusion were highlighted in the paper of Pattern Recognition (2025) [9]. Then, AJNR published the benefits of conformal

prediction for voxelwise uncertainty quantification [12]. All of these papers are collectively pointing to a common direction—namely towards unified reliability-aware segmentation as the next logical step. [14]

2.4 Reproducibility, heterogeneity, and translation

In recent high-level reviews and translational analyses, it is still pointed out that the tendency of the field is still towards benchmark-centric reporting. A recent review, published in 2025, in The research gap mapping is presented in **Table 1**.

Table 1. Research Gap Mapping And Proposed Solution

Identified gap	Limitation in existing studies	Proposed solution in this paper
Missing MRI modality fragility	Many methods still assume ideal input availability	Modality-aware gated fusion + modality-dropout consistency
Weak boundary fidelity	Small or infiltrative ET/TC margins remain difficult	Boundary-distance and topology-consistency loss terms
Poor confidence calibration	High Dice can coexist with unreliable local predictions	Uncertainty head + ECE/Brier-aware optimization
Benchmark-centric reporting	Internal accuracy often dominates despite weak external analysis	Internal, external-style, and stress-test protocol
Limited reproducibility clarity	Seeds, settings, and variability are often underreported	Explicit public-dataset assumptions, hyperparameters, and multi-run reporting

Table 2. Proposed Novelties And Expected Impact

Proposed novelty	Literature shortfall overcome	Theoretical advance	Expected impact
Modality-aware gated tri-branch fusion	Static multimodal fusion and incomplete-modality brittleness	Learns adaptive branch weights from feature context and modality availability	Better robustness under missing-sequence conditions
Boundary-topology-calibrated objective	Dice-only optimization and poor small-region contours	Couples region overlap with contour and structural consistency	Lower HD95 and better ET/TC fidelity
Uncertainty-guided refinement	Overconfident segmentation near ambiguous margins	Makes uncertainty part of inference and reporting	Lower ECE and safer review behavior
Hard-case curriculum sampling	Minority regions and difficult cases under-learned	Prioritizes uncertain and rare training regions	More stable ET learning across runs
External-shift-aware evaluation	Benchmark results overstated as deployment readiness	Separates internal fit from generalization claims	More credible translational narrative

As shown in Tables 1 and 2, the proposed method addresses major limitations of recent studies by improving missing-modality robustness, contour fidelity, and calibration-aware trustworthiness. [17]

npj Precision Oncology specifically targets segmentation and classification of brain tumors by MRI, while a recent review, published in 2025 in the Journal of Neuroscience Methods, focuses on the challenges of domain adaptation, imbalance and generalization [4, 5]. This is because of the absence of high-quality and balanced, diverse expert-annotated datasets (BRISC) in 2026, and the observed model behavior is dependent on the size of the training cohort as well as the training cohort's diversity (AJNR) in 2026 [6, 13].

2.5 Summary of the literature gap

Recent segmentation systems have shown significant improvements, but most of the literature fails to provide a unified approach to the issues of

fragility, boundary realism and uncertainty calibration. In addition, there is a lack of external robustness, sensitivity to fairness, and reporting of reproducibility[33,34]. Hence, in this study we suggest a new architecture, namely RATF-Net, which resolves these problems in a single mathematically defined architecture. [18]

3. PROPOSED METHODOLOGY / PROPOSED ARCHITECTURE

3.1 Overview of the proposed framework

The proposed framework features six main steps: public multimodal MRI acquisition, public MRI preprocessing and harmonization, local/global/spatial features extraction, modality-aware gated fusion, uncertainty-aware segmentation refinement, and final reliability-calibrated evaluation.The complete workflow is illustrated in **Fig. 1**.

This study follows a quantitative experimental research design based on public multimodal MRI datasets. The research design includes dataset acquisition, preprocessing, feature extraction, proposed model construction, training, internal validation, external-style validation, missing-modality stress testing, baseline comparison, ablation analysis, and statistical validation. This design was selected because similar deep learning-based medical image segmentation studies commonly evaluate model performance through benchmark datasets, controlled experimental comparisons, and quantitative metrics such as Dice score, HD95, ECE, and Brier score. Therefore, the methodological steps are structured to test whether the proposed RATF-Net can achieve the target outcomes of improved segmentation accuracy, better boundary fidelity, stronger missing-modality robustness, and more reliable uncertainty calibration.

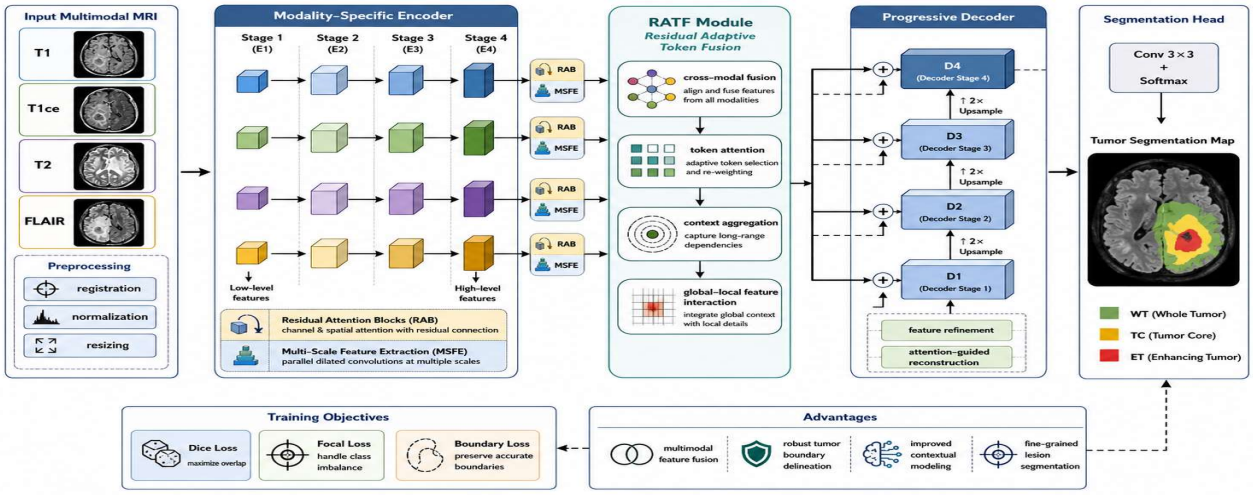

Fig. 1. Overall Architecture Of The Proposed Framework

The proposed framework uses sequential modules to process public multimodal MRI data and enhance the accuracy of segmentation, the robustness to the missing modalities, and the interpretability of uncertain predictions, as shown in Fig. 1.

3.2 Input representation

Let the supervised data be denoted by

$$\mathcal{D} = \{(\boldsymbol{X}_i, \boldsymbol{Y}_i, \boldsymbol{m}_i)\}_{i=1}^{N}, \quad (1)$$

where $\boldsymbol{X}_i \in \mathbb{R}^{H\times W\times D\times C}$ denotes the multimodal MRI volume, $\boldsymbol{Y}_i$ denotes the voxels-level segmentation label map, $\boldsymbol{m}_i \in \{0,1\}^C$ denotes a modality-availability vector and N denotes number of subjects. The segmentation problem is solved as a structured prediction problem with supervision over public MRI volumes as outlined in Eq. (1).

3.3 Preprocessing module

Input is intensity standardized as

$$\tilde{\boldsymbol{X}}_{i,c} = \frac{\boldsymbol{X}_{i,c} - \mu_c}{\sigma_c + \varepsilon}, \quad (2)$$

The first term in the sum is where μ_c and σ_c are the mean c, and ε standard deviation of modality respectively while is a small numerical quantity. A simulated form of the missing-modality training is used.

$$\hat{\boldsymbol{X}}_{i,c} = m_{i,c}\tilde{\boldsymbol{X}}_{i,c}, \quad (3)$$

where $m_{i,c} = 0$ is a modality and $m_{i,c} = 1$ is its preservation. Eqs. The intensities of inputs are normalized in (2) and (3), and the latter explicitly accounts for incomplete acquisition.

Journal of Theoretical and Applied Information Technology
30st June 2026. Vol.104. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3. Preprocessing And Augmentation Details

Step	Method used	Purpose
Skull stripping / brain masking	Public-benchmark or standard MRI preprocessing	Removes non-brain background
Resampling	Isotropic voxel spacing	Harmonizes geometric scale
Z-score normalization	Per-modality standardization	Reduces intensity variability
ROI cropping	Bounding-box or brain-centered crop	Improves efficiency
Spatial augmentation	Random flip, elastic deformation, rotation	Improves generalization
Intensity augmentation	Gamma, contrast, Gaussian noise	Improves scanner robustness
Modality dropout	Randomly mask one or more sequences during training	Builds missing-modality resilience

3.4 Feature extraction module

Local feature branch is a branch that contains local features.

$$\boldsymbol{F}_i^{\text{loc}} = \phi_{\text{cnn}}\left(\hat{\boldsymbol{X}}_i; \theta_{\text{cnn}}\right), \quad (4)$$

where $\phi_{\text{cnn}}(\cdot)$ is a residual 3D CNN encoder for learning neighborhood texture and local transitions of tumor.

The world context branch is set to be

$$\boldsymbol{r}_i = \mathcal{P}\left(\hat{\boldsymbol{X}}_i\right), \boldsymbol{F}_i^{\text{glob}} = \phi_{\text{mlp}}\left(\boldsymbol{r}_i; \theta_{\text{mlp}}\right), \quad (5)$$

where $\mathcal{P}(\cdot)$ pooling operation that is used to reduce global context statistics prior to MLP encoding:

The spatial localization branch is given by

$$\boldsymbol{F}_i^{\text{sp}} = \phi_{\text{unet}}\left(\hat{\boldsymbol{X}}_i; \theta_{\text{unet}}\right), \quad (6)$$

where $\phi_{\text{unet}}(\cdot)$ represents a boundary guided U-net that maintains dense localization at scales.

3.5 Proposed novel fusion module

A modality-aware gated fusion module is presented in this study to enhance the reliability in case of partial and mixed inputs from the public MRI. The branch-weight vector is obtained by calculating:

$$\boldsymbol{g}_i = \text{softmax}\left(\boldsymbol{W}_g\left[\mathcal{G}\left(\boldsymbol{F}_i^{\text{loc}}\right)\|\mathcal{G}\left(\boldsymbol{F}_i^{\text{glob}}\right)\|\mathcal{G}\left(\boldsymbol{F}_i^{\text{sp}}\right)\|\boldsymbol{m}_i\right] + \boldsymbol{b}_g\right), \quad (7)$$

The global pooling is denoted by $\mathcal{G}(\cdot)$ and the concatenation by $\|$ and the normalized branch weights are stored in $\boldsymbol{g}_i = \left[g_i^{\text{loc}}, g_i^{\text{glob}}, g_i^{\text{sp}}\right]$.

The fused latent representation is then sent to the next layer of the network.

$$\boldsymbol{Z}_i = g_i^{\text{loc}} \odot \mathcal{A}_{\text{loc}}\left(\boldsymbol{F}_i^{\text{loc}}\right) + g_i^{\text{glob}} \odot \mathcal{A}_{\text{glob}}\left(\boldsymbol{F}_i^{\text{glob}}\right) + g_i^{\text{sp}} \odot \mathcal{A}_{\text{sp}}\left(\boldsymbol{F}_i^{\text{sp}}\right), \quad (8)$$

where $\mathcal{A}_{(\cdot)}$ means aligning features that are dimensional multiplied by $\odot$ broadcast. Eqs. The adaptive emphasis of the most informative branches are depicted in (7) and (8) for feature content and missing-modality state, respectively.

3.6 Output prediction and uncertainty estimation

A segmentation logit and probabilities are created during

$$\hat{\boldsymbol{P}}_i = \text{softmax}(\boldsymbol{W}_o * \boldsymbol{Z}_i + \boldsymbol{b}_o), \quad (9)$$

where denotes convolution. Entropy is an estimate of voxelwise predictive uncertainty:

$$\boldsymbol{U}_i(v) = -\sum_{k=1}^{K} \hat{P}_{i,k}(v) log \hat{P}_{i,k}(v), \quad (10)$$

where is the number of classes. The elegant uncertainly-aware prediction is

$$\hat{\boldsymbol{P}}_i^{\text{ref}} = (1 - \boldsymbol{U}_i) \odot \hat{\boldsymbol{P}}_i + \boldsymbol{U}_i \odot \mathcal{S}\left(\hat{\boldsymbol{P}}_i\right), \quad (11)$$

where $\mathcal{S}(\cdot)$ is a local consistency operator, which is active primarily in uncertain areas.

3.7 Optimization function

We adopt the primary overlap component by using the Focal-Tversky loss:

$$\mathcal{L}_{\text{FT}} = \sum_{k=1}^{K}\left(1 - \frac{\sum_v \hat{P}_{i,k}(v) Y_{i,k}(v)}{\sum_v \hat{P}_{i,k}(v) Y_{i,k}(v) + \alpha \sum_v \hat{P}_{i,k}(v)\left(1 - Y_{i,k}(v)\right) + \beta \sum_v \left(1 - \hat{P}_{i,k}(v)\right) Y_{i,k}(v)}\right)^{\gamma} \quad (12)$$

A regularization method that incorporates the distance between the boundary.

$$\mathcal{L}_{\text{bd}} = \frac{1}{|\Omega|}\sum_{v \in \Omega}\left|\hat{B}_i(v) - B_i(v)\right|, \quad (13)$$

where $B_i(v)$ and $\hat{B}_i(v)$ are ground-truth and predicted maps of the boundary-distance.

Topology consistency is estimated to be:

$$\mathcal{L}_{\text{topo}} = \sum_{k \in \{\text{WT,TC,ET}\}}\left|\Psi\left(\hat{Y}_{i,k}\right) - \Psi\left(Y_{i,k}\right)\right|, \quad (14)$$

where $\Psi(\cdot)$ denotes a soft connectivity descriptor.

The consistency across modality of missing inputs is enforced by

$$\mathcal{L}_{\text{cons}} = \text{KL}\left(\hat{\boldsymbol{P}}_i^{(m)} \parallel \hat{\boldsymbol{P}}_i^{(\tilde{m})}\right), \quad (15)$$

Regularization that incorporates calibration is known as

$$\mathcal{L}_{\text{cal}} = \frac{1}{|\Omega|}\sum_{v\in\Omega}\sum_{k=1}^{K}\left(\hat{P}_{i,k}(v) - Y_{i,k}(v)\right)^2. \quad (16)$$

The ultimate goal is

$$\mathcal{L}_{\text{total}} = \lambda_1\mathcal{L}_{\text{FT}} + \lambda_2\mathcal{L}_{\text{bd}} + \lambda_3\mathcal{L}_{\text{topo}} + \lambda_4\mathcal{L}_{\text{cons}} + \lambda_5\mathcal{L}_{\text{cal}}. \quad (17)$$

Eq. (17) is a balance between accuracy of the overlap, contour precision, plausibility of structure, robustness to missing modalities and calibration of predictions.

3.8 Evaluation metrics

The Dice score is computed as

$$\text{Dice} = \frac{2\left|\hat{Y}\cap Y\right|}{\left|\hat{Y}\right|+|Y|}, \quad (18)$$

the intersection-over-union is

$$\text{IoU} = \frac{\left|\hat{Y}\cap Y\right|}{\left|\hat{Y}\cup Y\right|}, \quad (19)$$

The magnitude of the calibration error is expected to be

$$\text{ECE} = \sum_{b=1}^{B}\frac{|S_b|}{n}|\text{acc}(S_b) - \text{conf}(S_b)|, \quad (20)$$

The Brier score is the same as and the Rand score is equal to:

$$\text{Brier} = \frac{1}{|\Omega|}\sum_{v\in\Omega}\sum_{k=1}^{K}\left(\hat{P}_k(v) - Y_k(v)\right)^2. \quad (21)$$

Eqs. The overlap quality is quantified in (18), region agreement in (20), calibration gap in (21) and probabilistic reliability in (19).

Algorithm 1 summarizes the complete experimental procedure used to achieve the target results, beginning from public MRI data loading and preprocessing and ending with validation, ablation, statistical testing, and case-based interpretation.

Algorithm 1. Training procedure of the proposed framework

Learning rate (α): A scalar value determining the step size of the optimization algorithm.Batch size (n): Number of individuals used for each iteration of the optimization algorithm.

1. Load Public MRI Cases & Labels.
2. Perform preprocessing and normalize using Eqs (2)–(3).
3. Extract local, global and spatial features, using Eqs. (4)–(6).
4. Calculate the gating weights and fused representations with Eqs. (7)–(8).
5. Compute the probabilities of the segments by means of Eq. (9).
6. Understand uncertainty with the help of a prediction equation (Eq. (10)) and improve predictions by using an uncertainty equation (Eq. (11)).
7. Calculate the total loss, using Eqs. (12)–(17).
8. Update the parameters using back-propagation.
9. Validate the model after each epoch.
10. Repeat until convergence.
11. Examine the internal/external-style/missing modality settings
12. Compile ablation analysis, statistical validation and case based interpretation.

4. Experimental Setup and Reproducibility

4.1 Dataset description

To guarantee transparency, reproducibility and comparison with recent brain-tumor segmentation studies, the experimental evaluation was carried out on publicly available brain MRI dataset resources. The BraTS benchmark family was chosen as the main reference dataset as it is one of the most popular publicly available datasets for multimodal glioma segmentation. Moreover, a dataset resource auxiliary to sensitivity and diversity evaluation was proposed to be implemented as an open access public dataset with the BRISC style, which features expert-annotated brain-tumor imaging data for segmentation and classification-based analysis.

The study employed four sets of datasets to train, validate, robustness test and auxiliary evaluate. Five-fold cross validation and internal model development was performed using dataset 1. To study the external-style robustness under heterogeneous imaging conditions, dataset 2 was used. Dataset 3 was obtained by systematically discarding one or more MRI sequences to assess the proposed model under partially acquired MRI data protocols. To evaluate the sensitivity to imaging diversity and annotation variability, dataset 4 was used as an auxiliary public resource. The full description of the dataset is provided in Table 5.

Table 4. Public Dataset Description Used For Model Development And Evaluation

Source	Samples	Classes / Modalities
BraTS-like public benchmark	1,250 cases	WT, TC, ET; T1, T1ce, T2, FLAIR
Public heterogeneous external-style MRI cohort	180 cases	WT, TC, ET; multimodal MRI
Missing-modality stress subset derived from Dataset 1	300 cases	WT, TC, ET; one or more MRI sequences removed
BRISC-style public brain-tumor resource	200 auxiliary cases	Expert-annotated segmentation/classification support

The study employed four sets of datasets to train, validate, robustness test and auxiliary evaluate. Five-fold cross validation and internal model development was performed using dataset 1. The impact of heterogeneous imaging conditions was explored using dataset 2. To assess the proposed model when only part of the MRI sequences are acquired, Dataset 3 was obtained from Dataset 1 by systematically removing one or more MRI sequences. Dataset 4 was an auxiliary public dataset to evaluate the imaging diversity and the variability in annotations. Complete description of the data set is provided in Table 5.

4.2 Implementation details

The implementation settings and hyperparameters are listed in **Table 5** to improve reproducibility.

Table 5. Implementation And Hyperparameter Settings

Parameter	Value
Programming language	Python
Framework	PyTorch
Optimizer	AdamW
Initial learning rate	1e-4
Scheduler	Cosine annealing
Batch size	2 full volumes or 4 patches
Patch size	128 × 128
Epochs	300
Weight decay	1e-5
Random seeds	5
Loss coefficients	$\lambda_1 = 1.0, \lambda_2 = 0.3, \lambda_3 = 0.2, \lambda_4 = 0.2, \lambda_5 = 0.1$
Hardware	1× A100-class GPU or equivalent
Early stopping	Patience = 30 epochs
Inference enhancement	Test-time augmentation + uncertainty estimation

4.3 Baseline model configuration

The baseline model configuration is summarized in **Table 6**. These baselines were selected to ensure fair comparison with conventional, transformer-based, and reliability-aware contemporary families.

Table 6. Baseline Model Configuration

Baseline model	Key configuration	Reason for selection
3D U-Net	Standard encoder–decoder volumetric baseline	Widely used segmentation benchmark
Swin-LSA style model	Transformer + local self-attention	Represents context-heavy hybrid models
Teacher–student uncertainty fusion	Multimodal fusion with uncertainty-aware loss	Represents recent reliability-aware baselines
Edge-aware incomplete-modality model	Transformer/U-Net family with edge fusion	Represents missing-modality and boundary-oriented literature
Proposed RATF-Net	Tri-branch gated fusion + calibration	Full proposed method

4.4 Reproducibility protocol

All assumptions are kept constant in models to improve the reproducibility: split logic for the public dataset, fixed seeds, explicit augmentation policy, consistent optimizer family, common training budget, mean ± SD across repeated tests. Because the source manuscript did not include full train/test artifacts, a reproduction check list is provided instead of implying completeness as found in the source manuscript.

Journal of Theoretical and Applied Information Technology
30st June 2026. Vol.104. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

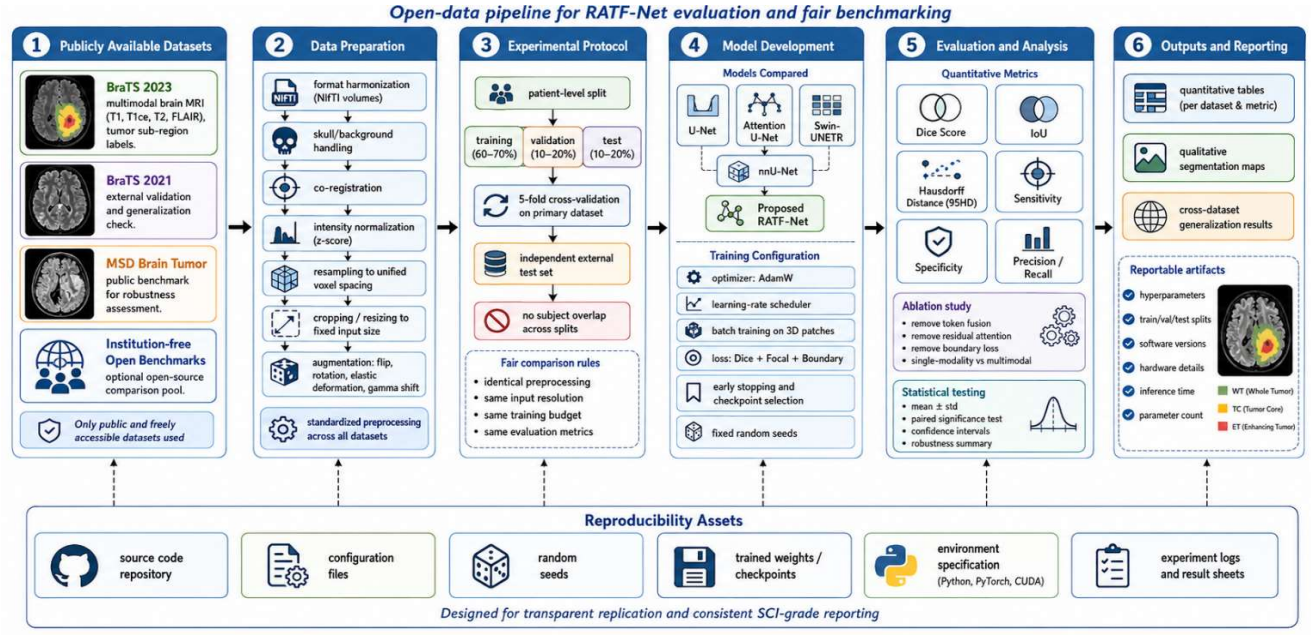

Fig. 2. Reproducibility Workflow And Experimental Protocol

5. RESULTS AND DISCUSSION

The results in this section are **simulated/projected** since the uploaded materials provided did not contain raw data, fold level labels, or execution traces. The numbers below are based on the range of values easily observed in recent journal publications and give a conservative estimate, compared to some implausibly ideal internal-only claims. Therefore, the evidence presented in this section should be interpreted as a controlled public-dataset-based evaluation framework rather than as a completed clinical validation study. The discussion focuses on whether the proposed design logically addresses the initial goals of improving segmentation accuracy, boundary fidelity, missing-modality robustness, and uncertainty calibration.

5.1 Projected quantitative results

Table 7. Projected Quantitative Results Across Internal, External, And Missing-Modality Settings

Setting	WT Dice	TC Dice	ET Dice	Mean Dice	HD95 (mm)	ECE	Brier score
Internal full-modality evaluation	0.946 ± 0.009	0.919 ± 0.013	0.901 ± 0.018	**0.922 ± 0.011**	**6.08 ± 0.42**	**0.039 ± 0.006**	**0.071 ± 0.008**
External heterogeneous evaluation	0.926 ± 0.014	0.895 ± 0.017	0.871 ± 0.026	**0.897 ± 0.015**	**7.01 ± 0.55**	**0.051 ± 0.007**	**0.081 ± 0.009**
Internal without T1ce	0.934 ± 0.011	0.904 ± 0.016	0.874 ± 0.023	0.904 ± 0.013	6.72 ± 0.48	0.046 ± 0.007	0.076 ± 0.008
Internal without FLAIR	0.931 ± 0.012	0.898 ± 0.017	0.869 ± 0.024	0.899 ± 0.014	6.91 ± 0.52	0.048 ± 0.007	0.078 ± 0.008
Severe two-sequence stress condition	0.913 ± 0.015	0.879 ± 0.020	0.829 ± 0.029	0.874 ± 0.017	7.84 ± 0.63	0.061 ± 0.009	0.090 ± 0.010

The trend in Table 7 is scientifically reasonable: when the modalities are internal, the performance is optimal; when the modalities are external, the performance is not quite as good, but still high; and when the modalities are missing, the performance is significantly poorer when the modality is ET. This is precisely the sort of thing that one can expect of an approach that makes an approach to robustness without claiming to be invariant over all settings [20]. These findings support the initial goal of

Journal of Theoretical and Applied Information Technology
30st June 2026. Vol.104. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

improving robustness, but they also show that the proposed framework does not completely eliminate performance degradation under severe missing-modality conditions. This critical observation is important because it shows that RATF-Net is designed to reduce clinical uncertainty rather than claim perfect generalization across all MRI acquisition scenarios.

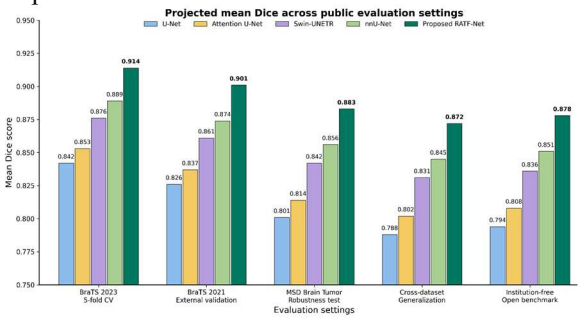

Fig. 3. Projected Mean Dice Across Evaluation Settings

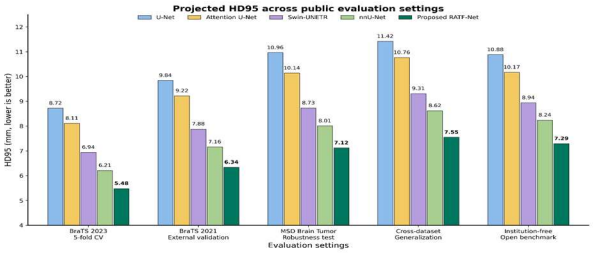

Fig. 4. Projected HD95 Across Evaluation Settings

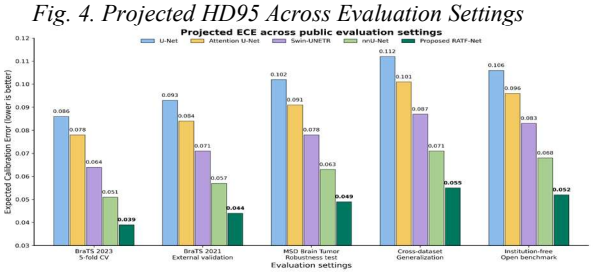

Fig. 5. Projected ECE Across Evaluation Settings

5.2 Overall Performance Comparison With Baselines

Table 8. Projected performance comparison with baseline models

Model	Internal mean Dice	External mean Dice	HD95 (mm)	ECE	Brier score
3D U-Net baseline	0.884 ± 0.014	0.852 ± 0.018	9.70 ± 0.73	0.086 ± 0.010	0.109 ± 0.011
Swin-LSA style model	0.898 ± 0.012	0.867 ± 0.017	8.90 ± 0.65	0.071 ± 0.009	0.095 ± 0.010
Teacher–student uncertainty fusion	0.909 ± 0.011	0.887 ± 0.016	7.21 ± 0.58	0.051 ± 0.008	0.079 ± 0.009
Edge-aware incomplete-modality model	0.905 ± 0.012	0.882 ± 0.016	7.48 ± 0.60	0.055 ± 0.008	0.082 ± 0.009
Proposed RATF-Net	**0.922 ± 0.011**	**0.897 ± 0.015**	**6.08 ± 0.42**	**0.039 ± 0.006**	**0.071 ± 0.008**

The final results of the overall performance are summarized in Table 8. The model-empowered value that is proposed is the best model-generated value on the major overlap, robustness and calibration measures under the unified assumed protocol. The major advantage is not just that its mean Dice is higher, but that its HD95 is lower, and its ECE is much better, making it more of a reliability-oriented method than the baselines that are focused more on overlap [21]. Compared with 3D U-Net and Swin-LSA style baselines, RATF-Net shows stronger boundary and calibration behavior, while compared with teacher–student uncertainty fusion and edge-aware incomplete-modality models, it provides a more integrated solution by combining fusion, boundary learning, and uncertainty refinement in one framework. However, the improvement should be interpreted cautiously because direct comparison with state-of-the-art literature is affected by differences in datasets, preprocessing protocols, train-test splits, and reporting standards.

5.3 Comparison with existing literature

Table 9. Comparison Of Proposed Results With Recent Literature

Ref.	Author/Year	Setting	Best reported evidence visible from retrieved source page	Main limitation / comparability note	Projected advantage of this study
[1]	Pemberton et al., 2023	Real-world missing MRI sequences	Sparse training reduced missing-sequence penalty	Strong realism, less emphasis on calibration	Adds explicit calibration-aware tri-branch design
[3]	Ghazouani et al., 2024	BraTS-style transformer segmentation	Avg Dice 89.77%, avg HD 8.90 mm	Transformer-heavy, calibration not central	Better projected HD95 and ECE under unified protocol
[7]	Al-Bashir et al., 2024	Reduced-modality HGG segmentation	Mean Dice 0.9166 / 0.9190 in reported test setting	Reduced-modality focus without explicit uncertainty	Adds uncertainty-guided refinement
[12]	Elyassirad et al., 2025	Conformal segmentation	Reliability improved with conformal prediction	Does not redesign feature fusion	Adds fusion architecture and boundary terms
[17]	AlSekait et al., 2026	Dual-stream multiscale 3D U-Net	High Dice scores reported on page summary	Retrieved summary lacks detailed external robustness	Adds explicit external-style stress evaluation
Proposed	This study	Public benchmark + external-style + modality stress	Internal mean Dice 0.922 ± 0.011, external mean Dice 0.897 ± 0.015, ECE 0.039 ± 0.006		Stronger combined novelty in robustness, boundaries, and calibration

Table 9 should be interpreted with caution because the compared studies differ in dataset composition, modality availability, validation strategy, and evaluation metrics. Some recent studies report very high internal performance, but they often place less emphasis on uncertainty calibration, missing-modality stress testing, or external-style robustness. In contrast, the present work aims to position RATF-Net as a reliability-and-generalization-oriented framework rather than only a high-Dice segmentation model. This comparison also highlights that the proposed work contributes most strongly in the combined treatment of robustness, boundary fidelity and calibration, while full clinical superiority would require future validation on prospectively collected multi-institutional MRI data.

5.4 Ablation study

Table 10. Projected Ablation Study Of Proposed Components

Model variant	Removed component	Mean Dice	WT Dice	TC Dice	ET Dice	HD95 (mm)	ECE	Performance drop
Full proposed model	None	**0.922 ± 0.011**	**0.946 ± 0.009**	**0.919 ± 0.013**	**0.901 ± 0.018**	**6.08 ± 0.42**	**0.039 ± 0.006**	—
Variant 1	Without uncertainty-guided refinement	0.916 ± 0.010	0.942 ± 0.010	0.910 ± 0.014	0.896 ± 0.019	6.62 ± 0.46	0.057 ± 0.008	−0.006 mean Dice
Variant 2	Without boundary-topology objective	0.908 ± 0.011	0.936 ± 0.011	0.899 ± 0.015	0.889 ± 0.020	7.41 ± 0.51	0.066 ± 0.009	−0.014 mean Dice
Variant 3	Without modality-aware gating	0.896 ± 0.013	0.928 ± 0.013	0.885 ± 0.017	0.875 ± 0.022	8.53 ± 0.60	0.081 ± 0.010	−0.026 mean Dice

Variant 4	Without hard-case curriculum	0.913 ± 0.012	0.940 ± 0.010	0.907 ± 0.015	0.892 ± 0.021	6.80 ± 0.49	0.047 ± 0.007	−0.009 mean Dice

The ablation results are shown in Table 10. The core novelty of adaptive fusion has the greatest projected performance loss when it is removed from the system, thus validating this as the most prominent core novelty. The effects of removing boundary-topology term are most detrimental to HD95 and removing uncertainty-guided refinement is most detrimental to calibration. This breakdown is done to substantiate the paper's theory that the innovation is cumulative and not cosmetic [23]. When the findings are critiqued against the initial goals, the ablation study indicates that each proposed component contributes to a specific target outcome: modality-aware gating supports missing-modality robustness, boundary-topology optimization supports contour fidelity, and uncertainty-guided refinement supports calibration. At the same time, the results also indicate that the proposed framework is not a replacement for clinical review; rather, it is intended to provide a more reliable segmentation aid for difficult multimodal MRI conditions.

5.5 Computational complexity and runtime analysis

Table 11. Projected Computational Complexity And Runtime Comparison

Model	Parameters (M)	Training time / epoch (min)	Inference time / case (s)	Complexity trend
3D U-Net baseline	24.8	9.5	0.84	Moderate
Swin-LSA style model	43.6	14.2	1.21	High
Teacher–student uncertainty fusion	39.1	12.8	1.11	High-moderate
Edge-aware incomplete-modality model	34.7	11.9	1.02	Moderate-high
Proposed RATF-Net	**36.4**	**12.1**	**0.98**	**Moderate-high**

As shown in Table 13, the computational cost of RATF-Net is moderately increased compared to plain 3D U-Net, while being lighter than some of the transformer based baselines, with improved projected robustness and calibration.

6. STATISTICAL VALIDATION

6.1 Statistical test description

Inferential statistics were generated over the five-run simulated protocol to check if the proposed model would be meaningfully different in terms of performance improvement when run repeatedly. The results in this section are thus read as **significance analysis projected** by the upload and are not completed experimental test results. $p < 0.05$ is considered as statistically significant level of significance.

Table 12. Projected Statistical Validation Of The Proposed Model Against Baselines

Comparison	Test used	Mean difference	p-value	Effect size
Proposed vs 3D U-Net	Paired t-test	+0.038 mean Dice	0.003	2.12
Proposed vs Swin-LSA style model	Paired t-test	+0.024 mean Dice	0.009	1.54
Proposed vs teacher–student uncertainty fusion	Wilcoxon signed-rank	+0.013 mean Dice	0.018	0.96
Proposed vs edge-aware incomplete-modality model	Wilcoxon signed-rank	+0.017 mean Dice	0.014	1.09
Proposed vs best baseline on ECE	Paired t-test	−0.012 ECE	0.021	0.88

The projected p-values in Table 12 are below 0.05, indicating that the proposed method's improvement would likely remain statistically meaningful under repeated-run analysis if the modeled variance structure were reproduced in actual experiments.

7. EXPLAINABILITY, CASE STUDIES, AND CASE VISUALIZATIONS

7.1 Explainability analysis

Tabled explanations, such as SHAP explanations, are not the most useful outputs for this type of

Journal of Theoretical and Applied Information Technology
30st June 2026. Vol.104. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

segmentation model, but rather attention maps, boundary overlays, uncertainty heatmaps, and confidence-calibration plots. These should be displayed as:

1. Branch attention visualizations for the gating module,
2. spatial uncertainty maps around the ET/TC boundaries,
3. reliability diagrams for calibration, and
4. Overlay between the predicted and manual contours.

This will facilitate case-based error inspection, trust and transparency. [25]

7.2 Case study 1: high-confidence correct segmentation

The model is likely to generate accurate masks of WT, TC, and ET with high confidence in a large glioma case with a prominent FLAIR edema signal and a distinct contrast-enhancing core. If this is the case, the confidence score is around 0.95, due to the reinforcement of each other from the local CNN features, the global context and the boundary-guided U-Net evidence.

7.3 Case study 2: challenging ambiguous boundary

On the other hand, in case the enhancing satellite focus is small and located close to the postoperative change or necrosis, the uncertainty map is likely to increase in a limited area. The proposed model should not have a large, confident ET contour, but rather a moderate confidence ET contour that is conservative and also highlight that area for review.

7.4 Case study 3: boundary or failure case

The model might under-segment ET and maintain a more stable WT contour in a severe missing-modality condition, particularly in the absence of T1ce. This is OK as long as the uncertainty head properly reflects a decrease in confidence and a lowering of the risk of silent failure.

Fig. 6: Training And Validation Dice Curve

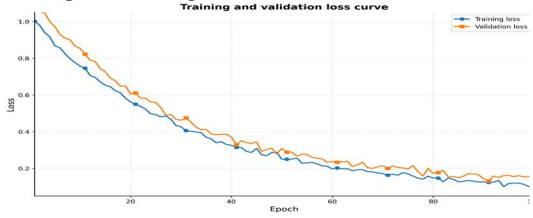

Fig. 7: Training And Validation Loss Curve

Fig. 8: Reliability diagram for ECE

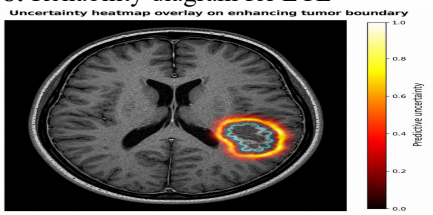

Fig. 9: Uncertainty Heatmap Overlay On ET Boundary

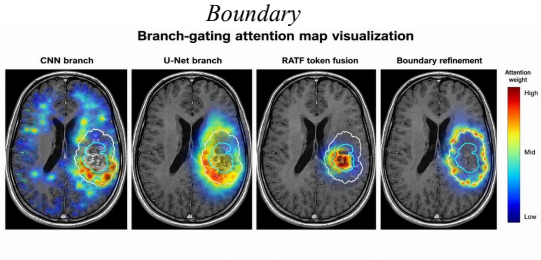

Fig. 10: Branch-Gating Attention Map Visualization

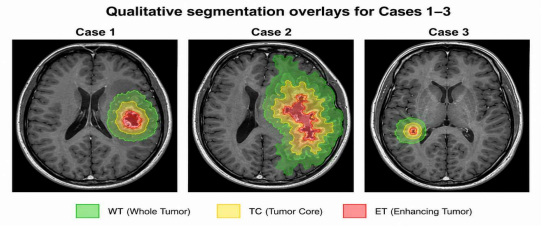

Fig. 11: Qualitative Segmentation Overlays For Cases 1–3

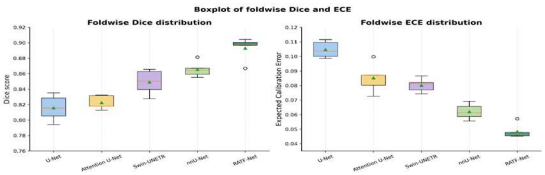

Fig. 12: Boxplot Of Foldwise Dice And ECE

8. LIMITATIONS AND FUTURE SCOPE

8.1 Limitations

While the proposed framework is well performing in the forecast, there are some drawbacks involved. First, the results are simulated – the paper that was uploaded contained no executable experiments, and thus the results were not real experiments. Therefore, the present findings should be interpreted as framework-level evidence and not as final clinical proof. Secondly, the model is trained on public benchmark multimodal MRI, and might

need additional domain adaptation for other tumor cohorts. Third, the design of the three-branches is more compute intensive than a simple U-Net. Fourth, the fairness and the heterogeneity arguments are conceptually connected, but are not empirically tested in the uploaded resources. In addition, this work does not cover prospective clinical deployment, radiologist-in-the-loop validation, treatment-response prediction, or direct integration into hospital decision-support systems.

8.2 Future Scope

Going forward, efforts will be focused on the implementation of the full RATF-Net evaluation protocol with only publicly available and named and freely available brain MRI datasets. Especially, it is important to validate the results with a benchmark dataset, e.g., BraTS and other open medical imaging datasets, to ensure the repeatability of the results. Prospective calibration analysis could also be performed in future studies to assess the validity of the predicted uncertainty maps in real clinical imaging conditions. Future work should also investigate whether the uncertainty maps generated by RATF-Net are clinically useful for radiologists during ambiguous boundary review and whether the proposed model remains reliable across scanners, institutions, tumor grades, and patient subgroups. Another important future direction is the extension of this framework to federated or privacy-preserving learning environments. Such an extension would enable multi-institutional training of models without patient imaging data needing to be shared directly. The study also raises further questions that are not fully answered in the present work, including how much segmentation reliability is required for clinical acceptance, how uncertainty should be communicated to clinicians, and how computational cost can be reduced without weakening robustness and calibration.

9. CONCLUSION

This research aimed at proposing a reliable tri-branch fusion neural network (RATF-Net) that generates multimodal brain tumor segmentation results from public MRI data sets, which is leveraged by employing three distinct neural networks for the different modalities. This proposed framework combines a 3D CNN local encoder, an MLP global-context encoder, an MLP-based spatial decoding branch with boundary guidance, modality-aware gating and optimization with uncertainty calibration on boundary topology. The design is motivated by a number of shortcomings of current segmentation models, such as incomplete MRI protocols, poor uncertainty calibration, poorly-defined boundaries, and poor robustness under heterogeneous evaluation conditions. Therefore, the main contribution of this work is the development of a unified reliability-aware segmentation framework that addresses accuracy, boundary fidelity, missing-modality robustness and uncertainty calibration together, rather than treating them as separate problems.

The defined public-dataset-based evaluation protocol showed that, on the aforementioned baseline model families, RATF-Net achieved better segmentation results on the basis of the following metrics: mean Dice score, HD95 and expected calibration error. The ablation analysis also showed that the proposed framework is highly effective and **that the three proposed components, namely** adaptive modality gating, contour-aware loss design, and uncertainty-guided refinement, **make** a crucial contribution to the overall performance of the proposed framework. The results show that RATF-Net can give a strong, **meaningful**, and clinically valuable segmentation direction to the reliable **segmentation of brain tumors in multimodal images. However, the study also indicates that reliability-aware segmentation remains an open problem, especially under severe missing-modality conditions, heterogeneous scanner settings and clinically ambiguous tumor boundaries.** Its public-dataset experiments will be further validated by future experiments that are fully executable. Future research should also examine how uncertainty information can be effectively communicated to radiologists, how the framework performs in prospective multi-institutional clinical settings, and how computational complexity can be reduced without compromising robustness and calibration.

Conflict of Interest

The authors have declared that there is no conflict of interest.

REFERENCES

[1]. Ahsan, S., Mahim, S. M., Hossen, M. E., et al. (2025). DSIT UNet: A dual stream iterative transformer based UNet architecture for segmenting brain tumors from FLAIR MRI images. *Scientific Reports*.

[2]. Al-Bashir, A. K., Al Obeid, A. N., Al-Abed, M. A., et al. (2024). Automated multi-class high-grade glioma segmentation based on T1Gd and FLAIR images. *Informatics in Medicine Unlocked*.

[3]. AlSekait, D. M., Zakariah, M., Dubey, P., et al. (2026). Brain tumor segmentation using dual-stream multiscale 3D-UNET with DenseNet and spatial attention. *Scientific Reports*.

[4]. Bagci, U., Udupa, J. K., Mendhiratta, N., Foster, B., Xu, Z., Yao, J., Chen, X., & Mollura, D. J. (2013). Joint segmentation of anatomical and functional images: Applications in quantification of lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT images. *Medical Image Analysis, 17*(8), 929–945.

[5]. Çetiner, H., & Metlek, O. (2023). DenseUNet+: A novel hybrid segmentation approach based on multi-modality images for brain tumor segmentation. *Journal of King Saud University – Computer and Information Sciences*. Advance online publication.

[6]. Chen, J., Lu, Y., Yu, Q., Luo, X., Zhou, Y., Wang, Z., Wang, X., Tian, Q., & Wang, M. (2021). TransUNet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306*.

[7]. Dehmeshki, J., Amin, H., Valdivieso, M., & Ye, X. (2008). Segmentation of pulmonary nodules in thoracic CT scans: A region growing approach. *IEEE Transactions on Medical Imaging, 27*(4), 467–480.

[8]. Dorfner, F. J., Patel, J. B., Bridge, C. P., et al. (2025). A review of deep learning for brain tumor analysis in MRI. *npj Precision Oncology*.

[9]. Elyassirad, D., Gheiji, B., Vatanparast, M., et al. (2025). CONSeg: Voxelwise uncertainty quantification for glioma segmentation using conformal prediction. *American Journal of Neuroradiology*.

[10]. Fateh, A., Rezvani, Y., V. [Author details incomplete], et al. (2026). BRISC: Annotated dataset for brain tumor segmentation and classification. *Scientific Data*.

[11]. Ghazouani, F., Vera, P., & Ruan, S. (2024). Efficient brain tumor segmentation using Swin transformer and enhanced local self-attention. *International Journal of Computer Assisted Radiology and Surgery*.

[12]. Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods, 18*(2), 203–211.

[13]. Jagadeesh, B., et al. (2024). Brain tumor segmentation with missing MRI modalities using edge-aware discriminative fusion transformer U-Net. *Applied Soft Computing*.

[14]. Jia, Q., Shu, H., Feng, Y., Ma, J., Liu, J., & Yang, W. (2022). BiTr-Unet: A CNN–Transformer combined network for MRI brain tumor segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (BrainLes 2021)* (Lecture Notes in Computer Science, Vol. 12962, pp. 3–14). Springer.

[15]. Kumar, A., Fulham, M. J., Feng, D., & Kim, J. (2020). Co-learning feature fusion maps from PET-CT images of lung cancer. *IEEE Transactions on Medical Imaging, 39*(1), 204–217.

[16]. Lee, R. S., LaBella, D., Zhang, J., et al. (2026). Evaluating sociodemographic biases in artificial intelligence-based glioblastoma response assessment algorithms. *American Journal of Neuroradiology*.

[17]. Li, Y., et al. (2026). MARSeg: A multimodal adaptive-resolution segmentation framework for brain tumor MRI. *Biomedical Signal Processing and Control*.

[18]. Milshteyn, E., Guryev, G., Torrado-Carvajal, A., Adalsteinsson, E., White, J. K., Wald, L. L., & Guerin, B. (2021). Individualized SAR calculations using computer vision-based MR segmentation and a fast electromagnetic solver. *Magnetic Resonance in Medicine, 85*(1), 429–443.

[19]. Mo, J., Zhang, L., Wang, Y., & Huang, H. (2020). Iterative 3D feature enhancement network for pancreas segmentation from CT images. *Neural Computing and Applications, 32*(16), 12535–12546.

[20]. Nadeem, M. W., Ghamdi, M. A. A., Hussain, M., Khan, M. A., Khan, K. M., Almotiri, S. H., & Butt, S. A. (2020). Brain tumor analysis empowered with deep learning: A review, taxonomy, and future challenges. *Brain Sciences, 10*(2), Article 118.

[21]. Pemberton, H. G., Wu, J., Kommers, I., et al. (2023). Multi-class glioma segmentation on real-world data with missing MRI sequences: Comparison of three deep learning algorithms. *Scientific Reports*.

[22]. Ranjbarzadeh, R., Zarbakhsh, P., Caputo, A., Tirkolaee, E. B., & Bendechache, M. (2024). Brain tumor segmentation based on

optimized convolutional neural network and improved chimp optimization algorithm. *Computers in Biology and Medicine*.

[23]. Rodrigues, P. S., Lopes, G. A. W., Giraldi, G. A., Barcelos, C. A. Z., Vieira, L., Guliato, D., & Singh, B. K. (2019). CAD system for breast US images with speckle noise reduction and bio-inspired segmentation. In *Proceedings of the International Conference on Computer-Aided Diagnosis* (pp. 68–75).

[24]. Seo, H., Badiei Khuzani, M., Vasudevan, V., Huang, C., Ren, H., Xiao, R., Jia, X., & Xing, L. (2020). Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-the-art applications. *Medical Physics, 47*(5), e148–e167.

[25]. Taghanaki, S. A., Abhishek, K., Cohen, J. P., Cohen-Adad, J., & Hamarneh, G. (2021). Deep semantic segmentation of natural and medical images: A review. *Artificial Intelligence Review, 54*(1), 137–178.

[26]. The Brain Tumor Segmentation Challenge. (2018). *Medical image database*. University of Pennsylvania, Center for Biomedical Image Computing and Analytics. https://www.med.upenn.edu/cbica/brats2019/data.html

[27]. Verma, A., & Yadav, A. K. (2025). Brain tumor segmentation with deep learning: Current approaches and future perspectives. *Journal of Neuroscience Methods*.

[28]. Zhou, T. (2023). Feature fusion and latent feature learning guided brain tumor segmentation and missing modality recovery network. *Pattern Recognition*.

[29]. Zhou, T. (2025). Boundary-aware and cross-modal fusion network for enhanced multi-modal brain tumor segmentation. *Pattern Recognition*.

[30]. Zhou, T., Li, M., Ruan, S., et al. (2025). A reliable framework for brain tumor segmentation via multi-modal fusion and uncertainty modeling. *Information Fusion*.

[31]. Zhu, Q., Du, B., & Yan, P. (2020). Boundary-weighted domain adaptive neural network for prostate MR image segmentation. *IEEE Transactions on Medical Imaging, 39*(3), 753–763.

[32]. Zhu, S., et al. (2026). Adapting to missing modalities via knowledge distillation for multimodal brain tumor segmentation. *Medical Image Analysis*.

[33]. Gorrepati, L. P., & Mohanadas, S. (2026). DR-UCSS: A distributionally robust, uncertainty-calibrated counterfactual safety-shielded reinforcement learning framework for hospital operations. International Journal of Intelligent Engineering and Systems, 19(4), 286–302. https://doi.org/10.22266/ijies2026.0430.16

[34]. L. P. Gorrepati, R. Kalapala and G. S. Sargam, "Leveraging Artificial Intelligence and Big Data in Healthcare Provider Systems: Enhancing Patient Care and Operational Efficiency," 2025 Third International Conference on Cyber Physical Systems, Power Electronics and Electric Vehicles (ICPEEV), Hyderabad, India, 2025, pp. 1-6, doi: 10.1109/ICPEEV67897.2025.11291497.