Cureus Journal of Computer Science
https://doi.org/10.7759/s44389-026-00262-0

Original Article

Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset

Tanya Foransbergher [1], Daniele Moș [2], Alex E. Debelka [3], Ana Magdalena Anghel [1]

1. *Department of Automatic Control and Industrial Informatics, Faculty of Automatic Control and Computers, National University of Science and Technology POLITEHNICA Bucharest, Bucharest, ROU*
2. *Computer Science, Independent Researcher, Erlangen, DEU*
3. *Biomedical Engineering, Université Paris Cité, Paris, FRA*

Received: May 17, 2026 | Review began: June 16, 2026 | Review ended: September 01, 2026 | Published: September 03, 2026

Abstract

Background

Hybrid deep learning pipelines combining You Only Look Once (YOLO) for lesion detection and U-Net for segmentation are increasingly applied to brain tumor MRI, yet the specific contribution of region of interest (ROI) cropping to segmentation performance remains poorly quantified. This study evaluates four ROI-based U-Net configurations for post-treatment glioma segmentation on the Brain Tumor Segmentation (BraTS) 2024 - GLI dataset, systematically varying the YOLO detector (YOLOv8m vs.YOLO11m) and the U-Net encoder backbone (ResNet-34 vs. ResNet-50).

Methods

Normalized T1 volumes and expert tumor masks were converted into axial 2D slices. Two YOLO detectors were trained for single-class tumor localization; their bounding-box predictions were used to crop 256 × 256 ROIs with adaptive padding. U-Net models with ResNet-34 and ResNet-50 encoders were then trained on each ROI set using combined Dice and binary cross-entropy loss. Performance was measured by Dice score and intersection over union (IoU) on held-out validation subjects.

Results

The best-performing configuration, YOLO11m-U-Net with a ResNet-34 encoder, achieved a validation Dice of 0.8824 and an IoU near 0.80. YOLOv8m-based ROIs with ResNet-34 produced comparable results (Dice 0.8797). ResNet-50 encoders yielded lower scores across both detectors (Dice 0.8656 and 0.8601).

Conclusions

Failure cases were predominantly linked to missed or poorly localized detections. These findings indicate that moderate-depth encoders paired with high-quality ROI generation offer the most favorable accuracy-efficiency tradeoff for post-treatment glioma segmentation in hybrid computer-aided diagnosis systems.

Categories: Health Informatics, Computer Vision, Deep Learning

Keywords: brain tumor segmentation, glioma, yolo, u-net, deep learning, region of interest, resnet, medical image analysis, brats, computer-aided diagnosis

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

Introduction

Post-treatment glioma assessment on magnetic resonance imaging (MRI) presents significant challenges due to surgical cavities, radiation-induced changes, and heterogeneous residual tumor characteristics that complicate manual contouring and introduce substantial inter-observer variability [1,2]. Gliomas represent the most common primary malignant brain tumors in adults, accounting for approximately thirty percent of all primary central nervous system tumors and up to eighty percent of malignant brain neoplasms [3]. Accurate segmentation of post-treatment tumor volumes is critical for treatment response assessment, progression monitoring, and surgical planning [2].

Deep learning architectures have transformed medical image segmentation over the past decade [3,4]. U-Net and its variants have become standard approaches for voxel-wise brain tumor segmentation, demonstrating strong performance on benchmark datasets, including the Brain Tumor Segmentation (BraTS) challenges [5,6]. In parallel, object detection frameworks such as You Only Look Once (YOLO) have been successfully adapted to medical imaging applications, offering rapid lesion localization in both two-dimensional and three-dimensional contexts [3,7].

A hybrid architecture that integrates YOLO-based detection with U-Net-based segmentation represents a promising paradigm for medical image analysis. These systems use detection to perform coarse localization of regions of interest (ROIs), then apply detailed segmentation networks to cropped regions, potentially improving computational efficiency and reducing false positives in background tissue. Several recent studies have reported successful implementation of YOLO-U-Net pipelines for brain tumor detection and segmentation [8-10].

However, most published work presents only end-to-end performance metrics without systematically examining how different components of the hybrid pipeline contribute to final segmentation accuracy, particularly in the context of post-treatment imaging, where anatomical distortion and treatment-related changes complicate both detection and segmentation tasks.

The BraTS 2024 - GLI challenge provides a large-scale standardized dataset of post-treatment multiparametric MRI with expert voxel-wise annotations, including explicit labeling of resection cavities as distinct regions. This dataset offers a realistic test bed for methodological investigations of training strategies in complex clinical scenarios [2,11].

This study addresses the following research question: how do detector architecture and encoder depth interact within a ROI-based YOLO-U-Net pipeline for post-treatment glioma segmentation? While hybrid YOLO-U-Net pipelines have been explored in prior work [8-10], existing studies typically evaluate a single end-to-end configuration, making it difficult to disentangle each component's contribution to overall performance. We address this gap through a controlled factorial comparison: four configurations formed by crossing two YOLO detectors (YOLOv8m and YOLO11m) with two U-Net encoder backbones (ResNet-34 and ResNet-50) are evaluated under identical experimental conditions. Additionally, we include a full-image segmentation baseline (without ROI cropping) to directly quantify the benefits of detection-guided training. We report segmentation performance with statistical significance testing, analyze convergence behavior, and characterize failure modes to provide practical guidance for hybrid computer-aided diagnosis (CAD) system design.

Materials And Methods

Dataset and preprocessing

This study utilized the BraTS 2024 - GLI (Glioma Segmentation on Post-treatment MRI) dataset, which comprises multiparametric brain MRI volumes acquired after surgical resection, radiotherapy, or chemotherapy, accompanied by expert voxel-wise segmentations [2,11]. Each case includes T1-weighted, contrast-enhanced T1-weighted (T1Gd), T2-weighted, and T2 fluid-attenuated inversion recovery (FLAIR) sequences, along with segmentation masks delineating enhancing tumor, non-enhancing tumor core, peritumoral edema, necrosis, and resection cavity as separate labels [1,2].

For this methodological comparison, we used only the normalized non-contrast T1-weighted sequence (T1n) as the imaging input. All tumor-related labels (enhancing tumor, edema, necrosis, and resection cavity) were merged into a single binary tumor class, with voxels labeled as 1 for tumor and 0 for background. This simplification enabled direct

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

comparison of configurations without confounding effects from multi-class balancing strategies.

Volumetric data were provided as NIfTI (.nii) files. For each subject, the T1n volume and corresponding segmentation volume were loaded using the NiBabel library. Axial 2D slices were extracted along the superior-inferior axis. Each slice underwent per-slice intensity normalization using min-max scaling to a range of 0 to 1 and was resized to 256 × 256 pixels using bilinear interpolation for images and nearest-neighbor interpolation for binary masks to preserve label integrity. Normalized slices were saved as 8-bit PNG images.

Subjects were divided into training and validation sets at the patient level to prevent data leakage. The same training-validation split was maintained across all experiments to ensure fair comparison. Approximately 273,657 axial slices were used for training and 21,365 slices for validation.

YOLO detection models and ROI generation

Two modern YOLO architectures, YOLOv8m and YOLO11m, were trained for single-class tumor detection on 2D slices [12]. Bounding box annotations were automatically generated from binary tumor masks by computing the minimal axis-aligned rectangle enclosing all tumor pixels in each slice. Coordinates were normalized to image dimensions and stored in YOLO text format (class index, x-center, y-center, width, height). Slices without tumors were retained as negative examples without annotation files.

Detection models were trained on an NVIDIA GPU using pre-trained COCO weights for transfer learning, with an image size of 640 × 640 pixels, a batch size of 16, and 50 training epochs. Data augmentation included random horizontal flips, slight rotation, and brightness adjustments. Performance was evaluated on the held-out validation set using precision, recall, mean average precision (mAP) at an Intersection over Union (IoU) threshold of 0.5 (mAP@0.5), and mAP at IoU thresholds from 0.5 to 0.95 (mAP@0.5:0.95) [13].

Both detectors achieved validation precision above 0.95, recall above 0.85, and mAP@0.5 above 0.91, indicating reliable tumor localization. Trained models were applied to all training and validation slices to generate ROIs for segmentation experiments. For each slice, bounding boxes with confidence scores above a threshold of 0.5 were used to crop square ROIs from both the original T1n image and the corresponding ground-truth tumor mask. A padding margin of 20 pixels was added around each detection to preserve peritumoral context and avoid truncating tumor boundaries.

Cropped ROI image-mask pairs were resized to 256 × 256 pixels and stored in separate caches for YOLOv8m-derived ROIs and YOLO11m-derived ROIs. Metadata files recorded subject identifier, slice index, bounding box coordinates, detection confidence, and crop dimensions. Slices without sufficiently confident detections (confidence < 0.5) were excluded from ROI-based training and logged separately for error analysis.

U-Net architecture and training configurations

A two-dimensional U-Net architecture with an encoder-decoder structure and skip connections served as the segmentation backbone for all experiments [14,15]. Two encoder backbones were evaluated: ResNet-34 and ResNet-50, both pre-trained on ImageNet [14]. The choice of these encoders was motivated by the need to assess whether increased encoder depth (50 layers versus 34 layers) translates into improved segmentation performance when applied to YOLO-derived ROIs, given the reduced spatial context and increased tumor-to-background ratio of cropped inputs. The decoder consisted of four up-sampling blocks with transposed convolutions and concatenation of skip connections from corresponding encoder levels. The final segmentation head comprised a 1 × 1 convolutional layer with sigmoid activation, producing a single-channel probability map for tumor versus background classification.

Four experimental configurations were defined by crossing the two YOLO detectors with the two encoder backbones:

(1) YOLOv8m ROIs + ResNet-34 encoder; (2) YOLOv8m ROIs + ResNet-50 encoder;

(3) YOLO11m ROIs + ResNet-34 encoder; (4) YOLO11m ROIs + ResNet-50 encoder.

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

ROI-Based U-Net: The network was trained exclusively on 256 × 256 ROI crops generated by YOLO detectors, where each ROI is centered on a detected tumor with added padding. Ground-truth masks were cropped identically to ensure perfect alignment between input images and targets.

Additionally, a full-image baseline was trained using an identical U-Net architecture (ResNet-34 encoder) on complete 256 × 256 axial slices without YOLO-derived ROI cropping. This baseline used the full training/validation slice set (227,504/18,196 slices) rather than the detection-positive subset used for ROI-based training, and shared the same optimizer, loss function, and training schedule as the ROI-based configurations, isolating the effect of ROI-based cropping from other pipeline components.

All configurations used the same loss function, optimizer, and training schedule. The loss function combined Dice loss and binary cross-entropy (BCE) loss with equal weighting:

$$L = L_{\text{Dice}} + L_{\text{BCE}} \tag{1}$$

where L_{Dice} is the Dice loss, which emphasizes overlap between the prediction and ground truth, and L_{BCE} (BCE loss) penalizes pixel-wise classification errors. This compound loss has been shown to improve convergence and boundary precision in medical image segmentation tasks [16].

The optimizer used for all experiments was AdamW (to minimize the compound loss) with an initial learning rate of 1×10^{-3}, consistent across all configurations. All U-Net models were trained for 25 epochs, with early stopping based on validation loss.

To stabilize training and leverage pre-trained encoder weights, the encoder was frozen during the initial epochs while only the decoder and segmentation head were updated, after which the entire network was unfrozen for end-to-end fine-tuning. Binary segmentation masks were obtained by thresholding the sigmoid probability outputs. All four configurations shared an identical optimizer, loss function, and training schedule so that observed performance differences could be attributed solely to detector architecture and encoder depth. The specific hyperparameter values are summarized in Table *1* (Section "Experimental Setup").

Evaluation metrics and statistical analysis

Segmentation performance was quantified using two standard overlap metrics computed at the pixel level on held-out validation subjects:

1. Dice Similarity Coefficient (DSC)

$$\text{DSC} = \frac{2\,|P \cap G|}{|P| + |G|} \tag{2}$$

where P is the set of pixels predicted as tumor by the segmentation network, G is the set of pixels labeled as tumor in the ground-truth mask, $|P \cap G|$ is the number of pixels common to both (true positives), and $|P|$ and $|G|$ are the total numbers of pixels in the predicted and ground-truth masks, respectively. Dice ranges from 0 (no overlap) to 1 (perfect agreement) [16].

2. IoU, also known as Jaccard index

$$\text{IoU} = \frac{|P \cap G|}{|P \cup G|} \tag{3}$$

where $|P \cap G|$ is the number of pixels in the intersection of the predicted and ground-truth tumor regions, and $|P \cup G|$ is the number of pixels in their union. IoU is more sensitive to segmentation errors than Dice and provides complementary information about prediction quality [16].

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

Binary predictions were obtained by applying a threshold of 0.5 to the sigmoid probability outputs. Metrics were computed for each validation slice and aggregated across all slices to obtain mean and standard deviation. Training dynamics were monitored by recording loss and Dice score trajectories across epochs.

Qualitative analysis included visual inspection of representative high-performing and failure cases. For error characterization in ROI-based regimes, we computed the ratio of predicted tumor area to ground-truth tumor area and categorized extreme cases such as: (1) under-segmentation (area ratio < 0.2), (2) over-segmentation (area ratio > 10), and (3) complete miss (Dice = 0). Failure modes were cross-referenced with detection quality by examining whether corresponding YOLO detections were correct, partial, or absent.

To assess whether observed performance differences are statistically significant, we conducted non-parametric statistical tests on per-slice Dice scores computed across the validation set. A Friedman test was applied to compare all four configurations simultaneously on the subset of slices common to all configurations (n = 15,002).

Because each detector produces a different number of valid ROI detections (confidence ≥ 0.5), configuration-specific validation sets differ in size (N = 15,280 for YOLO11m-derived ROIs; N = 15,630 for YOLOv8m-derived ROIs, as reported in Table *4*). The Friedman test requires paired observations across all four configurations and was therefore restricted to the subset of slices with valid detections common to all four configurations (n = 15,002).

For pairwise comparisons, Wilcoxon signed-rank tests were used to evaluate: (a) the effect of encoder depth (ResNet-34 vs. ResNet-50) within each detector, and (b) the effect of detector architecture (YOLO11m vs. YOLOv8m) within each encoder. All tests were two-sided, with a significance level of $\alpha = 0.05$. Statistical analyses were performed using SciPy (version 1.17.0).

Results And Discussion

Experimental setup

Table *1* summarizes the training configuration shared by all four ROI-based U-Net experiments. Detection and segmentation hyperparameters were held constant across configurations to isolate the effects of detector architecture and encoder depth.

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

Parameter	Value
Framework	PyTorch (Ultralytics, Segmentation Models PyTorch)
Hardware	NVIDIA GPU (CUDA)
Training/validation slices	273,657/21,365 (patient-level split)
YOLO input resolution/epochs	640 × 640/50
U-Net input/epochs	256 × 256 ROI crops (single-channel T1n)/25
Optimizer	AdamW
Learning rate	1×10^{-3}
Batch size	16
Encoder freezing	5 epochs, then unfrozen
Loss function	Dice + BCE (equal weighting)
Encoder weights	ImageNet – pretrained
Binarization threshold	0.5

TABLE 1: Training configuration

Detection performance

Both detectors achieved validation precision above 0.95, recall above 0.85, and mAP@0.5 above 0.91, indicating reliable tumor localization (Table *2*). YOLOv8m demonstrated a slight advantage in recall (0.862 vs. 0.851) and mAP@0.5 (0.926 vs. 0.918), while YOLO11m showed marginally higher box loss values, reflecting slightly less precise bounding-box regression.

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95	Box Loss
YOLOv8m	0.956	0.862	0.926	0.747	0.795
YOLO11m	0.953	0.851	0.918	0.736	0.810

TABLE 2: Comparison of detection performance on BraTS 2024 - GLI validation set

BraTS, Brain Tumor Segmentation; GLI, Glioma; mAP, mean Average Precision

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

Segmentation performance comparison

Table *3* summarizes the validation performance of all four ROI-based U-Net configurations on the BraTS 2024 - GLI validation set. The results are organized by YOLO detector and encoder backbone to facilitate analysis of both factors. The full-image baseline was trained on all 256 × 256 axial slices without ROI cropping, using identical hyperparameters to the ROI-based configurations.

Configuration	Encoder	Size	Number of Epochs With Frozen Encoder	Best Epoch	Best Dice Similarity Coefficient	Best IoU
YOLO11m + U-Net	ResNet-34	256	5	23	0.8824	~0.80
YOLOv8m + U-Net	ResNet-34	256	5	25	0.8797	~0.80
YOLOv8m + U-Net	ResNet-50	256	5	23	0.8656	~0.77
YOLO11m + U-Net	ResNet-50	256	5	24	0.8601	~0.77
Full-image baseline (no ROI)	ResNet-34	256	5	22	0.6625	~0.57

TABLE 3: Comparison of segmentation performance across training regimes on BraTS 2024 - GLI validation set

BraTS, Brain Tumor Segmentation; GLI, Glioma; IoU, Intersection over Union; ROI, Region of Interest

The best segmentation performance was achieved by the YOLO11m-U-Net configuration with a ResNet-34 encoder, reaching a validation Dice score of 0.8824 at epoch 23. The YOLOv8m-U-Net with a ResNet-34 encoder achieved closely comparable performance (Dice 0.8797 at epoch 25), indicating that both detectors produce ROIs of sufficient quality to support high-accuracy segmentation when paired with a moderate-depth encoder.

A consistent pattern emerged across both detectors: configurations using ResNet-34 encoders outperformed those using ResNet-50 by approximately 1.4-2.2 percentage points in Dice score. The YOLOv8m-ResNet-50 configuration achieved Dice 0.8656, while YOLO11m-ResNet-50 reached 0.8601, representing the lowest performance among the four configurations.

This finding suggests that the increased capacity of ResNet-50 did not translate into improved segmentation accuracy under the given training conditions, potentially due to overfitting on the relatively constrained ROI inputs or insufficient training duration to fully exploit the deeper architecture. These results demonstrate that restricting U-Net training to tumor-centered crops provides consistent and measurable improvements in segmentation accuracy.

Statistical Significance Analysis

To verify that the observed differences are not attributable to chance, we conducted statistical tests on per-slice Dice scores. Table *4* summarizes the descriptive statistics for each configuration. Because each detector generates a different number of valid ROI detections (confidence ≥ 0.5), the four configurations do not share an identical validation slice set. Per-configuration descriptive statistics (Table *4*) use each configuration's full validation set (N = 15,280 for YOLO11m-derived ROIs; N = 15,630 for YOLOv8m-derived ROIs). The Friedman test, which requires paired observations across all four configurations, was restricted to the subset of slices with valid detections in all four configurations (n = 15,002).

Configuration	Mean Dice	Standard Deviation	Median	N
YOLO11m + ResNet-34	0.8451	0.1196	0.8802	15,280
YOLOv8m + ResNet-34	0.8496	0.1182	0.8843	15,630
YOLOv8m + ResNet-50	0.7822	0.1667	0.8350	15,630
YOLO11m + ResNet-50	0.7587	0.1749	0.8108	15,280

TABLE 4: Per-slice descriptive statistics on the BraTS 2024 - GLI validation set

BraTS, Brain Tumor Segmentation; GLI, Glioma

The Friedman test confirmed a statistically significant difference across all four configurations ($\chi^2 = 19{,}973.21$, $p < 0.001$, n = 15,002 common slices). Pairwise Wilcoxon signed-rank tests (Table *5*) revealed that the encoder effect was both statistically significant and practically substantial: ResNet-34 outperformed ResNet-50 by a median of 3.8-5.5 Dice percentage points ($p < 0.001$) for both detectors. The detector effect (YOLO11m vs. YOLOv8m) was statistically significant but of smaller magnitude (median difference 0.4-1.8 percentage points, $p < 0.001$), reinforcing the conclusion that encoder depth has a larger practical impact on segmentation quality than detector choice.

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

Comparison	Mean A	Mean B	Median Difference	95% CI	r	p-value
ResNet-34 vs. 50 (YOLO11m)	0.8451	0.7587	+0.0545	[0.0532, 0.0560]	0.741	< 0.001 ***
ResNet-34 vs. 50 (YOLOv8m)	0.8496	0.7822	+0.0379	[0.0368, 0.0391]	0.707	< 0.001 ***
YOLO11m vs. YOLOv8m (R-34)	0.8490	0.8585	−0.0037	[−0.0041, −0.0032]	0.169	< 0.001 ***
YOLO11m vs. YOLOv8m (R-50)	0.7637	0.7945	−0.0182	[−0.0195, −0.0172]	0.323	< 0.001 ***

TABLE 5: Pairwise Wilcoxon signed-rank test results

Effect sizes were computed as $r = Z/\sqrt{N}$ for all Wilcoxon tests and confirmed the practical significance of the encoder effect: r = 0.707-0.741 (large effect), compared to the detector effect (r = 0.169-0.323, small-to-medium effect). 95% bootstrap confidence intervals (10,000 resamples) for the median differences excluded zero in all four comparisons, confirming robust and reproducible effects.

Means in Table *5* are calculated on the paired subset used for each Wilcoxon comparison and therefore may differ from the configuration-wide means in Table *4*. The reported median difference is the median of the paired, slice-level differences and should not be interpreted as the difference between the two displayed means.

Comparison With Full-Image Baseline

To quantify the benefit of ROI-based training, a full-image U-Net with a ResNet-34 encoder was trained on all 256 × 256 axial slices without detection-guided cropping, using identical hyperparameters. The full-image baseline achieved a validation Dice of 0.6625, compared to 0.8824 (batch-averaged) and 0.8451 (per-slice mean) for the best ROI-based configuration (YOLO11m + ResNet-34). This difference of 22.0 percentage points (batch-averaged comparison) or 18.3 percentage points (per-slice mean comparison) confirms that restricting training to tumor-centered ROIs improves segmentation accuracy by reducing background dominance and allowing the network to focus on fine-grained tumor boundary delineation.

Training dynamics and convergence

Figure *1* illustrates the training dynamics for the best segmentation performance. The validation loss decreased steadily from initial values around 0.21 to approximately 0.12 in the final epoch, indicating efficient optimization of the compound Dice + BCE loss function and strong correlation between loss minimization and metric improvement.

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

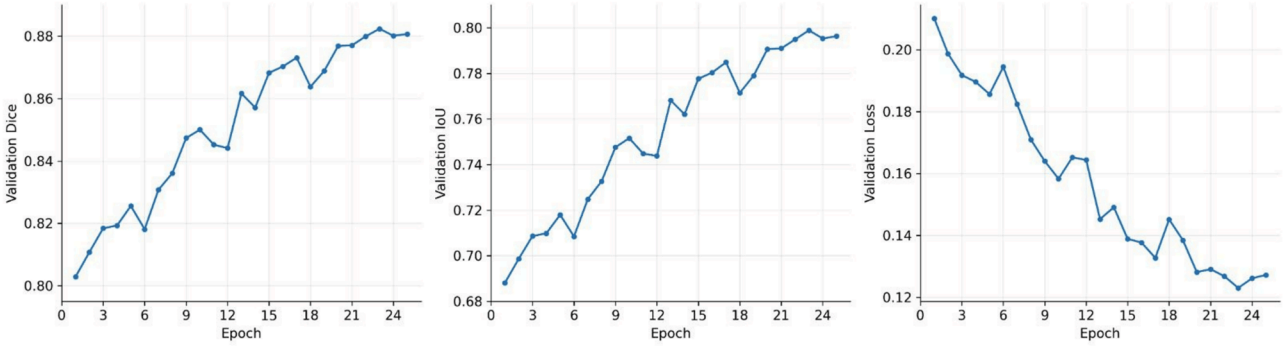

FIGURE 1: Validation Dice similarity coefficient (Dice)/Intersection over Union (IoU), and loss curves across epochs for YOLO11m-U-Net ResNet-34

Validation Dice score increased progressively from around 0.80 at early epochs to peak values near 0.88, with minimal overfitting. Stabilization of both Dice and IoU after approximately 20 epochs suggests that the model reached a generalization plateau without evidence of overfitting, confirming the robustness of the ROI-based training approach in the context of MRI segmentation.

ResNet-50 configurations exhibited slightly slower convergence and greater fluctuation in validation metrics during intermediate training epochs, consistent with the hypothesis that deeper encoders require more training data or longer schedules to realize their full potential on cropped ROI inputs.

Error analysis and failure modes

Visual inspection of segmentation outputs confirmed that all four configurations produced coherent tumor contours in many cases, with Dice scores exceeding 0.96 and IoU above 0.92 in the best individual cases. Limitations were observed primarily for very small or diffuse tumors, where the transition between tumoral and healthy tissue is poorly defined, leading to under-segmentation or over-segmentation (Figure *2*).

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

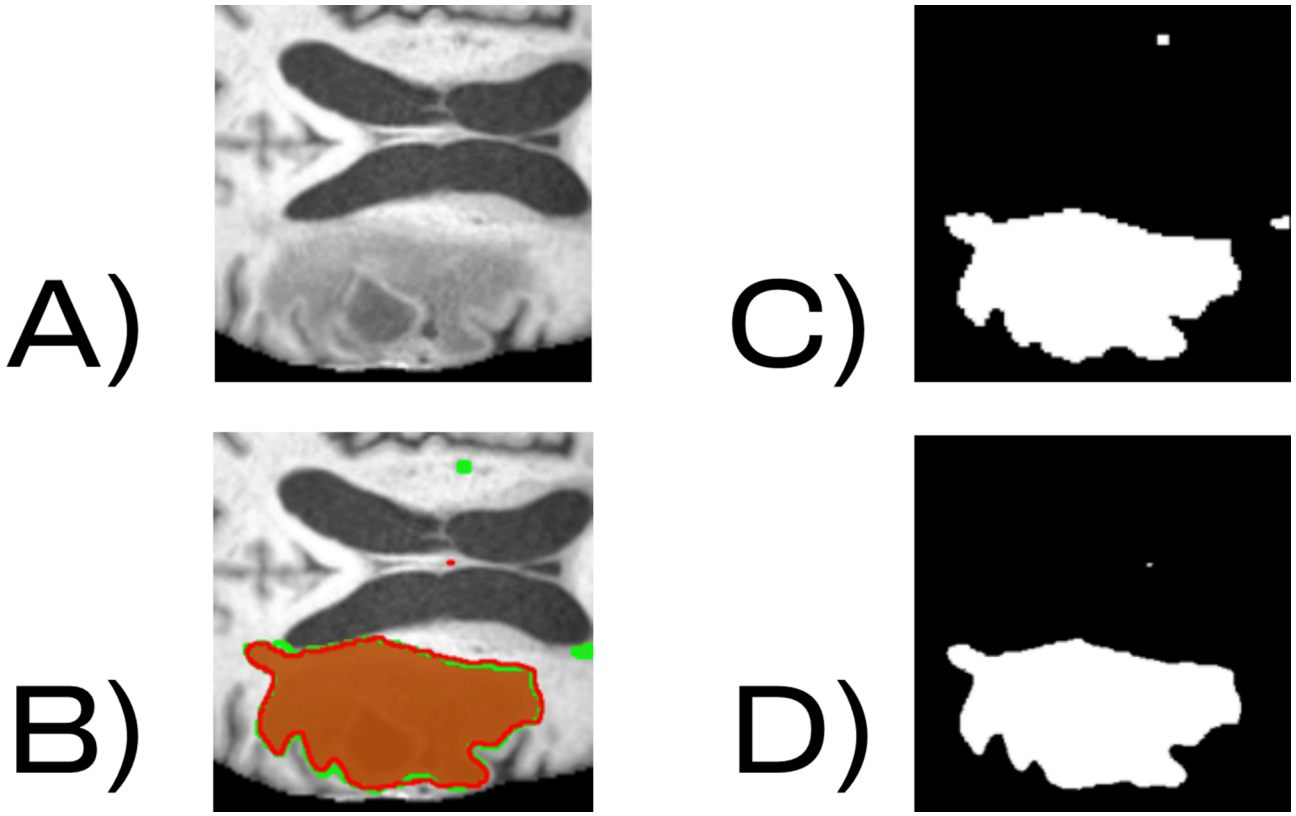

FIGURE 2: A) Region of interest from original image. B) Ground-truth mask. C) Overlay of the original image and the resulting segmentation. D) Mask generated by the segmentation model

However, ROI-based training introduced a new failure mode: complete segmentation failure (Dice = 0) in cases where YOLO detectors either missed the lesion entirely or produced severely misaligned bounding boxes.

Qualitative examination of best-case examples showed near-perfect contour agreement between predictions and ground truth in ROI-based training, particularly for compact, well-defined tumors. Worst-case examples typically involved very small lesions (< 50 pixels), low-contrast diffuse infiltration, or ROIs that did not fully encompass the tumor due to undersized detection boxes (Table *6*).

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

Case Type	Dice	IoU	Area Ratio	Interpretation
Best	0.9822	0.9650	0.9989	Near-perfect overlap
Worst	0.0000	0.0000	0.2984	Complete miss
Under-segmentation	0.1306	0.0698	0.0808	Partial coverage only
Over-segmentation	0.0874	0.0457	11.2822	Excessive expansion

TABLE 6: Performance metrics for selected edge cases (YOLO11m-U-Net ResNet-34 configuration)

Dice refers to Dice Similarity Coefficient, and Area Ratio refers to the predicted tumor area/ground-truth tumor area.
IoU, Intersection over Union

The worst case (Dice 0.0000) corresponded to a complete segmentation failure, typically associated with absent or severely misaligned YOLO detections. Under-segmentation cases were characterized by partial tumor coverage, particularly in zones with diffuse margins or low contrast. Over-segmentation was observed when predicted masks extended significantly beyond tumor boundaries, often in the presence of intensity artifacts or textures resembling tumoral tissue.

This observation confirms that detection quality is the primary bottleneck for overall system performance and supports prioritizing high-recall detectors in hybrid pipeline design.

Discussion

This study presents a systematic comparison of four ROI-based YOLO-U-Net configurations for post-treatment glioma segmentation on BraTS 2024, varying both the detection backbone and the segmentation encoder. The results reveal two principal findings with practical implications for hybrid CAD system design.

First, ResNet-34 consistently outperformed ResNet-50 across both detectors. This may appear counterintuitive, as deeper networks typically offer greater representational capacity. However, in ROI-based training, where inputs are cropped to tumor-centered regions, the reduced spatial context and limited training set size may favor architectures that generalize more efficiently with fewer parameters. Without sufficient data or training duration, the additional parameters of ResNet-50 may introduce overfitting rather than improved feature extraction, consistent with observations that moderate-depth encoders often suffice for constrained medical segmentation tasks [5,6].

Second, both YOLO detectors produced comparable downstream segmentation despite measurable differences in detection metrics. This near equivalence suggests that beyond a quality threshold, further improvements in detection precision yield diminishing returns for segmentation accuracy. The critical factor is whether detectors provide sufficiently accurate bounding boxes to enable the segmentation network to focus on tumor delineation rather than background discrimination.

Our findings extend prior YOLO-U-Net work by providing a controlled factorial comparison of detector and encoder choices, supported by formal statistical testing. Previous studies have typically evaluated a single YOLO version paired with a single U-Net configuration, making it difficult to disentangle each component's contribution.

The Wilcoxon signed-rank tests confirm that the encoder depth effect (median Dice difference of 3.8-5.5 percentage points) substantially exceeds the detector effect (0.4-1.8 percentage points), providing quantitative evidence for design priorities in hybrid CAD systems. Furthermore, the comparison with a full-image baseline (Dice 0.6625 vs. 0.8824 for ROI-

How to cite this article:
Foransbergher T, Moş D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

based) directly demonstrates that detection-guided cropping is not merely a computational convenience but a substantive contributor to segmentation quality.

YOLOv8m and YOLO11m yielded similar downstream Dice despite small mAP differences, and moderate-depth encoders such as ResNet-34 were sufficient, which is encouraging for resource-constrained settings. Several limitations should be acknowledged.

First, the use of single-modal non-contrast T1-weighted images, while enabling controlled comparison, excludes complementary information from T1Gd, T2, and FLAIR sequences that could improve segmentation accuracy in clinical settings. Second, the 2D slice-wise approach discards volumetric spatial information and inter-slice continuity that 3D architectures could exploit. Third, evaluation on a single public dataset (BraTS 2024 - GLI) limits generalizability claims; external validation on independent clinical cohorts with different acquisition protocols is needed before clinical deployment. Fourth, the training schedule of 25 epochs may have been insufficient for ResNet-50 to reach its optimal performance on cropped ROI inputs, potentially underestimating deeper encoders' capacity.

Finally, while all pairwise comparisons reached statistical significance ($p < 0.001$), the large sample size (>15,000 slices) means that even small differences become significant; the practical relevance of the detector effect (median Dice difference of 0.4 percentage points for ResNet-34) warrants cautious interpretation. Future work should address multimodal integration, 3D architectures, uncertainty-aware ROI selection, and external validation on independent clinical cohorts.

Conclusions

This study demonstrates that within a ROI-based YOLO-U-Net pipeline for post-treatment glioma segmentation, encoder depth and detector architecture interact meaningfully with segmentation quality. ResNet-34 encoders consistently outperformed ResNet-50 across both YOLOv8m and YOLO11m detectors, achieving a best validation Dice score of 0.8824 with YOLO11m-derived ROIs. The two YOLO detectors produced comparable downstream segmentation accuracy despite differences in detection metrics, suggesting that ROI quality above a certain threshold has diminishing impact on segmentation outcomes. The benefits of ROI-based training include faster convergence, reduced extreme over-segmentation, and more stable training dynamics. However, ROI-based training propagates detection errors, causing segmentation failure when tumors are missed or poorly localized by the detector.

The systematic comparison of four configurations revealed that encoder depth has a larger impact on segmentation quality than detector choice. These findings support the value of hybrid detection-segmentation pipelines and highlight the importance of high-quality tumor localization as a prerequisite for effective ROI-focused segmentation. Future work should investigate joint optimization strategies and uncertainty-aware ROI selection to maximize the advantages of hybrid approaches while mitigating their failure modes.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Acquisition, analysis, or interpretation of data: Alex E. Debelka, Tanya Foransbergher, Daniele Moș

Drafting of the manuscript: Alex E. Debelka, Tanya Foransbergher, Ana Magdalena Anghel

Critical review of the manuscript for important intellectual content: Alex E. Debelka, Tanya Foransbergher, Daniele Moș, Ana Magdalena Anghel

Supervision: Alex E. Debelka, Daniele Moș, Ana Magdalena Anghel

Concept and design: Tanya Foransbergher, Daniele Moș, Ana Magdalena Anghel

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue.

Animal subjects: All authors have confirmed that this study did not involve animal subjects or tissue.

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following:

- **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work.
- **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work.
- **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The datasets (and/or code) supporting this study are available from the corresponding author upon reasonable request. Data used in this publication were obtained as part of the Brain Tumor Segmentation (BraTS) Challenge project through Synapse ID: syn53708249.

Acknowledgements

The authors thank the organizers of the BraTS 2024 challenge for curating and providing access to the post-treatment glioma dataset.

References

1. Aboian M, Anazodo U, Baid U, et al.: MICCAI 2025 Lighthouse Challenge: brain tumor segmentation cluster of challenges (BraTS). Zenodo. International Conference on Medical Image Computing and Computer Assisted Intervention 2025 (MICCAI). 2024, 10.5281/zenodo.13981216
2. Correia de Verdier M, Saluja R, Gagnon L, et al.: The 2024 Brain Tumor Segmentation (BraTS) challenge: glioma segmentation on post-treatment MRI [PREPRINT]. arXiv. 2024, 10.48550/arXiv.2405.18368
3. Bouhafra S, El Bahi H: Deep learning approaches for brain tumor detection and classification using MRI images (2020 to 2024): a systematic review. Journal of Imaging Informatics in Medicine. 2024, 38:1403-1433. 10.1007/s10278-024-01283-8
4. Das S, Goswami RS: Advancements in brain tumor analysis: a comprehensive review of machine learning, hybrid deep learning, and transfer learning approaches for MRI-based classification and segmentation. Multimedia Tools and Applications. 2024, 84:26645-26682. 10.1007/s11042-024-20203-0
5. Umarani CM, Gollagi SG, Allagi S, Sambrekar K, Ankali SB: Advancements in deep learning techniques for brain tumor segmentation: A survey. Informatics in Medicine Unlocked. 2024, 50:101576. 10.1016/j.imu.2024.101576
6. Jiang H, Diao Z, Yao YD: Deep learning techniques for tumor segmentation: a review. The Journal of Supercomputing. 2022, 78:1807-1851. 10.1007/s11227-021-03901-6
7. Almufareh MF, Imran M, Khan A, Humayun M, Asim M: Automated brain tumor segmentation and classification in MRI using YOLO-based deep learning. IEEE Access. 2024, 12:16189-16207. 10.1109/access.2024.3359418
8. Iriawan N, Pravitasari AA, Nuraini US, et al.: YOLO-UNet architecture for detecting and segmenting the localized MRI brain tumor image. Applied Computational Intelligence and Soft Computing. 2024, 2024:1-14. 10.1155/2024/3819801
9. Ahsan R, Shahzadi I, Najeeb F, Omer H: Brain tumor detection and segmentation using deep learning. Magnetic Resonance Materials in Physics, Biology and Medicine. 2025, 38:13-22. 10.1007/s10334-024-01203-5
10. Zafar W, Husnain G, Iqbal A, et al.: Enhanced TumorNet: Leveraging YOLOv8s and U-net for superior brain tumor detection and segmentation utilizing MRI scans. Results in Engineering. 2024, 24:102994. 10.1016/j.rineng.2024.102994

11. BraTS 2024 Challenge. (2024). Accessed: January 17, 2026: https://www.synapse.org/Synapse:syn53708249/wiki/626323.

12. Sapkota R, Karkee M: Ultralytics YOLO evolution: an overview of YOLO26, YOLO11, YOLOv8 and YOLOv5 object detectors for computer vision and pattern recognition [PREPRINT]. arXiv. 2026, 10.48550/arXiv.2510.09653

13. Ultralytics: Performance metrics deep dive. (2023). Accessed: January 14, 2026: https://docs.ultralytics.com/guides/yolo-performance-metrics/.

14. Azad R, Khodapanah Aghdam E, Rauland A, et al.: Medical image segmentation review: the success of U-Net. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2024, 46:10076-10095. 10.1109/tpami.2024.3435571

15. Jiangtao W, Ruhaiyem NIR, Panpan F: A comprehensive review of U-Net and its variants: advances and applications in medical image segmentation. IET Image Processing. 2025, 19:e70019. 10.1049/ipr2.70019

16. Alirr O, Alshatti R, Altemeemi S, Alsaad S, Alshatti A: Automatic brain tumor segmentation from MRI scans using U-Net deep learning. 2023 5th International Conference on Bio-engineering for Smart Technologies (BioSMART), Paris, France. 2023, 1-5. 10.1109/BioSMART58455.2023.10162093

How to cite this article:
Foransbergher T, Moș D, Debelka A E, et al. (September 03, 2026) Impact of You Only Look Once (YOLO)-Derived Regions of Interest on U-Net Segmentation of Post-Treatment Glioma in Magnetic Resonance Imaging: A Methodological Study on the BraTS (Brain Tumor Segmentation) 2024 Dataset. Cureus J Comput Sci 3 : es44389-026-00262-0. DOI https://doi.org/10.7759/s44389-026-00262-0