

Modality-Efficient Deep Learning for Automated Multi-Class Segmentation of High-Grade Gliomas

Gabriel Costa^{1*}, Lucas Ribeiro¹, Ricardo Alves², Mariana Lopes³

¹ Department of Neuro-oncology and Deep Learning, Faculty of Medicine, Federal University of Minas Gerais, Belo Horizonte, Brazil.

² Department of Glioma Segmentation and Medical Imaging, Faculty of Medicine, University of Coimbra, Coimbra, Portugal.

³ Department of Modality-Efficient AI in Radiology, Faculty of Medicine, University of São Paulo, São Paulo, Brazil.

*E-mail ✉ gabriel.costa@ufmg.br

Abstract

Glioma constitutes the most frequent primary malignant neoplasm of the brain. Precise delineation of glioma regions on magnetic resonance imaging (MRI) is vital for clinical treatment planning. Conventional approaches to glioma segmentation typically depend on four MRI sequences: T1, T2, T1Gd, and FLAIR. Obtaining all four sequences prolongs patient exposure to the scanner and extends the overall computational time required for segmentation. In practice, however, complete sets of these modalities are not always available because of scanner time limitations or patient-related challenges. To address this issue, the present study applied U-Net-based fully convolutional neural networks for automatic segmentation, specifically investigating whether a reduced set of MRI modalities significantly impairs segmentation performance. The developed framework was trained, validated, and tested on 100 high-grade glioma (HGG) cases under two conditions: first with the complete set of MRI sequences and second with only FLAIR and T1Gd. Evaluation on the held-out test data revealed that the baseline U-Net achieved a mean Dice score of 0.9166 across all sequences and 0.9190 when restricted to FLAIR and T1Gd. To improve performance using only these two modalities, we adopted an ensemble strategy, incorporating pre-trained VGG16, VGG19, and ResNet50 architectures as modified encoders within the U-Net structure for glioma segmentation. We trained, validated, and tested these models on a larger cohort of 259 high-grade glioma (HGG) cases. The experimental outcomes demonstrated mean Dice scores of 0.9395 for the baseline U-Net and 0.9360, 0.9359, and 0.9356 for the ensembles based on VGG16, VGG19, and ResNet50 encoders, respectively. We also compared performance with prior research using the full complement of four MRI modalities. This investigation underscores that FLAIR and T1Gd contribute most to accurate tumor segmentation. The baseline U-Net model proved robust in segmenting sub-tumoral components of HGG and exhibited results comparable to current state-of-the-art methods.

Keywords: Brain tumor segmentation, Convolutional neural networks (CNNs), HGG, Pre-trained CNNs, U-Net

Introduction

Brain tumors have shown a notable rise in occurrence across all age groups globally over the past several

decades. Data from the American Cancer Society indicate that roughly 23,800 to 24,500 patients in the United States were newly diagnosed with malignant central nervous system (CNS) tumors in 2019 [1-3]. Among these, gliomas—tumors arising from glial cells—account for approximately 30% of all primary brain tumors and 80% of malignant cases, with an age-standardized incidence rate of 4.7 per 100,000 person-years [4]. Gliomas are also the leading cause of mortality among primary brain tumors. The World Health Organization (WHO) categorizes gliomas into four grades reflecting their degree of malignancy [5]: low-

Access this article online

<https://smerpub.com/>

Received: 29 September 2025; Accepted: 15 December 2025

Copyright CC BY-NC-SA 4.0

How to cite this article: Costa G, Ribeiro L, Alves R, Lopes M. Modality-Efficient Deep Learning for Automated Multi-Class Segmentation of High-Grade Gliomas. J Med Sci Interdiscip Res. 2025;5(2):188-206. <https://doi.org/10.51847/RjebAFYcFF>

grade gliomas (LGG, grades I and II), which are less aggressive and permit survival over multiple years, and high-grade gliomas (HGG, grades III and IV), which are highly malignant and associated with a median survival of about two years post-resection [6]. Contemporary glioma management typically encompasses surgical intervention, radiotherapy, chemotherapy, or integrated multimodal regimens [7].

A fundamental requirement in glioma therapy is the exact determination of tumor margins by neurosurgeons and oncologists, given that imprecise boundaries can lead to either inadequate treatment or excessive damage, thereby exacerbating patient harm. This delineation process is termed segmentation. In standard clinical workflows, segmentation still relies heavily on manual outlining by specialist neuroradiologists. However, the labor-intensive nature of manually processing numerous MRI slices makes the procedure inefficient and introduces substantial inter- and intra-observer inconsistencies and potential diagnostic errors [8]. These drawbacks can produce grave clinical outcomes. To overcome these limitations, developing robust computer-aided diagnosis (CAD) systems with precise, consistent, and rapid segmentation algorithms has become essential [9]. Although several semi-automated brain tumor segmentation tools exist, their reliance on human oversight makes them vulnerable to operator-dependent mistakes, selection bias, and processing delays. A fully automated solution capable of comprehensively segmenting the entire tumor mass and its internal substructures in one pass is therefore highly desirable. Over time, numerous sophisticated fully automated segmentation techniques have been proposed. The annual Multimodal Brain Tumor Segmentation (BraTS) challenge provides a standardized platform for assessing such algorithms on glioma cases using a common dataset [8].

Many entries in the BraTS challenge have leveraged convolutional neural network (CNN) designs, often enhanced with dense blocks or ResNet elements [10, 11]. Alternative submissions have employed cascaded processing pipelines that exploit the layered anatomical hierarchy of tumor sub-regions to improve segmentation of both the overall tumor and its internal components [12-14]. Although researchers have explored a range of network configurations for medical image segmentation, the U-Net architecture remains the most prevalent framework for brain tumor applications. First introduced in 2015 by Ref. [15] for biomedical image analysis, U-

Net represents a specialized fully convolutional network (FCN) consisting of a contracting encoder path and an expanding decoder path. Its strong performance has led to widespread adoption, with U-Net-based models comprising more than 50% of top-performing submissions in recent BraTS editions [12, 16-19]. Notably, the 2019 BraTS champion [12] utilized a cascaded two-stage U-Net for refined substructure prediction. The runner-up in that challenge similarly relied on U-Net components [20]. In Ref. [21], the DeepSCAN architecture—a lightweight U-Net variant featuring densely linked dilated convolutions and an uncertainty-aware loss—advanced a prior third-place BraTS 2018 model by integrating instance normalization and a compact attention module, again securing third place.

Several BraTS 2019 participants further advanced cascaded methodologies that respect the hierarchical nature of tumor sub-compartments. These systems feed the output of earlier stages as contextual priors into subsequent ones, enabling incremental refinement. This design underpinned the 2019 winning entry [12] and the TU-Net framework [13], which chains three U-Net-style networks to sequentially extract the whole tumor, tumor core, and enhancing tumor core from multimodal scans. Parallel multi-stage cascaded systems built around 3D U-Net cores were likewise described in Refs. [14, 22], and [23].

Other researchers have fused U-Net with complementary CNN backbones such as VGG16, VGG19, ResNet, and DenseNet to harness their respective advantages. Saluja *et al.* [24], working with the BraTS 2018 dataset, combined a modified U-Net for detailed segmentation with VGG-16 for classification, achieving Dice scores of 0.96 (whole tumor), 0.92 (tumor core), and 0.89 (enhancing tumor). Yang *et al.* [25] introduced a 2D ResUNet model on the same dataset, which exhibited enhanced sensitivity and feature representation relative to conventional U-Net, although Dice scores were marginally lower.

Despite these advances, most high-performing brain tumor segmentation models assume the complete availability of T1, T2, T1Gd, and FLAIR sequences. In real-world hospital environments, however, imaging protocols often differ due to time pressures, patient factors, or logistical constraints, resulting in incomplete modality sets. These gaps hinder methods that depend on full multimodal input and can degrade segmentation quality when fewer sequences are used. This situation

prompts inquiry into the optimal subset of modalities that can sustain strong performance. Research by Ref. [26] evaluated modality combinations and determined that T1Gd exerts the strongest influence on segmentation outcomes. Ref. [27] identified T1Gd and FLAIR as the most critical sequences in the BraTS manual annotation process. A subsequent investigation by Ref. [28] tested U-Net variants across heterogeneous BraTS datasets with missing sequences and still reported robust Dice similarity coefficients (internal = 0.86, external = 0.93). Building on these insights, the present study establishes U-Net as the baseline architecture to improve glioma segmentation accuracy and sub-region localization while reducing computational demands, memory usage, and runtime compared with earlier efforts on equivalent data. The research also systematically assesses how restricting the number of MRI modalities affects model effectiveness. Ultimately, these contributions aim to reduce resource requirements and improve the practicality and efficiency of clinical MRI workflows.

Materials and Methods

This investigation aims to show that combining T1Gd and FLAIR modalities is sufficient for reliable segmentation of high-grade gliomas (HGG), achieving

performance comparable to models trained on the complete four-modality BraTS 2019 dataset using a U-Net backbone. Additionally, the study evaluates whether substituting the standard U-Net encoder with independent pre-trained networks (VGG16, VGG19, and ResNet50) can yield further improvements when operating exclusively on T1Gd and FLAIR inputs. **Figure 1** provides a schematic of the overall pipeline. As shown in the diagram, the methodology primarily encompasses data preparation, followed by model training and performance assessment. In the initial stage, 2D axial slices from multimodal brain MRI volumes were imported. These images then received a series of preprocessing treatments. Next, the data were divided into training, validation, and test partitions. The training portion, paired with its ground-truth masks, was used to optimize the network using selected loss functions. We then used the final model parameters to generate predictions on unseen validation and test slices. We systematically compared the resulting segmentation outputs with the reference ground truth masks using several quantitative metrics. When performance fell short of expectations, we refined the network configuration and training settings until the predicted masks and associated evaluation indicators reached acceptable levels.

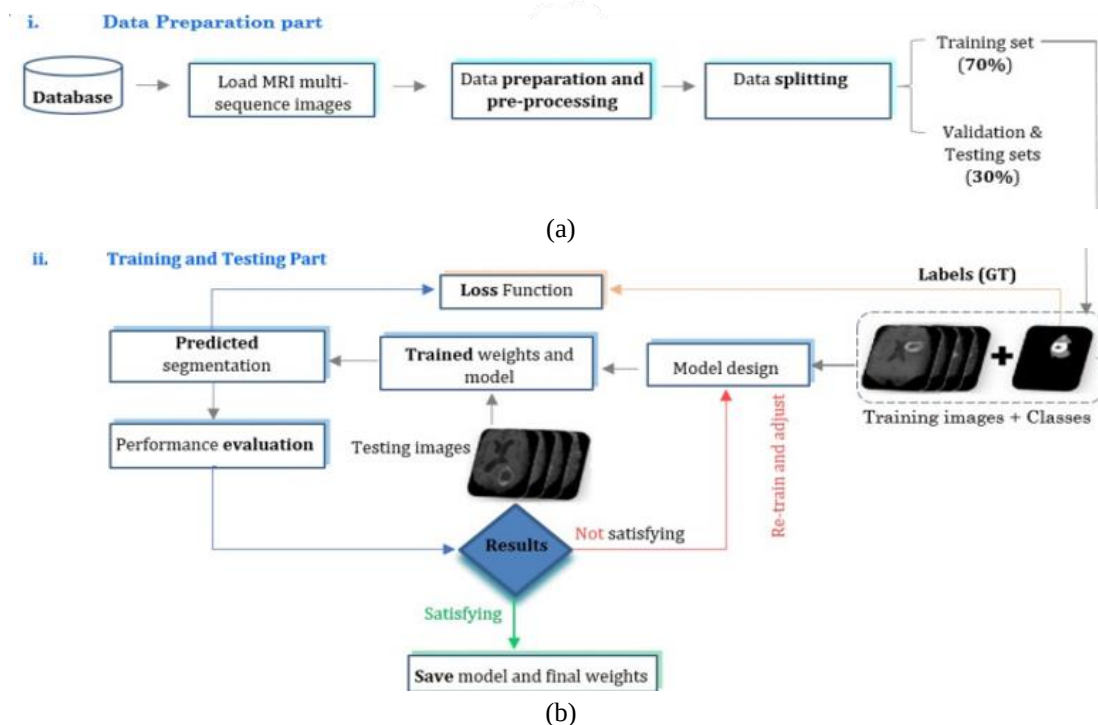


Figure 1. Workflow of the proposed methodology.

Dataset description

This work employed 259 high-grade glioma cases from the BraTS 2019 repository [8, 27, 29]. This dataset features pre-operative multi-parametric MRI acquisitions collected from diverse clinical centers and is dedicated to advancing research on brain tumor segmentation, particularly for gliomas. It supplies a comprehensive collection of MRI volumes with expert annotations of tumor sub-compartments. Multiple annotators independently performed manual segmentations according to a uniform protocol, and an experienced board-certified neuroradiologist reviewed and finalized the ground truth labels. For each case, we supplied four MRI sequences: FLAIR, T1-weighted, T1Gd, and T2-weighted, along with the associated ground truth (GT)

maps. The GT labels categorize voxels into four classes: background, necrotic and non-enhancing tumor (NCR/NET), edema (ED), and enhancing tumor (ET) [30, 31]. **Figure 2** shows sample axial slices from the BraTS 2019 dataset, illustrating the four MRI modalities alongside their ground-truth annotations. Before distribution, the BraTS organizers aligned all scans to a standardized anatomical template, removed non-brain tissue, and resampled the volumes to an isotropic voxel size of 1 mm^3 . Each volumetric sequence measures $155 \times 240 \times 240$ voxels and is provided in the Neuroimaging Informatics Technology Initiative (NIfTI) file format. The complete BraTS 2019 database is available for free at <https://www.kaggle.com/datasets/aryashah2k/brain-tumor-segmentation-brats-2019>.

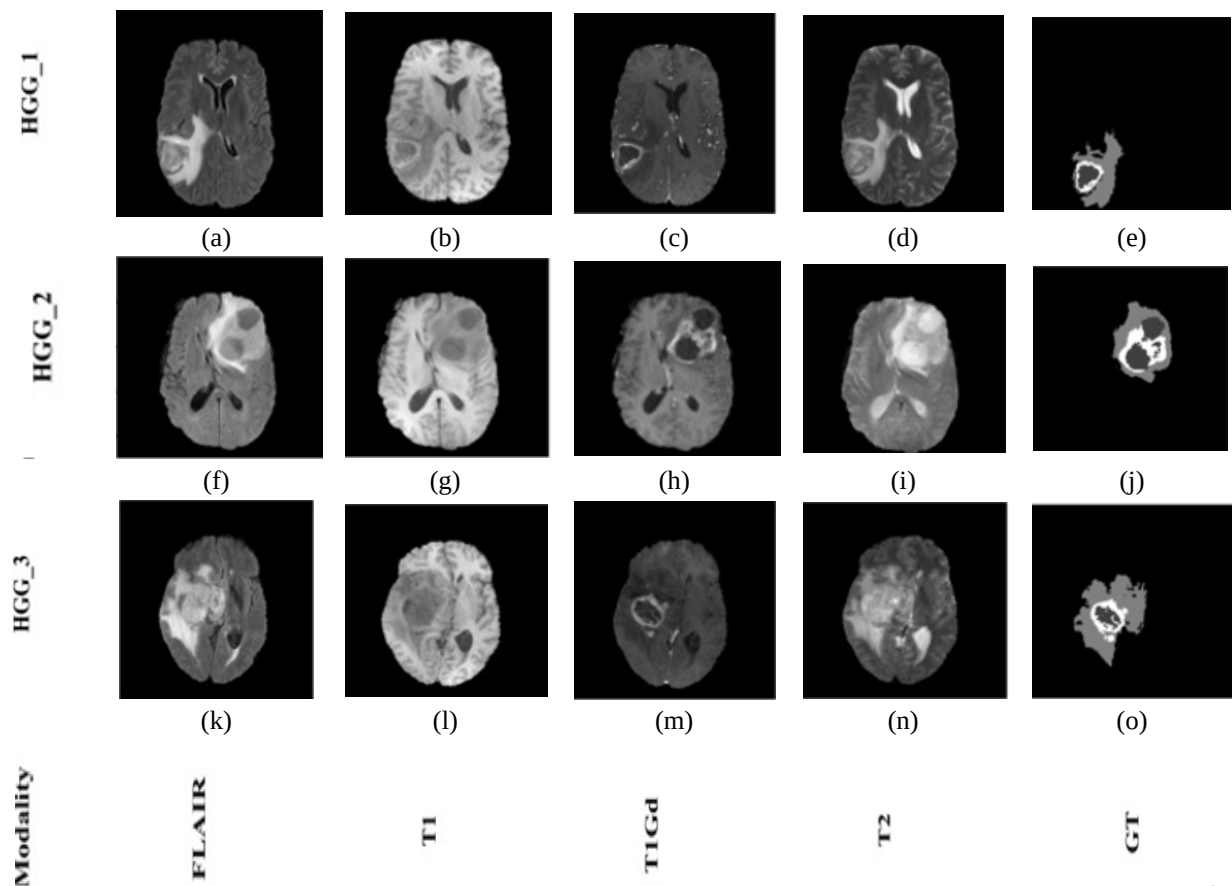


Figure 2. An illustration of the BraTS2019 HGG dataset. The figure displays axial slices from three different patients across the four MRI modalities, along with the corresponding ground-truth segmentation [29].

Dataset preparation

The initial dimensions of the dataset were structured as (N, M, D, H, W), in which N represents the number of

patients, M denotes the number of MRI sequences, D corresponds to the number of axial slices (depth), H indicates the height, and W the width of each slice. For

the HGG dataset, these dimensions were (259, 4, 155, 240, 240), with matching ground truth (GT) labels of size (N, D, H, W).

Each MRI sequence comprised 155 axial slices of 240×240 pixels. Processing the full volume across all 259 patients would impose substantial computational and memory burdens. To address this, we implemented targeted preprocessing measures. Examination of each subject's images revealed that many slices lacked relevant tumor or even brain tissue; specifically, only slices 30 through 120 proved informative for the segmentation task. This observation was corroborated by the ground truth maps, in which discarded slices contained solely zero-intensity values. The same selection process was applied across all cases, removing 65 non-informative slices per volume and retaining 90 slices. Consequently, the dataset dimensions were reduced to (259, 4, 90, 240, 240).

Next, visual inspection showed that a large proportion of pixels represented extraneous background areas. Each axial slice was therefore centrally cropped to 192×192 pixels to retain only the brain region. This step further adjusted the dataset dimensions to (259, 4, 90, 192, 192). Finally, the 2D axial slices were reorganized from the original (N, M, D, H, W) format into (S, H, W, M), where S equals the product of N and D, representing the aggregate number of 2D slices across the entire patient cohort.

Dataset preprocessing

A key obstacle in automated MRI segmentation is the considerable variability of image intensity distributions within and between scans. To mitigate this, every slice in the dataset was normalized to achieve a zero mean and unit variance. Normalization was performed by subtracting the mean intensity value from each pixel and subsequently dividing by the maximum intensity value, as expressed in Eq. (1):

$$\text{Normalized Slice} = \frac{(x - x_{\text{mean}})}{x_{\text{max}}} \quad (1)$$

Finally, the patient cohort was randomly partitioned into training, validation, and testing sets using a 70%, 15%, and 15% allocation, respectively.

Proposed network architecture

The study began by implementing a standard U-Net as the baseline architecture for segmenting glioma

subregions. This was subsequently extended through three variants featuring modified encoders, as detailed in the following sections.

1. The baseline U-Net was trained from random initialization using Xavier (Glorot) Initialization [32]. Training utilized 70% of the dataset along with the corresponding ground truth labels.
2. U-Net training was performed in two distinct configurations: one employing the complete set of modalities (FLAIR, T1, T1Gd, and T2) and another restricted to FLAIR and T1Gd only. Due to GPU memory constraints within the Google Colab environment, each configuration was trained on a randomly selected subset of 100 HGG cases ($n = 100$).
3. To assess potential gains in segmentation performance when using only FLAIR and T1Gd, the encoder component of the U-Net was substituted with convolutional layers from pre-trained CNN models. These modified networks were then trained on the full set of 259 HGG cases ($n = 259$). The pre-trained encoders included VGG16, VGG19, and ResNet50, whose weights had been previously optimized on the ImageNet dataset.

The subsequent subsections examine these architectures in greater depth.

U-Net

The U-Net architecture employed in this study, illustrated in **Figure 3**, closely follows the established design introduced in Ref. [15], with slight adjustments to suit the input dimensions and required output format. This symmetric U-Net consists of four resolution levels, each containing encoder and corresponding decoder blocks linked by skip connections. Within every encoder level, two successive 3×3 conv2D operations are applied, each followed by batch normalization and a Rectified Linear Unit (ReLU) activation, before downsampling via a 2×2 maxpool2D layer. The filter counts (channels) increase progressively across levels as 64, 128, 256, and 512, reaching 1024 at the lowest (bottleneck) level. In the decoder pathway, the process is inverted: upsampling and feature concatenation precede the application of equivalent convolutional blocks from the encoder. The final output layer performs multi-class segmentation through a 1×1 convolution followed by a Softmax activation. Unlike the original U-Net configuration [15], which omitted padding and caused progressive shrinkage of feature maps, we incorporated zero-padding to enable

seamless feature concatenation. Furthermore, we inserted a batch normalization (BN) layer after each convolution to promote faster convergence. This addition also serves

a regularizing effect, reducing the necessity for techniques such as dropout [33].

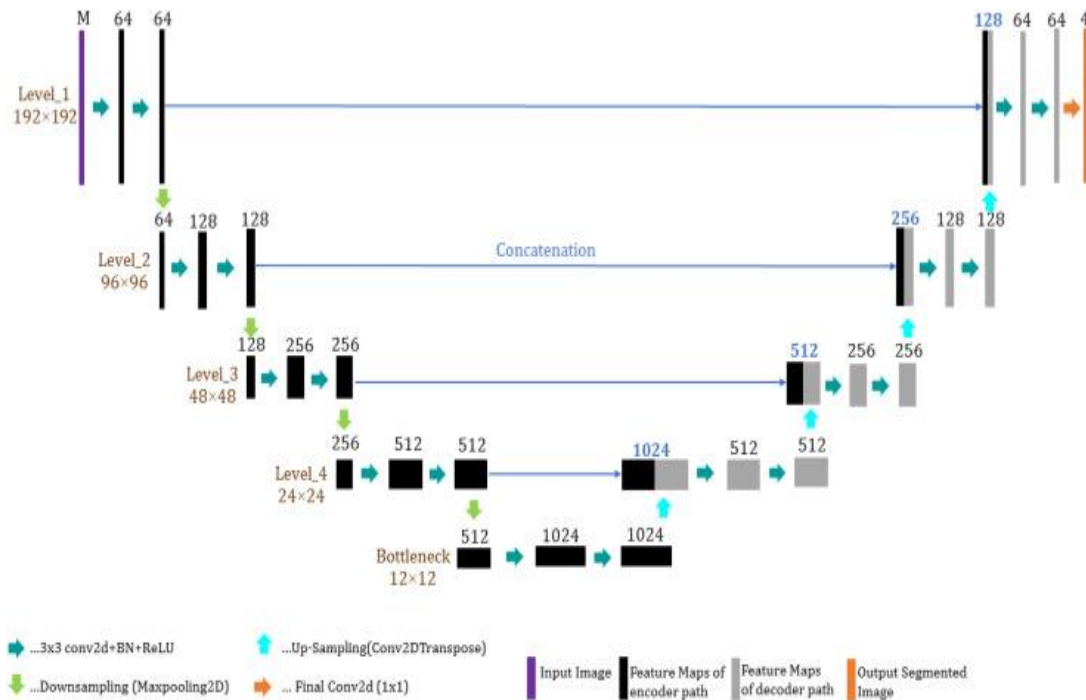


Figure 3. The proposed U-Net schematic.

The figure illustrates the U-Net architecture configured for 192×192 input image patches. This encoder-decoder network features a contracting path (encoding section on the left) that progressively reduces spatial dimensions and an expansive path (decoding section on the right) that restores the original image size. Black and gray rectangles denote feature maps with varying channel counts. Rectangle height reflects relative spatial dimensions (in pixels), while width corresponds to the number of channels, with explicit values indicated at the top. The left side of the diagram marks overall image dimensions. Long arrows indicate skip connections that carry feature maps from the contracting path. Short arrows represent the operations detailed in the legend. The symbol 'M' above the input layer indicates the number of MRI modalities utilized.

VGG-UNet

The VGG family of networks [34], encompassing VGG16 and VGG19 variants, consists of deep convolutional architectures renowned for their strong

results on the ImageNet classification benchmark [35]. The numerals 16 and 19 refer to the total count of weighted layers. The primary distinction between VGG16 and VGG19 is the three additional layers in the latter model.

As depicted in **Figures 4a and 4b**, both VGG16 and VGG19 comprise five convolutional blocks (CBs) succeeded by three fully connected (dense) layers. Each convolutional block includes multiple 3×3 convolutional layers employing ReLU activation, followed by MaxPooling2D downsampling. Within successive blocks, the number of feature maps doubles. After the convolutional blocks, the dense layers map the extracted features to class predictions. In this work, the feature-extraction backbone extends from the input through the final max-pooling layer, with the remaining components constituting the classification head. VGG16 and VGG19 were chosen for integration because their structure is compatible with the contracting pathway of the standard U-Net encoder.

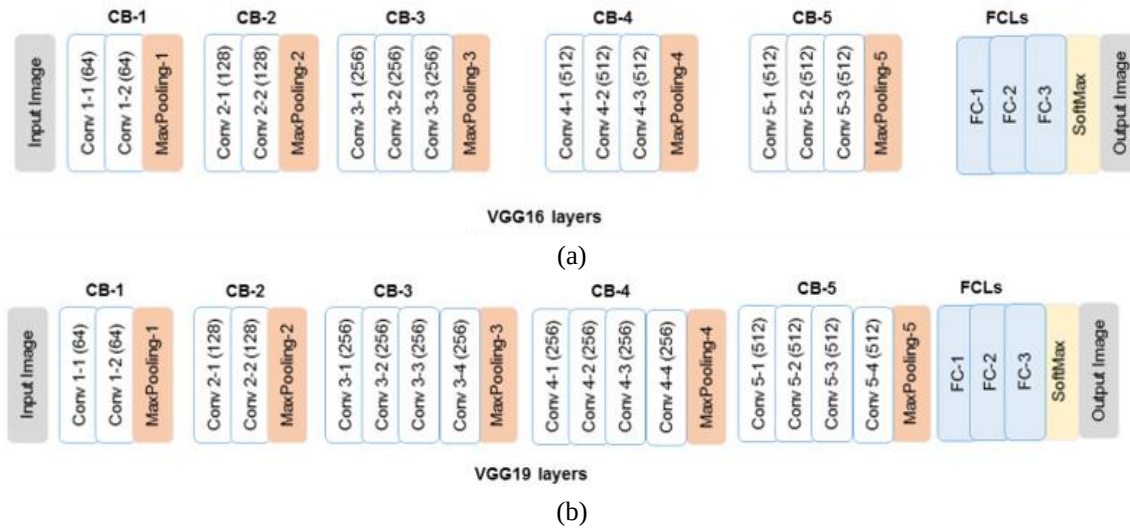


Figure 4. VGG layers arrangement.

The VGG16 and VGG19 networks were adapted in this study to align with the U-Net framework via the following modifications:

- 1) Retaining the first four convolutional blocks (CB-1 to CB-4) to form the encoder,
- 2) Utilizing the fifth convolutional block as the bridge (bottleneck) while omitting its max-pooling operation, and

- 3) Substituting the three fully connected layers of each VGG model with the expansive decoder pathway from the baseline U-Net shown in **Figure 3**.

The resulting overall structure retains the characteristic U-shape, as presented in **Figures 5 and 6**. Light red rectangles in the diagrams emphasize the distinctions between the modified encoders and the original U-Net encoder.

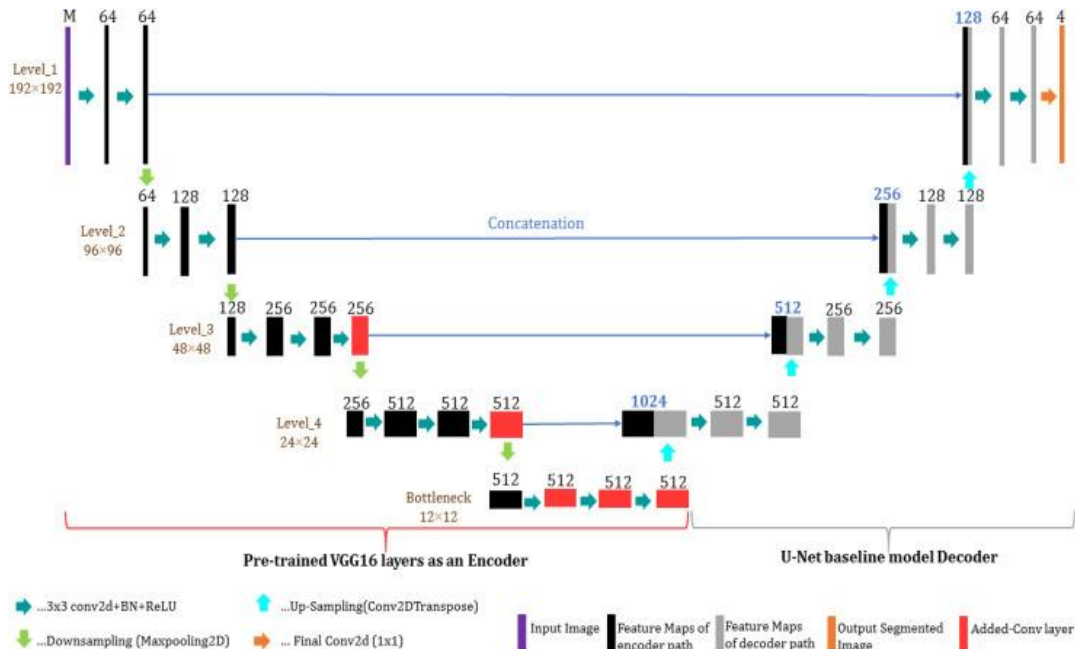


Figure 5. An illustration of a modified U-Net architecture with pre-trained VGG16 as an encoder part.

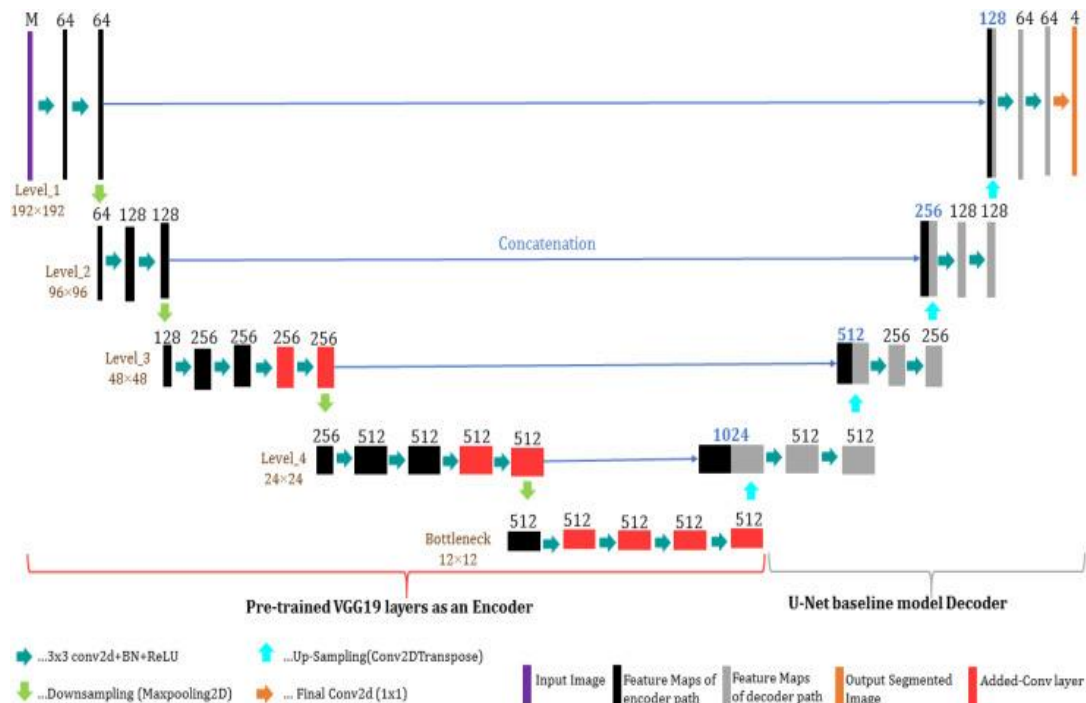


Figure 6. An illustration of a modified U-Net architecture with pre-trained VGG19 as an encoder part.

ResNet50-UNet

Residual Networks constitute a prominent family of deep neural architectures frequently employed as backbones across a wide range of computer vision applications. ResNet50, in particular, won the 2015 ImageNet competition [36]. These networks build upon conventional stacks of convolutional, pooling, activation, and fully connected layers by incorporating residual

connections. ResNet50 specifically comprises 48 convolutional layers, one MaxPool layer, and one average pooling layer. It is organized into four residual blocks, as described in Ref. [36]. For the encoder section of the U-Net in this study, we used the first three residual blocks. We incorporated the convolutional layers from the fourth block to form the bottleneck, as illustrated in **Figure 7**.

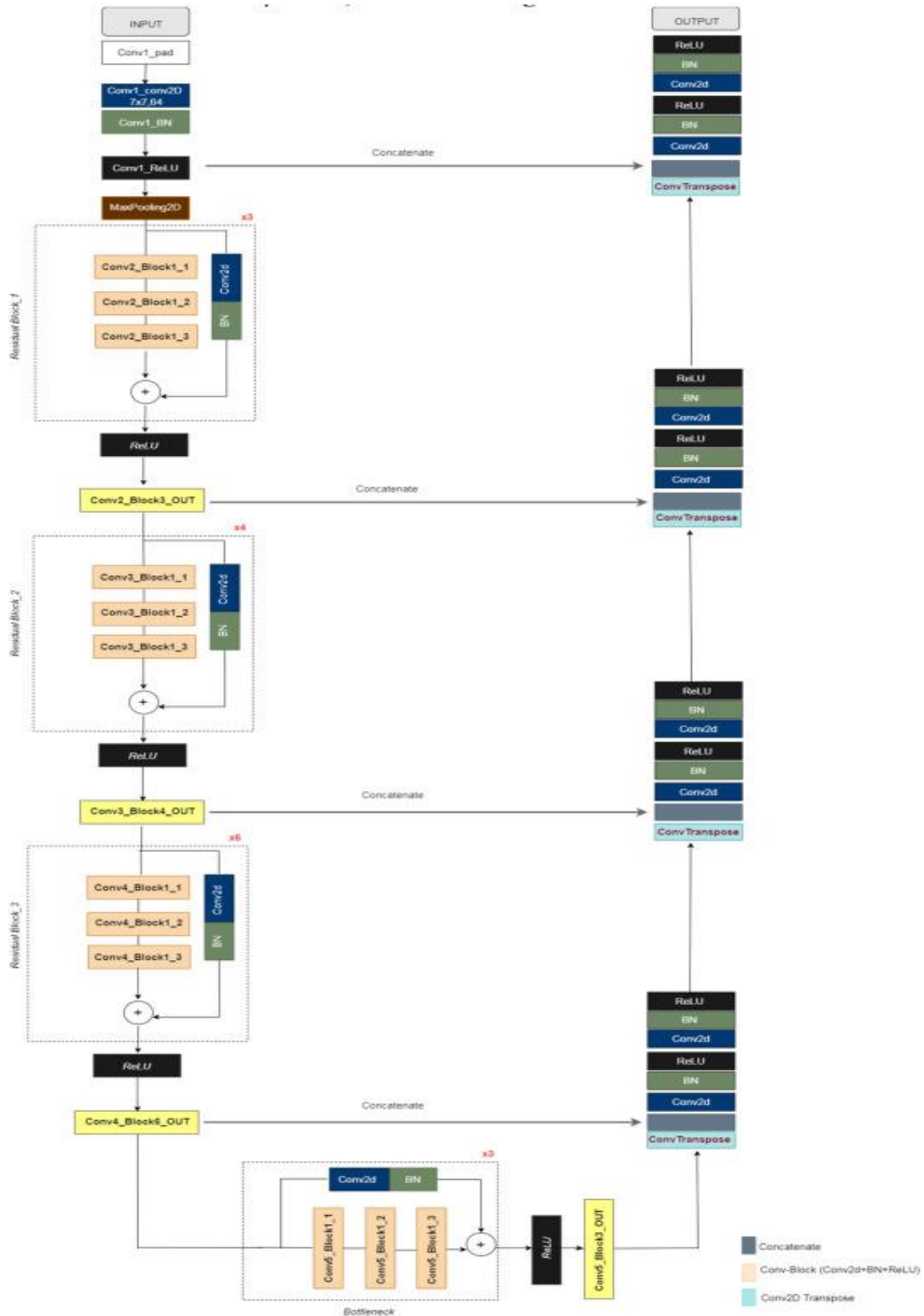


Figure 7. A schematic view of the U-Net architecture using ResNet50 as an encoder.

Training configuration

All models developed in this work were trained for 100 epochs with a batch size of 32. Analysis of the class distribution within the segmentation masks revealed a pronounced imbalance, with the vast majority of voxels belonging to the background (healthy tissue) and only a small fraction representing tumor subregions. Such class imbalance tends to bias networks toward the dominant background class. To counteract this effect, we applied class-specific pixel weighting. Training proceeded in two phases using distinct loss functions: the first 50 epochs employed weighted categorical cross-entropy (WCCE) loss, while the subsequent 50 epochs utilized soft Dice loss (SDL) [12, 37]. This staged approach first enabled the model to recognize all classes (including tumor substructures) and then refined boundary precision and overlap quality. In the WCCE loss, higher weights were assigned to underrepresented (minority) classes that are more difficult to predict. In contrast, soft Dice loss inherently handles class imbalance without requiring explicit weighting.

The standard weighted categorical cross-entropy loss is defined as follows (2) [38]:

$$L_{wcce} = \frac{-1}{M} \sum_{k=1}^K \sum_{m=1}^M w_k \times y_m^k \times \log(h_{\theta}(x_m, k)) \quad (2)$$

Where:

M: number of training samples

K: number of desired classes

w_k: the assigned weight for each class K.

: the actual class for the training sample m for the specific class K.

x_m: the input for a training sample m.

h_θ: model weight

The soft Dice loss quantifies the spatial overlap between the predicted probability map (P) and the ground truth map (G). It is formulated as (3):

$$L_{SDL} = 1 - \frac{2 * \sum(G * P)}{(\sum G^2 + \sum P^2) + \epsilon} \quad (3)$$

Where: P denotes the model's predicted output, G represents the ground truth, and ε is a small smoothing constant (1×10⁻⁶) introduced in the denominator to prevent division by zero. This formulation constrains the loss value to the interval [0,1].

The Adam optimizer [39] was used with an initial learning rate of 1e⁻⁵. The output layer employed the Softmax activation function, which is well-suited for multi-class segmentation as it converts raw logits into normalized probabilities ranging from 0 to 1 for each class.

Implementation environment

All models in this study were implemented in Python version 3.7.10. The deep learning components were developed and trained using TensorFlow 1.15.2 and Keras 2.3.1, alongside supporting libraries. Given the substantial size of the dataset and the computational demands of the networks, resource limitations posed a significant challenge. Therefore, all experiments were executed on Google Colab, using its free GPU runtime and Google Drive for efficient dataset storage and retrieval.

Evaluation metrics

Performance of the developed models was assessed with respect to the BraTS-defined tumor regions, which aggregate specific intra-tumoral classes as follows:

- Whole Tumor (WT): encompasses all tumor components, namely edema, enhancing tumor, and tumor core.
- Tumor Core (TC): comprises enhancing tumor, necrosis, and non-enhancing tumor regions, while excluding edema.
- Enhancing Tumor (ET): corresponds exclusively to the enhancing tumor class.

The principal evaluation measure was the Dice Similarity Coefficient (DSC), widely regarded as the benchmark metric for medical image segmentation tasks. Unlike simple accuracy, DSC provides a balanced assessment by penalizing false positives and false negatives equally. Originally proposed by Dice (1945), DSC quantifies the degree of spatial overlap between the ground truth segmentation (G) and the predicted segmentation (P). It is computed according to (4):

$$DSC = \frac{2 |P \cap G|}{|P| + |G|} \quad (4)$$

DSC values range from 0 (indicating no overlap) to 1 (indicating complete overlap).

In relation to the confusion matrix, DSC can equivalently be expressed in terms of true positives, false positives, true negatives, and false negatives as (5):

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (5)$$

Where True Positive (TP) represents correctly identified positive samples, True Negative (TN) denotes correctly identified negative samples, False Positive (FP) indicates samples incorrectly classified as positive, and False Negative (FN) refers to samples incorrectly classified as negative.

For multi-class pixel-wise classification with $L = 4$ classes, the mean Dice Similarity Coefficient is given by Eq. (6):

$$\overline{DSC} = \frac{1}{L} \sum_{i=1}^L \frac{2 TP_i}{2 TP_i + FP_i + FN_i} \quad (6)$$

Where TP_i , FP_i , TN_i , and FN_i represent the true positives, false positives, true negatives, and false negatives for class i , respectively.

Additional performance metrics employed include the following:

- Sensitivity measures the proportion of actual positive instances correctly identified. It reflects the model's ability to detect pixels belonging to the positive class and is calculated as (7):

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (7)$$

- Specificity, the complement of sensitivity, measures the proportion of actual negative instances correctly identified. It indicates the model's effectiveness in correctly rejecting pixels that do not belong to the target class and is formulated as (8):

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (8)$$

- Cohen's Kappa coefficient offers a robust alternative to overall accuracy by adjusting for agreement occurring by chance [40]. It helps provide a more reliable measure of classification quality while reducing the influence of random predictions. Kappa scores range from -1 to 1 , where 1 signifies perfect agreement, -1 indicates complete disagreement, and 0 corresponds to agreement no better than chance. A comprehensive interpretation of Kappa values is available in Ref. [41]. The coefficient is defined by the following expression (9):

$$\text{Kappa} (\kappa) = \frac{c \times N - \sum_i^M p_i \times t_i}{N^2 - \sum_i^M p_i \times t_i} \quad (9)$$

Where,

N : the total number of elements,

c : the sum of the diagonal elements of the confusion matrix (correctly classified instances),

$p_i = \sum_k^M C_{ik}$: the total number of times class i was predicted (column total),

$t_i = \sum_k^M C_{ki}$: the total number of times class i actually occurred (row total).

Results and Discussion

This section reports the key findings obtained from evaluating various MRI input modality combinations. The initial analysis focuses on the effectiveness of the baseline U-Net model under differing modality configurations. Training was carried out on a randomly chosen subset of 100 HGG cases in two separate experiments: one employing the complete collection of four MRI sequences (FLAIR, T1, T1Gd, and T2) and the other relying solely on FLAIR and T1Gd. **Table 1** summarizes these outcomes.

Table 1. U-NET baseline model's mean DSC performance across different MRI modalities ($n = 100$).

MRI modality configuration	Dice similarity coefficient (DSC)	Sensitivity	Specificity	Model training duration
Complete MRI modalities	0.9166	0.9265	0.9913	2 h 35 min
FLAIR and T1Gd modalities	0.9190	0.9123	0.9849	1 h 31 min

The subsequent analysis assesses the modified U-Net architectures when restricted exclusively to T1Gd and FLAIR inputs. For this evaluation, we used the full set of 259 HGG cases.

Table 2 reports mean DSC values, sensitivity, specificity, and Kappa statistics. **Figure 8** shows segmentation accuracy for the primary tumor subregions: whole tumor (WT), tumor core (TC), and enhancing

tumor (ET). In addition, **Figure 9** displays the normalized confusion matrices for all tested models. In these matrices, rows represent the actual ground truth classes, while columns indicate the predicted classes.

Diagonal percentages reflect correctly assigned pixels per class, and off-diagonal entries show the distribution of classification errors.

Table 2. Comparison OF U-NET and modified U-NET models' mean performance (n = 259). The best result IS SHOWN IN bold

Model	DSC	Kappa (κ)	Sensitivity	Specificity
U-Net baseline	0.9395	0.9432	0.9359	0.9890
VGG16-UNet	0.9360	0.9383	0.9346	0.9886
VGG19-UNet	0.9359	0.9379	0.9291	0.9874
ResNet50-UNet	0.9356	0.9378	0.9327	0.9872

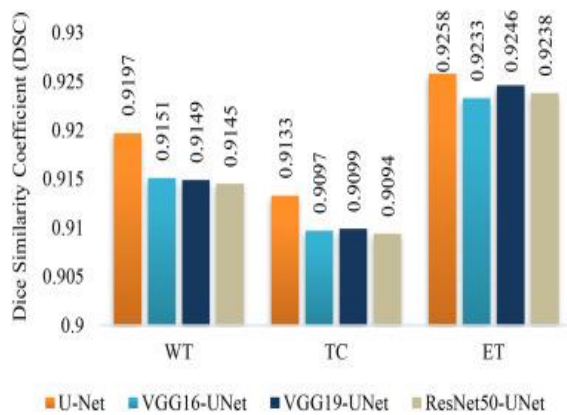
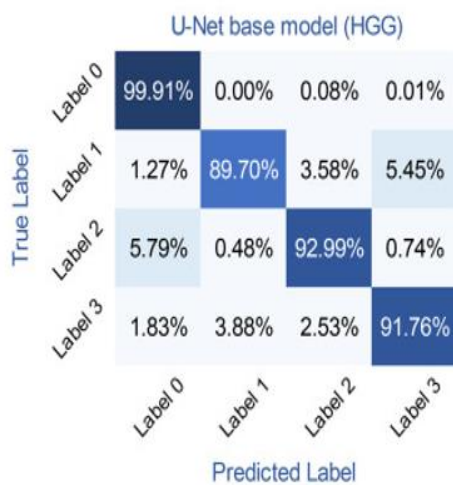
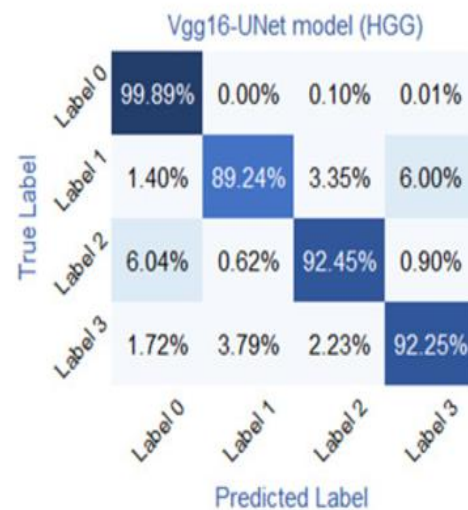


Figure 8. Bar plots showing the DSC computed for the target structures of the BraTS2019 HGG cases (n = 259).



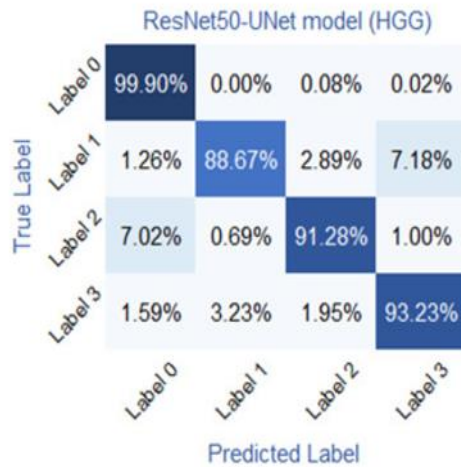
(a)



(b)



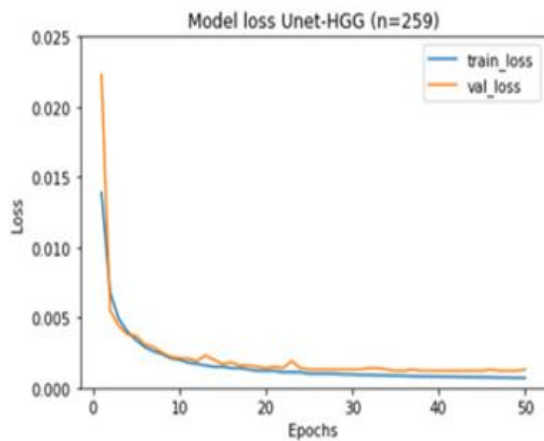
(c)



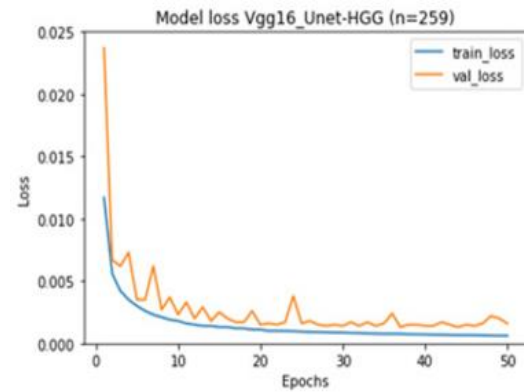
(d)

Figure 9. Normalized Confusion matrix using different approaches on the HGG dataset (n = 259). Columns denote the predicted class, while rows denote the true class. The data is expressed as a percentage of classified pixels.

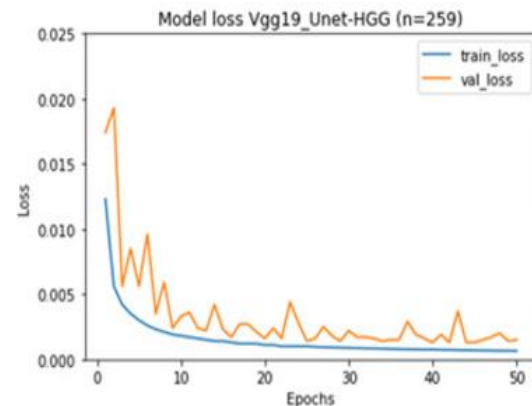
Figure 10 illustrates the progression of training and validation losses for the standard U-Net and its encoder-modified counterparts. Loss values (y-axis) are plotted against the number of training epochs (x-axis), with training loss shown in blue and validation loss in orange.



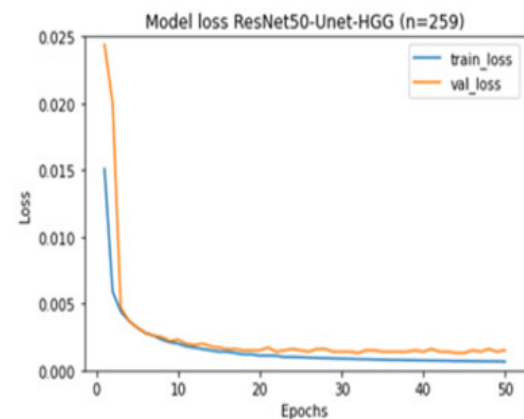
(a)



(b)



(c)



(d)

Figure 10. Training and validation losses versus the number of epochs for the four proposed models: (a) U-Net and (b–d) modified encoder U-Net.

Figure 11 provides qualitative examples of the segmentation outputs. These examples highlight the strong correspondence between the baseline U-Net predictions and the expert-annotated ground truth masks.

Further interpretation of these visual results appears in the discussion section.

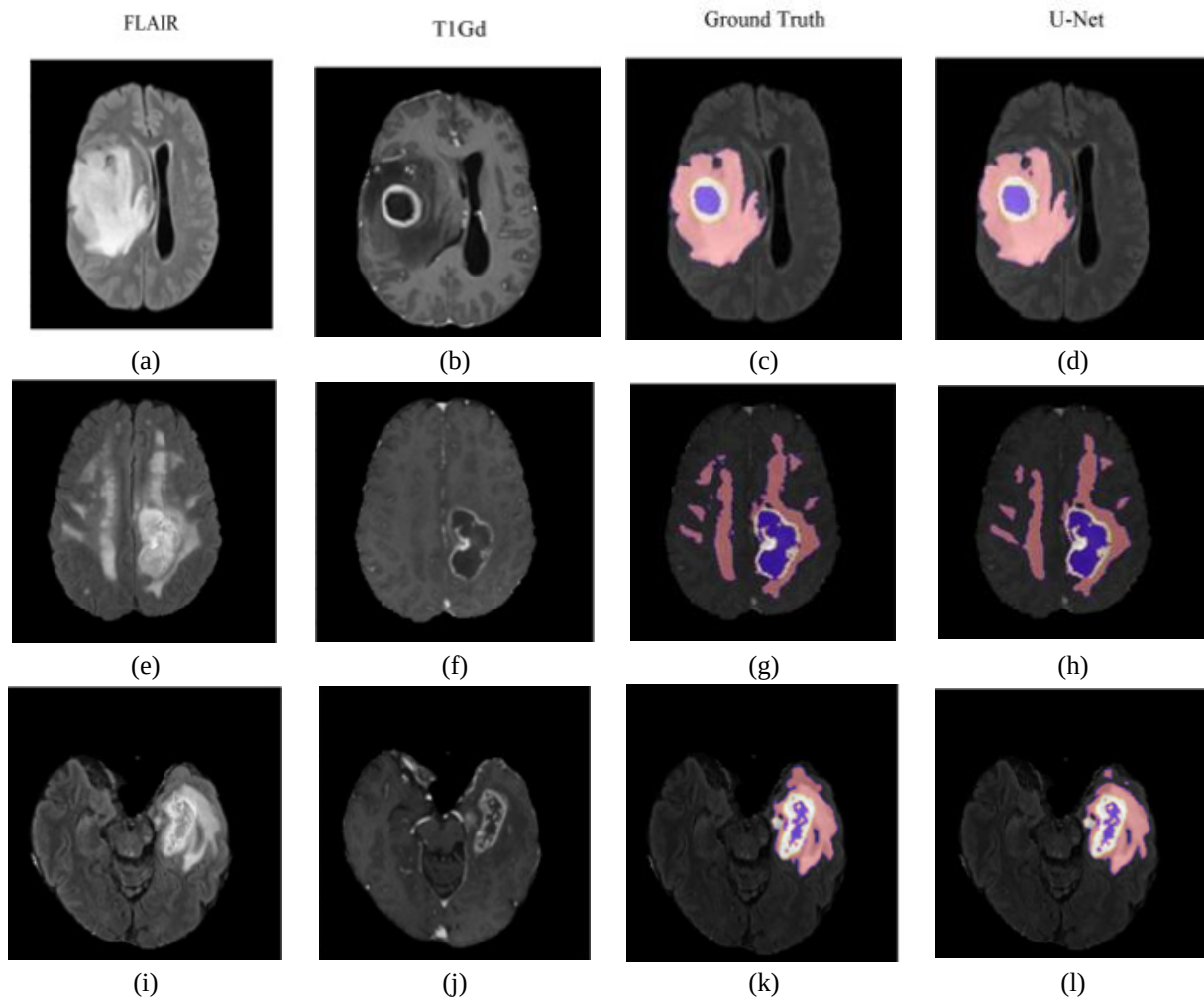


Figure 11. A qualitative comparison of the predicted segmentation using the baseline U-Net model (best-proposed approach) and the actual segmentation (GT) on three different representative cases of the BraTS 2019 HGG dataset. Each column, from left to right, shows the input FLAIR scan, T1Gd scan, ground truth segmentation, and automated segmentation using U-Net. Each color represents a tumor class: Pink indicates ED, white indicates ET, whereas purple indicates NCR/NET (rest of tumor).

Performance of U-Net Baseline Model using different MRI modalities

As indicated in **Table 1**, the experimental outcomes confirm that employing only the T1Gd and FLAIR modalities enables satisfactory segmentation of the essential sub-tumoral components within the HGG dataset, yielding acceptable Dice Similarity Coefficient values. Furthermore, this reduced-modality approach requires substantially fewer computational resources than using all four sequences. Consequently, the

remainder of the paper concentrates on this promising observation.

Performance of modified U-Net Architecture using T1Gd and FLAIR only

Building on the observations from subsection 4.1, this analysis extends the evaluation to the entire high-grade glioma cohort of 259 subjects, using only T1Gd and FLAIR sequences. In this phase, we replaced the baseline U-Net encoder with convolutional layers from three established pre-trained CNN architectures—VGG16,

VGG19, and ResNet50—while preserving the original U-Net decoder. This modification aimed to determine whether replacing the encoder could improve overall segmentation performance. **Table 2** presents the quantitative results, including mean DSC scores for the tumor substructures as well as Kappa, sensitivity, and specificity. **Figure 8** further visualizes the segmentation accuracy across the whole tumor, tumor core, and enhancing tumor regions.

The differences in performance among the models, as shown in **Table 2**, are relatively small. This suggests that replacing the baseline encoder with pre-trained alternatives does not yield meaningful improvements in HGG segmentation. Notably, the standard U-Net achieved the highest overall mean DSC score of 0.9395 across the full dataset. Segmentation of the enhancing tumor (ET) region slightly surpassed that of the tumor core (TC) and whole tumor (WT), as depicted in **Figure 8**. This pattern may stem from the exclusion of the T2 sequence, whose information regarding edema largely overlaps with that provided by FLAIR. Consequently, the network could allocate greater attention to other relevant features and structures.

In addition, the elevated Kappa coefficients reflect strong agreement between the model predictions and the ground truth annotations. This metric underscores the model's robustness and reliability, as it evaluates pixel-level classification accuracy rather than merely reflecting class frequency distributions, in contrast to simpler baseline classifiers.

Figure 10 shows the learning curves, tracking the evolution of training and validation losses across the four U-Net variants during training. In all cases, both loss values showed a consistent downward trend as epochs increased. Although the convergence trajectories differed slightly among models, stabilization generally occurred around 20 epochs. The baseline U-Net and ResNet50-U-Net converged faster and showed smoother validation loss profiles than the VGG-based variants, whose loss curves tended to plateau more noticeably as training advanced.

Figure 11 presents qualitative examples of the segmentation outputs for a representative HGG case. These visualizations reveal that the U-Net model effectively captures the contours and spatial extent of individual tumor subregions in a manner closely aligned with expert manual delineations (ground truth).

Comparison with other works

Table 3 presents a comparative analysis between the outcomes of the present study and those reported in recent investigations employing the same BraTS HGG dataset. The comparison encompasses the primary methodological approach, the number of MRI sequences utilized, and the achieved Dice Similarity Coefficient (DSC) values for the key tumor subregions: whole tumor (WT), tumor core (TC), and enhancing tumor (ET). The proposed model demonstrates superior performance relative to existing approaches, most notably in delineating the enhancing tumor region. Notably, the baseline U-Net was trained exclusively on T1Gd and FLAIR modalities. Omitting T1 and T2 sequences did not substantially degrade performance compared with prior studies that incorporated all four modalities. Moreover, the current model markedly outperformed earlier works that also relied solely on FLAIR and T1Gd. These observations reinforce the premise that restricting the input to fewer modalities encourages the network to concentrate on tumor features that may be less discernible in the excluded sequences (T1 and T2). The results corroborate earlier findings regarding the pivotal importance of FLAIR and T1Gd for distinguishing tumor subregions [42], align with the established expert annotation protocols used in the BraTS dataset [27], and affirm the robustness of a properly trained U-Net architecture for tumor segmentation tasks [18]. With only modest adaptations, the baseline U-Net in this study attained competitive Dice scores on the BraTS dataset despite employing large image patches, a relatively modest dataset size, limited preprocessing, and carefully selected loss functions.

Table 3. Performance comparison (DSC) with recently comparable works that used the BraTS HGG dataset.

Reference	Segmentation model	Dataset	MRI modalities	Whole tumor (WT) DSC	Enhancing tumor (ET) DSC	Tumor core (TC) DSC
Rafi <i>et al.</i> [43]	U-Net	BraTS 2019	FLAIR, T1, T2, T1Gd	0.810	0.790	0.610
Di Ieva <i>et al.</i> [44]	CNN	BraTS 2019	FLAIR, T1, T2, T1Gd	0.872	0.849	0.815

Di Ieva et al. [44]	CNN	BraTS 2019	FLAIR, T1Gd	0.887	0.856	0.817
Ahmad et al. [45]	3D U-Net	BraTS 2019	FLAIR, T1, T2, T1Gd	0.871	0.846	0.784
Mzoughi et al. [46]	Cascaded U-Net	BraTS 2018	FLAIR, T1, T2, T1Gd	0.900	0.840	0.800
Rehman et al. [47]	BU-Net	BraTS 2018	FLAIR, T1, T2, T1Gd	0.901	0.867	0.835
Pemberton et al. [28]	nn-U-Net	BraTS 2021	Incomplete MRI Sequences	0.95	0.84	0.96
Saluja et al. [24]	U-Net	BraTS 2018	FLAIR, T1, T2, T1Gd	0.96	0.92	0.89
Yang et al. [25]	2D ResUNet	BraTS 2018	FLAIR, T1, T2, T1Gd	0.85	0.74	0.71
Abd-Allah et al. [23]	Cascaded U-Net	BraTS 2017	FLAIR, T1, T2, T1Gd	0.91	0.80	0.83
Proposed Approach	U-Net	BraTS 2019	FLAIR, T1Gd	0.919	0.926	0.913

Clinical implications and future outlooks

Traditional assessment of brain tumors relies on manual interpretation by specialist neuroradiologists. The outcomes of this study therefore carry important clinical relevance, supporting the integration of computer-aided diagnostic tools within radiology workflows. Automated glioma segmentation can enhance diagnostic precision, reduce radiologist workload, and ultimately improve patient outcomes and treatment effectiveness. In future work, the research team intends to validate the proposed model on independently acquired local clinical datasets and to extend its application beyond segmentation. Potential directions include automated tumor grading, survival prediction, and forecasting tumor progression or regression dynamics.

Conclusion

Accurate and reproducible brain tumor segmentation is essential for guiding optimal diagnosis and therapeutic strategies. Given the difficulties and inherent variability associated with manual tumor annotation, automated approaches are highly advantageous. To address this need, the current study introduced a deep learning framework based on fully convolutional U-Net networks to detect and segment high-grade gliomas. Appropriate preprocessing procedures were applied to prepare the data for training while accommodating GPU memory constraints. Due to the severe class imbalance—where tumor voxels constitute only a small fraction of the total brain volume—specialized loss functions, namely

weighted categorical cross-entropy and soft Dice loss, were employed during model optimization.

The investigation initially evaluated the baseline U-Net architecture under both four-modality and two-modality input conditions. The findings show that the proposed method delivers efficient, reliable segmentation using only the two most informative sequences, FLAIR and T1Gd, without relying on the full complement of four modalities. Furthermore, the baseline U-Net proved highly effective for HGG segmentation, achieving strong results without requiring extensive architectural alterations. Future efforts will extend the evaluation of the proposed framework to low-grade gliomas (LGGs).

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin.* 2019;69(1):7-34. doi:10.3322/caac.21551
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020;70(1):7-30. doi:10.3322/caac.21590

3. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA Cancer J Clin.* 2021;71(1):7-33. doi:10.3322/caac.21654
4. Miller KD, Ostrom QT, Kruchko C, Patil N, Tihan T, Cioffi G, et al. Brain and other central nervous system tumor statistics, 2021. *CA Cancer J Clin.* 2021;71(5):381-406. doi:10.3322/caac.21693
5. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol.* 2016;131(6):803-20. doi:10.1007/s00401-016-1545-1
6. Suchorska B, Weller M, Tabatabai G, Senft C, Hau P, Sabel MC, et al. Complete resection of contrast-enhancing tumor volume is associated with improved survival in recurrent glioblastoma—results from the DIRECTOR trial. *Neuro Oncol.* 2016;18(4):549-56. doi:10.1093/neuonc/nov326
7. Tabatabai G, Stupp R, van den Bent MJ, Hegi ME, Tonn JC, Wick W, et al. Molecular diagnostics of gliomas: the clinical perspective. *Acta Neuropathol.* 2010;120(5):585-92. doi:10.1007/s00401-010-0750-6.
8. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans Med Imaging.* 2014;34(10):1993-2024. doi:10.1109/TMI.2014.2377694
9. Petrick N, Sahiner B, Armato SG 3rd, Bert A, Correale L, Delsanto S, et al. Evaluation of computer-aided detection and diagnosis systems. *Med Phys.* 2013;40(8):087001. doi:10.1118/1.4816310
10. McKinley R, Meier R, Wiest R. Ensembles of densely-connected CNNs with label-uncertainty for brain tumor segmentation. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lecture Notes in Computer Science*, vol 11384. Cham: Springer; 2018. p. 456-65. doi:10.1007/978-3-030-11726-9_40
11. Pawar K, Chen Z, Shah NJ, Egan GF. An ensemble of 2D convolutional neural network for 3D brain tumor segmentation. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science*, vol 11993. Cham: Springer; 2020. p. 359-67. doi:10.1007/978-3-030-46640-4_34
12. Jiang Z, Ding C, Liu M, Tao D. Two-stage cascaded U-Net: 1st place solution to BraTS challenge 2019 segmentation task. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science*, vol 11993. Cham: Springer; 2020. p. 231-41. doi:10.1007/978-3-030-46640-4_22
13. Vu MH, Nyholm T, Löfstedt T. TuNet: end-to-end hierarchical brain tumor segmentation using cascaded networks. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science*, vol 11993. Cham: Springer; 2020. p. 174-86. doi:10.1007/978-3-030-46640-4_17. arXiv:1910.05338
14. Li X, Luo G, Wang K. Multi-step cascaded networks for brain tumor segmentation. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science*, vol 11993. Cham: Springer; 2020. p. 163-73. doi:10.48550/arXiv.1908.05887
15. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science*, vol 9351. Cham: Springer; 2015. p. 234-41. doi:10.1007/978-3-319-24574-4_28
16. Feng X, Tustison NJ, Patel SH, Meyer CH. Brain tumor segmentation using an ensemble of 3D U-Nets and overall survival prediction using radiomic features. *Front Comput Neurosci.* 2020;14:25. doi:10.3389/fncom.2020.00025
17. Isensee F, Kickingereder P, Wick W, Bendszus M, Maier-Hein KH. Brain tumor segmentation and radiomics survival prediction: contribution to the BRATS 2017 challenge. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science*, vol 11993. Cham: Springer; 2020. p. 163-73. doi:10.48550/arXiv.1908.05887

2017. Lecture Notes in Computer Science, vol 10670. Cham: Springer; 2018. p. 287-97. doi:10.48550/arXiv.1802.10508
18. Isensee F, Kickingereder P, Wick W, Bendszus M, Maier-Hein KH. No new-net. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lecture Notes in Computer Science, vol 11384. Cham: Springer; 2019. p. 234-44. doi:10.48550/arXiv.1809.10483
19. Myronenko A. 3D MRI brain tumor segmentation using autoencoder regularization. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lecture Notes in Computer Science, vol 11384. Cham: Springer; 2019. p. 311-20. doi:10.1007/978-3-030-11726-9_2
20. Zhao Y, Zhang YM, Liu CL. Bag of tricks for 3D MRI brain tumor segmentation. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science, vol 11993. Cham: Springer; 2020. p. 210-20. doi:10.1007/978-3-030-46640-4_20
21. McKinley R, Rebsamen M, Meier R, Wiest R. Triplanar ensemble of 3D-to-2D CNNs with label-uncertainty for brain tumor segmentation. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science, vol 11993. Cham: Springer; 2020. p. 379-87. doi:10.1007/978-3-030-46640-4_36
22. Cheng X, Jiang Z, Sun Q, Zhang J. Memory-efficient cascade 3D U-Net for brain tumor segmentation. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science, vol 11993. Cham: Springer; 2020. p. 242-53. doi:10.1007/978-3-030-46640-4_23
23. Abd-Allah MK, Awad AI, Khalaf AAM, Ibraheem AM. Automatic brain-tumor diagnosis using cascaded deep convolutional neural networks with symmetric U-Net and asymmetric residual-blocks. Sci Rep. 2024;14(1):9501. doi:10.1038/s41598-024-59566-7
24. Saluja S, Trivedi MC, Sarangdevot SS. Advancing glioma diagnosis: integrating custom U-Net and VGG-16 for improved grading in MR imaging. Math Biosci Eng. 2024;21(3):4328-50. doi:10.3934/mbe.2024191
25. Yang Y, Zhang H, Li J, Wang X, Chen L, Liu Z, et al. Segmentation method of magnetic resonance imaging brain tumor images based on improved UNet network. Transl Cancer Res. 2024;13(3):1567-83. doi:10.21037/tcr-23-1858
26. Kronberg RM, Meskelevicius D, Sabel M, Kollmann M, Rubbert C, Fischer I. Optimal acquisition sequence for AI-assisted brain tumor segmentation under the constraint of largest information gain per additional MRI sequence. Neurosci Inform. 2022;2(4):100053. doi:10.1016/j.neuri.2022.100053
27. Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. arXiv:1811.02629. 2018. doi:10.48550/arXiv.1811.02629
28. Pemberton HG, Zaki LAM, Goodkin O, Das RK, Stezoski A, Goh V, et al. Multi-class glioma segmentation on real-world data with missing MRI sequences: comparison of three deep learning algorithms. Sci Rep. 2023;13(1):18911. doi:10.1038/s41598-023-44794-0
29. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby JS, et al. Advancing the Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. Sci Data. 2017;4:170117. doi:10.1038/sdata.2017.117
30. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby J, et al. Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection [data set]. The Cancer Imaging Archive. 2017. doi:10.7937/K9/TCIA.2017.KLXWJJ1Q
31. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby J, et al. Segmentation labels for the pre-operative scans of the TCGA-GBM collection [data set]. The Cancer Imaging Archive. 2017.
32. Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In:

- Teh YW, Titterton M, editors. Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (AISTATS 2010). JMLR Workshop and Conference Proceedings. 2010;9:249-56. <https://proceedings.mlr.press/v9/glorot10a.html>.
33. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv:1502.03167*. 2015. doi:10.48550/arXiv.1502.03167
 34. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556*. 2014. doi:10.48550/arXiv.1409.1556
 35. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Miami, FL: IEEE; 2009. p. 248-55. doi:10.1109/CVPR.2009.5206848
 36. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV: IEEE; 2016. p. 770-8. doi:10.1109/CVPR.2016.90
 37. Agravat RR, Raval MS. Brain tumor segmentation and survival prediction. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science, vol 11993. Cham: Springer; 2020. p. 338-48. doi:10.48550/arXiv.1909.09399
 38. Ho Y, Wookey S. The real-world-weight cross-entropy loss function: modeling the costs of mislabeling. *IEEE Access*. 2020;8:4806-13. doi:10.48550/arXiv.2001.00570
 39. Kingma DP, Ba J. Adam: a method for stochastic optimization. *arXiv:1412.6980*. 2014. doi:10.48550/arXiv.1412.6980
 40. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37-46. doi:10.1177/001316446002000104
 41. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74. doi:10.2307/2529310
 42. Shen Y, Gao M. Brain tumor segmentation on MRI with missing modalities. *arXiv:1904.07290*. 2019. doi:10.48550/arXiv.1904.07290
 43. Rafi A, Rizvi STH, Ali H, Ahmad A, Khan ZA, Ahmed M. U-Net based glioblastoma segmentation with patient's overall survival prediction. In: Arai K, Kapoor S, Bhatia R, editors. Intelligent Computing Systems (ISICS 2020). Communications in Computer and Information Science, vol 1188. Cham: Springer; 2020. p. 22-32. doi:10.1007/978-3-030-43364-2_3
 44. Di Ieva A, Russo C, Liu S, Jian A, Rizzo G, Gagliardo C, et al. Application of deep learning for automatic segmentation of brain tumors on magnetic resonance imaging: a heuristic approach in the clinical scenario. *Neuroradiology*. 2021;63(8):1253-62. doi:10.1007/s00234-021-02649-3
 45. Ahmad P, Qamar S, Hashemi SR, Shen L. Hybrid labels for brain tumor segmentation. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science, vol 11992. Cham: Springer; 2020. p. 158-66. doi:10.1007/978-3-030-46643-5_15
 46. Mzoughi H, Njeh I, Slima MB, Ben Hamida A, Mhiri C, Mahfoudh KB. Towards a computer aided diagnosis (CAD) for brain MRI glioblastomas tumor exploration based on a deep convolutional neuronal networks (D-CNN) architectures. *Multimed Tools Appl*. 2021;80(1):899-919. doi:10.1007/s11042-020-09786-6
 47. Rehman MU, Cho S, Kim JH, Chong KT. BU-Net: brain tumor segmentation using modified U-Net architecture. *Electronics*. 2020;9(12):2203. doi:10.3390/electronics9122203