

Article

M²-YOLO: Multi-Task Instance Segmentation and Non-Invasive Molecular Profiling of Gliomas via Preoperative MRI

Skandar Hadj Abdallah  and Moulay A. Akhloufi * 

Perception, Robotics and Intelligent Machines Lab (PRIME), Université de Moncton,
Moncton, NB E1A 3E9, Canada; esh7911@umoncton.ca

* Correspondence: moulay.akhloufi@umoncton.ca

Abstract

Gliomas are the most common malignant primary brain tumors. The 2021 WHO classification requires the integration of molecular biomarkers, including IDH mutation, MGMT promoter methylation, and 1p/19q codeletion, alongside histological criteria, yet their assessment still relies on invasive surgical biopsy. We present M²-YOLO (Multi-task Molecular YOLO), a unified deep learning framework that simultaneously performs glioma instance segmentation and non-invasive molecular biomarker prediction from preoperative multi-sequence MRI. The proposed approach was evaluated on the multi-site UTSW-Glioma dataset comprising 625 patients. M²-YOLO achieved Dice scores of 0.853 for edema and 0.697 for tumor core segmentation, together with AUC-ROC values of 0.849 for IDH mutation and 0.718 for 1p/19q codeletion prediction. MGMT prediction yielded an AUC of 0.520, consistent with the limited discriminative capacity of anatomical MRI for this biomarker. These results demonstrate the potential of a unified multi-task framework for combining tumor delineation and molecular characterization within a single non-invasive workflow, supporting future applications in precision neuro-oncology.

Keywords: gliomas; instance segmentation; molecular classification; IDH; MGMT; 1p/19q; YOLO; multi-sequence MRI; multi-site

1. Introduction

Gliomas are the most common primary malignant brain tumors, accounting for approximately 80% of malignant neoplasms of the central nervous system, with an average annual incidence rate of 7.06 per 100,000 population [1]. The World Health Organization (WHO) 2021 classification of central nervous system tumors has fundamentally redefined the diagnosis of gliomas by making the integration of molecular biomarkers into histological criteria mandatory [2]. In particular, isocitrate dehydrogenase (IDH) mutation status, O6-methylguanine-DNA methyltransferase (MGMT) promoter methylation, and 1p/19q codeletion have become indispensable for accurate tumor classification, prognostic stratification, and treatment planning.

Recent studies matched with propensity scores have shown that although overall survival is similar, mol-GBM has significantly better progression-free survival than histological GBM, particularly in patients with unmethylated MGMT [3]. Furthermore, in elderly patients, composite scores such as ACCI (Age-adjusted Charlson Comorbidity Index) and the OFO classification (Onco-Functional Outcome, incorporating postoperative functional status) have been shown to be independent predictors of survival [4,5], showing the importance of non-invasive tools for preoperative stratification.



Academic Editor: Alexander Zhbanov

Received: 27 June 2026

Revised: 19 July 2026

Accepted: 28 July 2026

Published: 3 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\) license](#).

Despite their clinical importance, these molecular markers are currently obtained exclusively by invasive surgical biopsy, exposing patients to procedural risks such as neurological deficits, bleeding, and infection while imposing diagnostic delays of 7 to 14 days that delay critical clinical decision-making. A recent longitudinal study conducted at 21 UK centers covering 84% of the population revealed that the average time to molecular diagnosis increased from 16 to 21 days between 2021 and 2024, with only 30% of centers meeting the 14-day target [6]. The development of non-invasive methods for predicting molecular biomarkers using imaging is, therefore, a major priority in neuro-oncology.

Magnetic resonance imaging (MRI) contains detailed morphological and functional information on the phenotype of gliomas. Radiogenomics, the study of associations between imaging characteristics and genomic profiles, has demonstrated that molecular subtypes exhibit distinct visual signatures on multi-sequence MRI. As recently summarized by Christodoulou et al. [7], artificial intelligence approaches, including radiomics and deep learning, have demonstrated AUCs ranging from 0.80 to 0.95 for the non-invasive prediction of molecular markers, outperforming traditional radiomic approaches.

A systematic review by Contreras et al. [8] confirmed that AI models achieve AUCs greater than 0.97 in distinguishing GBM from other intracranial tumors, often outperforming expert radiologists. IDH-mutated gliomas typically present as well-defined homogeneous masses with little necrosis, whereas wild-type IDH glioblastomas exhibit heterogeneous enhancement, central necrosis, and perilesional edema features that can be directly exploited by deep learning models. While recent studies have proposed unified multi-task frameworks for jointly performing glioma segmentation and molecular characterization, most existing approaches rely on computationally intensive encoder–decoder architectures or two-stage pipelines, limiting their suitability for real-time clinical deployment. Moreover, their performance is often evaluated on highly curated benchmark datasets such as BraTS [9], which do not capture the substantial inter-scanner variability encountered in real-world multi-site clinical deployment.

In this work, we present M²-YOLO (Multi-task Molecular YOLO), an end-to-end framework that simultaneously performs glioma instance segmentation and molecular biomarker classification from preoperative multi-sequence MRI in a single 3.8 ms inference. Built on the YOLO26m-seg architecture [10], M²-YOLO incorporates a shared backbone with two output branches: a segmentation head that delineates the necrotic core (TC) and edema (ED) as distinct instances, and three molecular classification heads that predict IDH status, MGMT methylation, and 1p/19q codeletion via a Squeeze-and-Excitation channel attention mechanism. The model is trained and evaluated on the UTSW-Glioma dataset comprising 625 patients acquired on three scanner manufacturers (Siemens, GE, Philips), with a stratified split ensuring balanced distributions of IDH and scanners across the training, validation, and test sets.

The main contributions of this work are, first, the demonstration that M²-YOLO achieves a Dice ED of 0.8535 and a Dice TC of 0.6968 on the multi-site UTSW-Glioma dataset, achieving the highest Dice scores among the evaluated baseline models. Furthermore, M²-YOLO is the first YOLO-based framework to simultaneously integrate glioma instance segmentation and multi-task classification of molecular biomarkers, achieving an AUC ROC of 0.849 for IDH and 0.718 for 1p/19q with a maximum recall of 82% and a specificity of 96%.

The remainder of this paper is organized as follows. Section 2 presents related work. Section 3 describes the dataset and the preprocessing pipeline and details the M²-YOLO architecture. Section 4 presents the experimental results. Section 5 discusses the results and limitations, and Section 6 concludes the paper.

2. Related Work

Of the numerous applications of deep learning in neuro-oncology, glioma segmentation from MRI has received the most research attention to date. Early work is based on patch-based convolutional neural networks like DeepMedic, which optimized deep hierarchical feature extraction and were among the first works showcasing promising deep learning approaches successfully addressing brain tumor delineation. U-Net [11] introduced the encoder–decoder paradigm for learning to segment medical images, where multi-scale contextual information is combined with high quality spatial localization, which remained a de facto standard framework in the field, and nnU-Net [12] retaining the most influential among them as a self-configuring pipeline reaching state-of-the-art by inference through automated configuration and 3D convolutional processing at the expense of high inference cost preventing real-time clinic deployment. Recent variants addressing some of the limitations: Jadhav et al. In BraTS 2023, ref. [13] use dual squeeze-and-excitation attention blocks with N3 bias-field correction and obtain a Dice of 0.958, while Bhagyalaxmi and Dwarakanath [14] use a Wiener filter followed by a 3D ViT feature extractor with chaotic optimization to reach 0.987 in the tumor core. Aslam et al. [15] present a lightweight ResUNet50 variant reducing parameters by 60% while maintaining competitive segmentation performance.

Prior to the transformer era, a rich body of work extended the basic encoder–decoder paradigm through residual and dense connectivity patterns. ResNet-based segmentation backbones [16] demonstrated that deeper feature hierarchies improve delineation of infiltrative tumor boundaries, while DenseNet architectures [17] promoted feature reuse across layers to compensate for limited annotated training data. Attention U-Net [18] introduced soft spatial attention gates that selectively suppress irrelevant activations, yielding consistent gains on the BraTS challenge without increasing inference cost. TransUNet [19] further bridged convolutional feature extraction with global self-attention by embedding a Vision Transformer within a U-Net decoder, achieving competitive performance while retaining spatial detail. Collectively, these CNN variants demonstrated that architectural choices, skip connection design, attention gating, and residual learning have a greater impact on glioma segmentation accuracy than raw model depth, a finding that motivates our use of SE-channel attention within the YOLO detection head rather than a heavier transformer block.

More recently, hybrid CNN–Transformer architectures have emerged as powerful alternatives to purely convolutional networks. TransBTS, Swin UNETR [20], nnFormer and other models demonstrate that self-attention captures long-range spatial dependencies and complex tumor morphology not adequately modeled by local convolutions alone. Neural network-based architectures have demonstrated competitive performance on BraTS benchmarks and also robustness to heterogeneous data. However, their higher computational complexity and memory requirements, as well as longer inference latencies, are still major barriers for deploying in real clinical workflows. All of these segmentation-only methods do not focus on molecular biomarker classification, requiring independent downstream models for IDH, MGMT and 1p/19q prediction.

To overcome these limitations, recent research has explored one-stage detection and segmentation frameworks derived from the YOLO family. Unlike conventional encoder–decoder architectures, YOLO-based methods are designed to perform localization and segmentation within a unified and computationally efficient framework. NeuroYOLO [21] combines MobileNetV3 with YOLOv10 to reach mAP@0.5 of 97.2% in under 10 ms. YOLO-LS [22] employs a ShuffleNet backbone and DySample upsampling to obtain a Dice of 0.91 at only 8.1 GFLOPs. YOLO-Scan [23] uses multi-scale dilated attention to reach 62 FPS. As pointed out by Eun et al. [24], gliomas show lower detection performance than menin-

gliomas, with sensitivity for small tumors significantly increasing when the IoU threshold is relaxed. Importantly, while these YOLO variations show the potential for real-time detection, they do not build on their architecture to perform molecular biomarker classification and all are tested on single-site datasets that do not reflect variability due to scanner differences. Alongside deep learning, radiomics-based methods have been established as explainable and low-complexity benchmarks for the prediction of molecular biomarkers. Handcrafted features with classical classifiers: Handcrafted features of tumor regions of interest (ROIs), including first-order statistics, grey-level co-occurrence matrices, and wavelet-derived texture descriptors, have been used together with traditional classifiers such as support vector machines and random forests to predict IDH mutation status and MGMT promoter methylation from pre-operative MRI [25]. While radiomics pipelines offer transparency and require smaller training cohorts, their performance is highly sensitive to segmentation accuracy, feature selection strategy, and scanner-specific imaging protocols, limiting their reproducibility across institutions [26].

More recent hybrid approaches integrate handcrafted radiomic features with deep-learned representations to combine interpretability with discriminative power [27]. Despite these efforts, radiomic models consistently underperform with end-to-end deep learning methods on large multi-site cohorts, and none achieve the simultaneous segmentation and classification capability targeted by our framework.

The non-invasive prediction of molecular biomarkers (IDH, MGMT, 1p/19q) has become a key issue since their inclusion in the 2021 WHO classification. Li et al. [28] demonstrate the power of foundation models with UMBIF, a model pre-trained using self-supervised learning on 51,029 multi-institutional MRI scans, achieving AUCs of 0.916 for 1p/19q, 0.896 for IDH, and 0.859 for MGMT. Nalentzi et al. [29] compare the Transformer-based ViT and BEiT models for the classification of gliomas and meningiomas and demonstrate that ViT achieves an AUC of 0.962 and an accuracy of 90.4% on the test set. Catal Reis and Turk [30] present an extensive review of segmentation and classification methods, including re-transfer learning, hybrid methods and metaheuristics. Ketabi et al. [31] propose a novel approach that incorporates radiological reports into a multimodal contrastive learning framework. Brandl et al. [32] present MethyLYZR, a naive Bayes classifier for intraoperative diagnosis by nanopore sequencing, with an end-to-end turnaround time of less than 1 h. Despite these advances, most molecular prediction models rely on classification-only architectures that do not exploit the spatial information provided by accurate tumor segmentation, potentially limiting their performance and clinical utility.

Multi-task frameworks jointly addressing segmentation and molecular classification remain rare but increasingly promising. Cheng et al. [33] propose a fully automated multimodal MRI-based multi-task network for simultaneous glioma segmentation and IDH genotyping. MTAF-Net integrates multi-attention guidance and multi-scale feature fusion for joint 3D segmentation and IDH prediction [34]. TAUM-Net combines CNNs with a Transformer architecture for simultaneous glioma segmentation and MGMT methylation prediction [35]. Bjorkeli and Esmaeili [36] introduced a CNN encoder–decoder for joint segmentation and classification of IDH mutation and 1p/19q codeletion, making it structurally the closest approach to our proposed M²-YOLO framework. Using the BraTS 2020 dataset, their model achieved a mean Dice score of 0.73 for segmentation and an overall molecular classification accuracy of 0.98. However, the reported sensitivity for IDH mutation detection was only 0.56, likely reflecting the class imbalance and the limited number of patients in the dataset. Furthermore, their model does not predict MGMT status, relies on a conventional encoder–decoder architecture, and is not designed for real-time inference. Iqbal et al. [37] propose a two-stage pipeline merging 3D ResU-Net segmentation with 3D ResNet10 classification for MGMT prediction, attaining an AUC of only 0.66. Rhanoui

et al. [38] demonstrate the effectiveness of multi-task learning across four cancer types using UNet with pretrained backbones, but their architecture is not optimized for real-time inference. Beyond neuro-oncology, recent advances in other medical imaging domains further highlight the potential of modern deep learning strategies for improving feature representation and model robustness. For example, Shu et al. [39] proposed an active learning framework based on image similarity to reduce annotation costs while maintaining high segmentation performance for skin lesion analysis. Similarly, Xiong et al. [40] introduced HCMNet, a hybrid CNN-Mamba architecture that combines local and global feature extraction for accurate breast ultrasound segmentation. Although these methods address different clinical applications, they provide valuable insights into efficient feature learning, robust representation learning, and architecture design that may inspire future improvements of multi-task medical imaging frameworks such as M²-YOLO. Despite these advances, existing multi-task methods suffer from three limitations: (i) they rely on separate, non-unified architectures with error propagation between stages; (ii) they do not achieve real-time inference suitable for intraoperative use and (iii) they are predominantly evaluated on curated single-site datasets.

A major challenge for the clinical translation of all aforementioned approaches is their lack of generalizability to the heterogeneity of MRI acquisition protocols. Histogram normalization, N4 bias-field correction and z-score intensity standardization are data-level approaches (at the preprocessing level) that partially reduce intensity variations caused by scanner effects but not bigger differences in distributions due to different acquisition protocols [41]. Some of these include applying domain adversarial training [42] at the model level and instance normalization to brain MRI to reduce the feature-level gap between source and target scanners. Recently, methods to allow models to automatically predict on new unseen scanner distributions without target-domain labels include test-time adaptation and continual learning strategies [43]. Kurian et al. [6] conducted a longitudinal study across 21 UK centers, which showed sustained geographic disparities and average delays to diagnosis of 16 to 21 days, highlighting the need for scalable models. Canisius et al. Moreover, ref. [44] demonstrate that a minimal contrast-enhanced T1 and T2-FLAIR two-sequence protocol is adequate to yield segmentations comparable to those from full protocols, signaling the need for protocol-robust architectures.

Our M²-YOLO framework, trained and evaluated on the same dataset from three scanner manufacturers (Siemens, GE, Philips), is the first to introduce an end-to-end solution that simultaneously implements YOLO-based instance segmentation alongside SE-channel-attention molecular classification heads providing ms inference explicitly targeting the domain-shift problem.

3. Materials and Methods

3.1. Dataset UTSW-Glioma

The UTSW-Glioma dataset comprises 625 patients with histologically confirmed gliomas (WHO grades 2–4 according to the 2021 WHO classification) treated at the University of Texas Southwestern Medical Center between 2006 and 2023. Each patient record includes four pre-operative MRI sequences: pre-contrast T1-weighted (T1w), post-contrast T1-weighted (T1w + C, hereafter referred to as T1ce), T2-weighted (T2w), and T2-weighted fluid-attenuated inversion recovery (T2w-FLAIR), as illustrated in Figure 1. MRI examinations were acquired using scanners from multiple vendors with magnetic field strengths ranging from 0.3 T to 3 T, as shown in Figure 1.

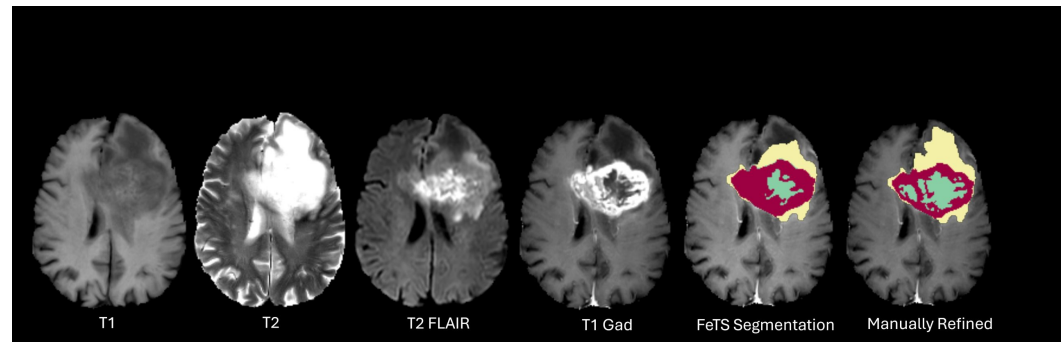


Figure 1. Example axial slice from the UTSW-Glioma dataset showing the four pre-operative MRI sequences and their corresponding tumor segmentations (T1, T2, T2-FLAIR, T1-Gad). Segmentation color code: yellow = peritumoral edema (ED), red = enhancing tumor (ET), green = necrotic core (NCR).

All scans were pre-processed using the FeTS platform [45], which registered each MRI scan to the SRI24 brain atlas at 1 mm³ isotropic resolution and performed skull stripping. The FeTS pipeline generated automated segmentations of three tumor subregions: enhancing tumor (ET), necrotic core (NCR), and peritumoral edema (ED) for all 625 cases. An initial subset of 362 cases underwent manual refinement using 3D Slicer by trained research personnel and subsequently validated by expert neuroradiologists [46]. The manually refined annotations provide a more accurate delineation of tumor subregions over the automated FeTS output, as shown in Figure 1. Since ET accounted for only 6.5% of the annotated instances, it was merged with NCR to form the Tumor Core (TC) region following BraTS convention, yielding two final segmentation classes used in this work: TC and ED. Molecular marker status, including IDH mutation, MGMT promoter methylation, and 1p/19q codeletion was assessed by biopsy and genomic analysis. A detailed summary of the dataset characteristics is provided in Table 1.

Table 1. Summary of the UTSW-Glioma Dataset.

Attribute	Details
Source	University of Texas Southwestern Medical Center
Period	2006–2023
Patients	625 (age > 18, no prior resection)
MRI Sequences	T1w, T1w + C (T1ce), T2w, T2w-FLAIR
Scanner Manufacturers	Siemens (<i>n</i> = 268), GE (<i>n</i> = 172), Philips (<i>n</i> = 141), Other (<i>n</i> = 44)
Field Strengths	0.3–3T
Original Tumor Regions	Enhancing Tumor (ET), Necrotic Core (NCR), Edema (ED)
Regions Used in This Work	Tumor Core TC (ET + NCR merged [†]), Edema (ED)
Segmentations	625 automated (FeTS) + 362 manually refined
IDH status	625 patients (100%)
MGMT methylation	281 patients (44.9%)
1p/19q codeletion	150 patients (24.0%)
Tumor Grades	WHO 2021: Grade 2, 3, 4
Preprocessing	FeTS (SRI24 atlas, 1 mm ³ , skull stripping)
Annotated Slices	~31,000 (tumor-containing only)
License	CC BY 4.0

[†] ET (6.5% of instances) merged with NCR following BraTS convention. Note: The remaining 44 patients were acquired using scanner manufacturers represented by a limited number of cases (e.g., Hitachi and Toshiba). These manufacturers were grouped as “Other” and were not analyzed separately because of the limited sample size.

3.2. Preprocessing

Four MRI sequences (T1w, T1w + C, T2w, T2w-FLAIR) were used to extract axial cross-sections to convert raw NIfTI volumes into 2D PNG slices. Each slice was normalized using percentile-based intensity scaling and resized to 640×640 pixels. Three sequences, T2w-FLAIR, T1w + C (T1ce), and T2w, were stacked into a three-channel PNG image to form the model input, following the standard BraTS protocol, which demonstrates that this combination is sufficient for TC and ED segmentation. Tumor sub-region annotations include three original regions: enhancing tumor (ET), necrotic core (NCR), and peritumoral edema (ED). Since ET represents only 6.5% of annotated instances, it was merged with NCR to form a single Tumor Core (TC) region following BraTS convention, yielding two segmentation classes: TC and ED. Annotations from the NIfTI segmentation masks were converted to YOLO polygon format using binary contour extraction. Slices without any tumor annotation were excluded from training, yielding approximately 31,000 annotated slices across the full dataset.

3.3. Dataset Split

The dataset was partitioned at the patient level into training (number of patients (n) = 500, 80%), validation (n = 62, 10%), and test (n = 63, 10%) sets using stratified sampling to preserve the joint distribution of IDH mutation status and scanner manufacturer across all splits. Stratification was applied to these two variables only, as MGMT methylation and 1p/19q codeletion had insufficient per-stratum coverage for stable multi-variable stratification. This stratification reduced the validation-to-test AUC gap for IDH from 13.7 to 5.2 points compared with random splitting. All three molecular biomarkers were used for classification training and evaluation, with per-task masking applied to exclude patients with unknown status for each biomarker.

3.4. M^2 -YOLO Architecture

The YOLO26m-seg backbone [10] is a real-time instance segmentation architecture from the Ultralytics YOLO26 family (23.6 M parameters, 121.5 GFLOPs, mask mAP_{50–95} of 44.1% on COCO). As shown in Figure 2, it is comprised of three main components: (1) a CSP-based DarkNet backbone that processes the input image through convolutional and C2f blocks followed by a Spatial Pyramid Pooling Fast (SPPF) module, producing hierarchical feature maps at three scales (P3: 80×80 , P4: 40×40 , P5: 20×20); (2) a Path Aggregation Network (PAN-FPN) neck that fuses multi-scale features through top-down and bottom-up pathways and (3) a decoupled output stage with a detection head performing bounding box regression, classification and Distribution Focal Loss (DFL) on P3, P4, and P5, and a segmentation head operating on P3 and P4 to generate instance masks. The official Ultralytics YOLO26m-seg implementation was used without modifying the default detection loss; therefore, Distribution Focal Loss (DFL) was retained for bounding box regression.

M^2 -YOLO is based on the YOLO26m-seg backbone, augmented with specialized heads for the concurrent tasks of instance segmentation and molecular biomarker classification. The shared backbone extracts hierarchical features from the three-channel PNG input ($640 \times 640 \times 3$), and produces multi-scale feature maps at three resolutions (P3: $80 \times 80 \times 256$, P4: $40 \times 40 \times 512$, P5: $20 \times 20 \times 512$). The segmentation head uses the Segment26 decoder operating on P3 and P4 features to produce two instance masks: TC (necrotic core) and ED (peritumoral edema), trained with a combined box, classification, DFL, and mask loss. The three molecular classification heads operate on P5 features ($20 \times 20 \times 512$) obtained by a forward hook on the backbone. Each head employs Squeeze-and-Excitation (SE) channel attention, in which separate global average pooling (GAP)

and global max pooling (GMP) streams are each passed through a fully connected shared bottleneck layer ($512 \rightarrow 32 \rightarrow 512$) with sigmoid activation, and then added to generate the channel-wise attention weights to recalibrate the feature responses. The attended features are aggregated by global average pooling and passed through a fully connected classifier ($512 \rightarrow 256 \rightarrow 128 \rightarrow 32 \rightarrow 1$) with batch normalization and task-specific dropout (IDH: $p = 0.25$, MGMT: $p = 0.30$, 1p/19q: $p = 0.35$). A single forward pass simultaneously produces the segmentation masks and three molecular predictions, as depicted in Figure 3.

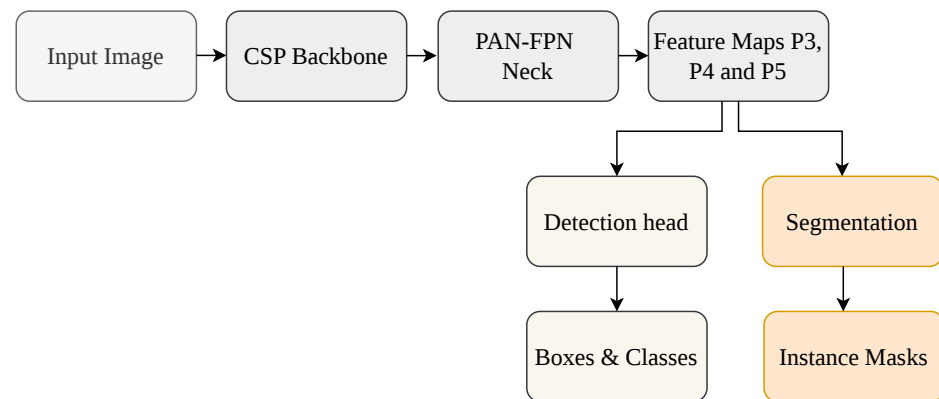


Figure 2. General architecture of the YOLO26m-seg backbone, adapted from the official Ultralytics documentation [10], consisting of a CSP backbone for multi-scale feature extraction (P3, P4, P5), a PAN-FPN neck for feature fusion, a detection head for bounding box regression and classification, and a segmentation head for instance mask generation.

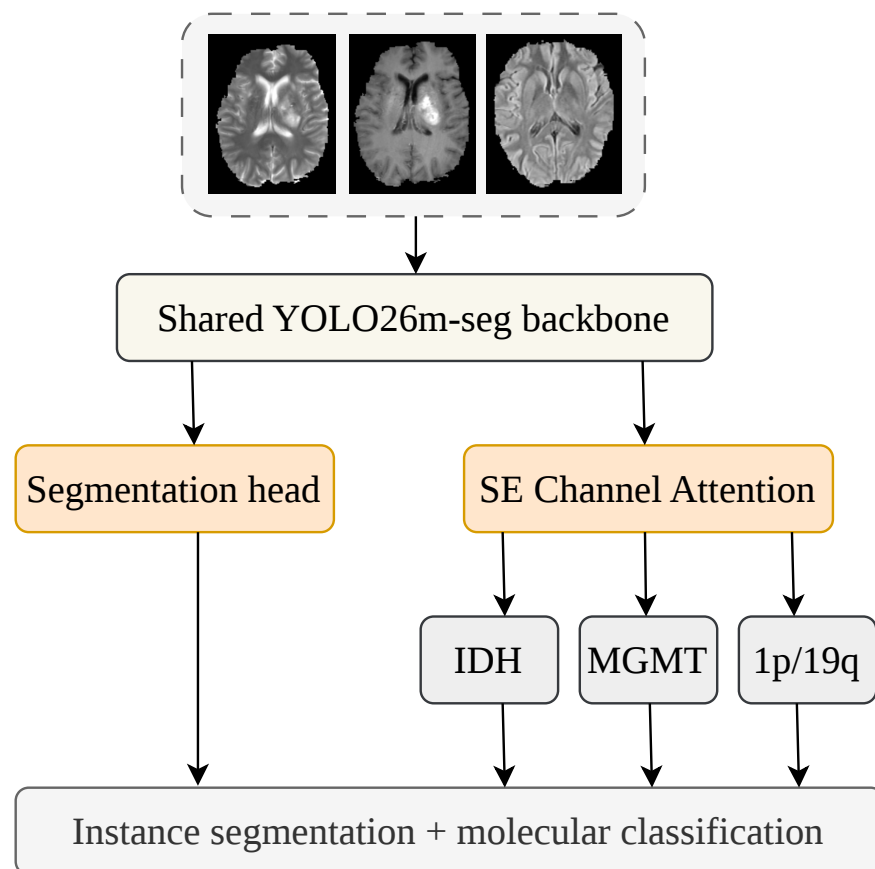


Figure 3. Overview of the proposed M²-YOLO architecture. The shared YOLO26m-seg backbone extracts multi-scale features, which are fed simultaneously into the instance segmentation head and three molecular classification heads (IDH, MGMT, 1p/19q) equipped with SE Channel Attention.

3.5. Training

M²-YOLO is trained in two phases on an NVIDIA RTX 3090 GPU (24 GB). Although M²-YOLO performs tumor segmentation and molecular biomarker prediction simultaneously during a single forward inference pass, the model is optimized using a two-stage training strategy rather than joint end-to-end optimization. In the first phase, the segmentation backbone is trained to learn robust tumor representations. In the second phase, the learned backbone is frozen while the molecular classification heads are optimized. This sequential training strategy improves optimization stability while preserving a unified inference pipeline. **Phase 1:** The YOLO26m-seg backbone was trained for instance segmentation for 60 epochs with batch size 16, an AdamW optimizer ($\text{lr} = 5 \times 10^{-4}$, weight decay = 10^{-3}) and a cosine scheduler ($\text{lrf} = 0.01$). The data augmentation techniques applied were mosaic ($p = 0.5$), mixup ($\alpha = 0.1$), copy-paste ($p = 0.1$), random horizontal and vertical flips ($p = 0.5$), scaling ($\pm 50\%$) and RandAugment.

The corresponding learning curves are presented in Figure 4, confirming stable convergence of the segmentation backbone over 60 epochs.

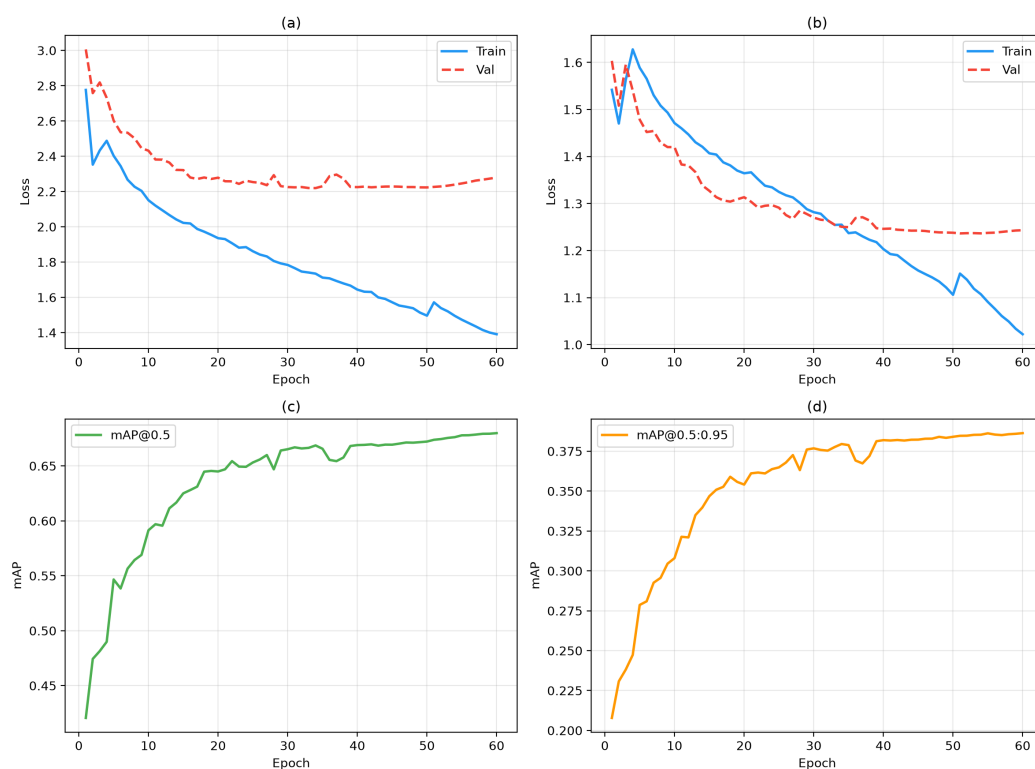


Figure 4. Learning curves for the segmentation stage (Phase 1). (a) Training and validation segmentation loss. (b) Training and validation box loss. (c) Validation mask mAP@0.5. (d) Validation mask mAP@0.5:0.95. All metrics are reported over 60 training epochs.

Phase 2: The SE channel attention molecular classification heads were trained for 20 epochs with the backbone frozen, an AdamW optimizer ($\text{lr} = 5 \times 10^{-4}$, weight decay = 5×10^{-4}), and CosineAnnealing ($T_{\text{max}} = 20$). The batch size is 32. The best checkpoint was selected at epoch 8 based on the validation AUC (AUC = 0.724). The molecular loss function is the weighted sum of three task-specific losses:

$$\mathcal{L}_{\text{mol}} = 0.40 \cdot \mathcal{L}_{\text{IDH}} + 0.30 \cdot \mathcal{L}_{\text{MGMT}} + 0.30 \cdot \mathcal{L}_{1p/19q} \quad (1)$$

where $\mathcal{L}_{\text{IDH}}$ is a weighted BCE loss (positive weight $w^+ = 2.5$), and $\mathcal{L}_{\text{MGMT}}$ and $\mathcal{L}_{1p/19q}$ are Focal BCE losses ($\alpha = 0.25$, $\gamma = 2.0$ and $\alpha = 0.35$, $\gamma = 2.5$, respectively) to ad-

dress class imbalance. All training slices were subject to horizontal and vertical flipping ($p = 0.5$). Additional augmentation (rotation $\pm 20^\circ$, brightness jitter $\pm 20\%$, Gaussian blur) was applied to 1p/19q-labeled slices to address class rarity (~ 150 known patients). Figure 5 presents the validation AUC-ROC curves during Phase 2 training, confirming that the molecular classification heads converge within the first 6–9 epochs, with no significant improvement observed beyond this point, justifying the use of early stopping for model selection.

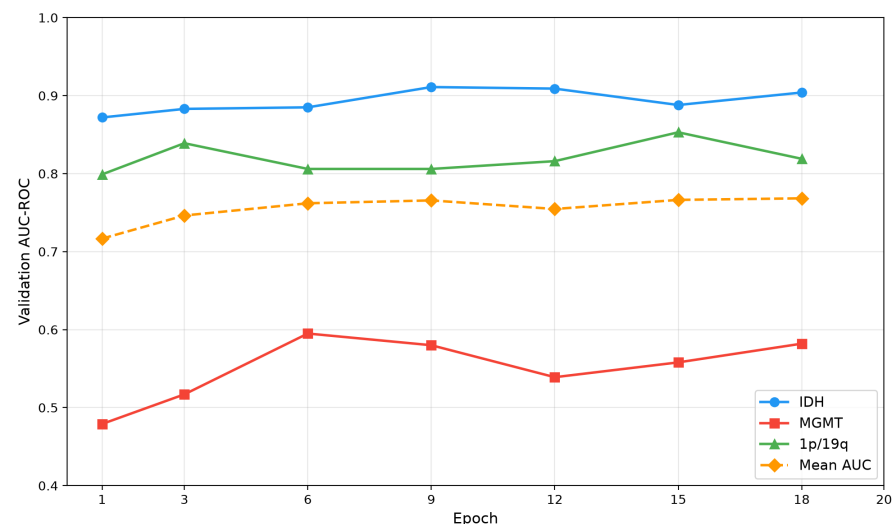


Figure 5. Validation AUC-ROC curves during Phase 2 training (molecular classification heads, 20 epochs), obtained from a reproduction experiment using the same architecture, dataset split, and training protocol. IDH prediction converges rapidly and remains stable, MGMT remains near chance level throughout training, and 1p/19q shows progressive improvement.

4. Experimental Results

This section presents the experimental results of M²-YOLO across its two complementary tasks. Segmentation performance on the UTSW-Glioma dataset is first reported with subgroup analysis by manufacturer, IDH status, and tumor grade. Molecular biomarker classification results are then detailed for IDH mutation, MGMT methylation, and 1p/19q codeletion prediction, followed by an internal ablation study comparing segmentation backbones and classification heads, and a contextual comparison with representative state-of-the-art methods.

4.1. Segmentation Performance

The segmentation ability of M²-YOLO on the UTSW-Glioma dataset is presented in Table 2. Figure 6 presents qualitative segmentation results of three selected test cases showing the ability of the model to segment Tumor Core (TC) and Edema (ED) of different tumor shapes.

The model obtained an overall Dice ED of 0.8535 and an overall Dice TC of 0.6968. The much better edema segmentation compared to necrotic core segmentation can be explained by the better delineation of edema on FLAIR and T2 sequences. A clear difference is observed depending on the MRI manufacturer: Philips images exhibit the best performance for TC segmentation (Dice = 0.7614), whereas GE scanners show the lowest performance (Dice TC = 0.6083), a difference of 16.2 points. TC segmentation is significantly influenced by IDH status: IDH-mutated tumors achieve a higher Dice score (0.7471) than IDH wild-type tumors (0.6671), confirming the better-defined contours of IDH-mutated gliomas. Finally, the best TC segmentation (Dice = 0.7556) is achieved for grade 2 tumors, whereas

grade 3 tumors have the worst performance (Dice = 0.5879), a 16.8-point drop compared to grade 2.

Table 2. Segmentation performance of M²-YOLO on the UTSW-Glioma dataset, TC and ED Dice scores by manufacturer, IDH status, and grade.

Group	TC (Dice)	ED (Dice)	N
Global	0.6968	0.8535	1760
Siemens	0.7112	0.8485	783
GE	0.6083	0.8601	470
Philips	0.7614	0.8509	415
Other	-	-	92
IDH Wild-type	0.6671	0.8717	1107
IDH Mutant	0.7471	0.8060	653
Grade 2	0.7556	0.7989	549
Grade 3	0.5879	0.8739	159
Grade 4	0.6825	0.8713	1052

TC: Tumor Core. ED: Edema. N: number of annotated slices. Other includes slices acquired from scanner manufacturers represented by a limited number of cases (e.g., Hitachi and Toshiba), which were not analyzed separately because of their small sample size.

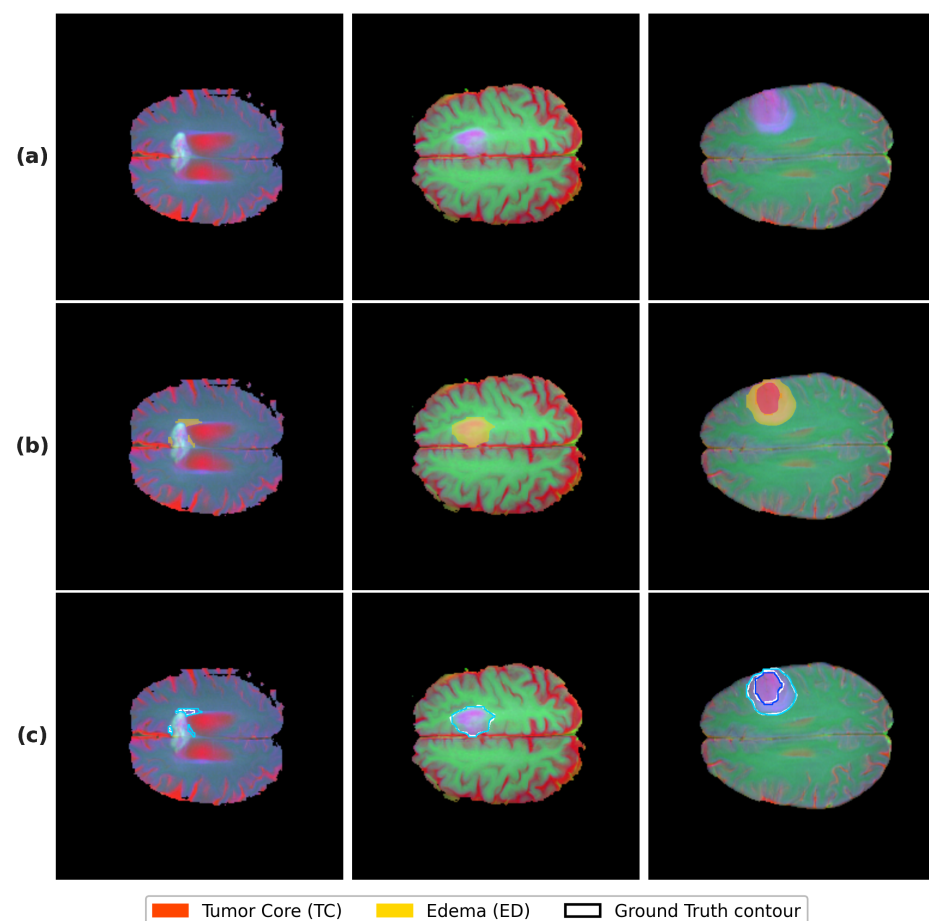


Figure 6. Qualitative segmentation results of M²-YOLO on three representative test cases (BT0023_z098, BT0120_z104, BT1051_z098). Row (a): input multi-sequence MRI (T1, T1ce, T2, FLAIR fused). Row (b): M²-YOLO predicted masks Tumor Core (TC, red) and Edema (ED, yellow). Row (c): overlay of ground truth contours (white) on predicted masks, illustrating boundary agreement for well-delineated tumors (BT1051_z098) and challenging cases (BT0023_z098).

To further characterize boundary quality, we computed three surface-based metrics—Hausdorff Distance 95th percentile (HD95), Average Surface Distance (ASD, unidirectional prediction-to-ground-truth), and Average Symmetric Surface Distance (ASSD)—in physical units (mm), using the 0.375 mm in-plane pixel spacing corresponding to the 640×640 resampled image resolution. Each metric was computed per tumor-bearing test slice and aggregated in two stages: first averaged within each patient, then averaged across patients with 95% confidence intervals. As shown in Table 3, M²-YOLO achieves sub-5 mm HD95 for both classes, indicating strong boundary agreement, which is critical for clinical applications such as surgical planning.

Table 3. Boundary-based segmentation metrics of M²-YOLO, patient-level aggregation with 95% confidence intervals (mm, 0.375 mm in-plane spacing).

Class	HD95 (mm)	ASD (mm)	ASSD (mm)	N Patients
Tumor Core (TC)	4.84 ± 1.01	0.91 ± 0.33	1.58 ± 0.39	54
Edema (ED)	4.86 ± 0.77	0.68 ± 0.05	1.18 ± 0.14	63

Note: Patient counts differ slightly between Dice (N = 59 for TC) and boundary metrics (N = 54 for TC), as boundary distances require both predicted and ground-truth masks to be non-empty on the same slice, whereas patient-level Dice aggregates volumetrically. The five additional patients excluded from HD95/ASD/ASSD had ground-truth TC present but no slice with a spatially overlapping non-empty prediction, indicating detection failure on their (typically few) TC-containing slices rather than a boundary quality issue. Slices with an empty predicted or ground-truth mask were excluded from the slice-level pool (1883/3329 for TC; 78/3329 for ED).

To further validate the segmentation performance of the proposed framework, we compared M²-YOLO with four widely used segmentation models, namely U-Net, nnU-Net, EfficientSeg, and YOLO11. All models were trained and evaluated on the UTSW-Glioma dataset using the same preprocessing pipeline and identical training, validation, and test splits to ensure a fair comparison. As summarized in Table 4, M²-YOLO achieved the highest Dice scores for both Tumor Core (TC) and Edema (ED), among the evaluated baseline models. Specifically, M²-YOLO improved the TC Dice score by 3.2 percentage points over EfficientSeg (0.697 vs. 0.665) and by 6.6 percentage points over U-Net (0.697 vs. 0.631). Similarly, the proposed model achieved the highest ED Dice score (0.853), exceeding nnU-Net (0.805), U-Net (0.802), EfficientSeg (0.799), and YOLO11 (0.783). These results demonstrate the effectiveness of the proposed architecture for accurate glioma segmentation while maintaining the real-time inference capability of the YOLO framework.

Table 4. Comparison of segmentation performance on the UTSW-Glioma dataset with representative baseline models.

Model	TC (Dice)	ED (Dice)
YOLO11	0.651	0.783
U-Net	0.631	0.802
nnU-Net	0.647	0.805
EfficientSeg	0.665	0.799
M²-YOLO	0.697	0.853

Bold indicates the proposed method.

To complement the slice-level results, we additionally evaluated segmentation at the patient level (Table 5), where a single volumetric Dice score is computed per patient and averaged across the test cohort with 95% bootstrap confidence intervals. The patient-level scores (TC = 0.706, ED = 0.872) are consistent with the slice-level values, confirming that the reported segmentation performance is not an artifact of slice-level aggregation.

Table 5. Patient-level volumetric Dice scores of M²-YOLO on the UTSW-Glioma test set, with 95% bootstrap confidence intervals.

Class	Slice-Level Dice	Patient-Level Dice (95% CI)	Patients
Tumor Core (TC)	0.697	0.706 (0.629–0.779)	59
Edema (ED)	0.853	0.872 (0.851–0.890)	63

Patient-level Dice was computed by reconstructing the 3D volume of each test patient from its axial slices and evaluating a single volumetric Dice per patient. TC was evaluated on 59 patients, as 4 test patients contained no annotated tumor-core region.

The quantitative gap is visually confirmed in Figure 6: BT1051_z098, a well-delineated tumor, shows tight contour agreement (row c) and BT0023_z098 illustrates a challenging TC case with diffuse boundaries, in agreement with the lower Dice TC for IDH wild-type tumors (0.667).

4.2. Molecular Classification Performance

Table 6 shows the classification performance of molecular biomarkers on the stratified test dataset. In predicting the IDH mutation status, the model achieved an AUC-ROC of 0.849, an accuracy of 0.81, sensitivity of 0.70, and specificity of 0.86. This is among the best reported performances of deep learning models (0.85–0.89), and shows the potential of the YOLO multi-task framework for non-invasive molecular classification. The AUC-ROC for 1p/19q codeletion prediction is 0.718 with optimal sensitivity and specificity of 0.82 and 0.95, respectively. While the AUC is moderate, the specificity and sensitivity have clinical importance: a specificity of 0.96 limits false positives and thus prevents inappropriate therapies for patients without the codeletion, which is crucial for the diagnosis of oligodendroglioma. We nevertheless retain MGMT prediction in the proposed multi-task framework for two reasons. First, according to the 2021 WHO classification of central nervous system tumors, MGMT promoter methylation remains one of the key molecular biomarkers for glioma characterization and treatment planning. Including this task, therefore, provides a more comprehensive molecular assessment. Second, the MGMT task is intentionally assigned a lower weight in the multi-task loss ($\lambda_{\text{MGMT}} = 0.30$ versus $\lambda_{\text{IDH}} = 0.40$, Equation (1)) and is predicted through a dedicated classification head, thereby reducing its influence on the optimization of the remaining molecular tasks.

Table 6. Molecular classification performance of M²-YOLO on the stratified test set, SE Channel Attention heads.

Biomarker	Patients	Slices	Slice AUC	Patient AUC (95% CI)	Accuracy	Sensitivity	Specificity	F1 (Class 0)	F1 (Class 1)	F1-Macro
IDH mutation	63	3329	0.849	0.867 (0.760–0.953)	0.811	0.704	0.863	0.872	0.686	0.779
MGMT methylation	31	1622	0.520	0.314 (0.105–0.544)	0.564	0.420	0.690	0.662	0.448	0.555
1p/19q codeletion	39	1985	0.718	0.745 (0.527–0.921)	0.754	0.820	0.959	0.959	0.821	0.890

Class 0: IDH Wild-type/MGMT Unmethylated/1p/19q Non-deleted. Class 1: IDH Mutant/MGMT Methylated/1p/19q Co-deleted. Patient-level AUCs were obtained by averaging slice-level prediction probabilities for each patient, followed by ROC analysis. Confidence intervals were estimated using bootstrap resampling (95% CI).

Prediction of MGMT promoter methylation has an AUC-ROC of 0.520, which is similar to random prediction. This result confirms an important observation in the literature that the methylation status of MGMT cannot be accurately predicted using only conventional anatomical MRI sequences and that advanced techniques such as diffusion weighted imaging (ADC) or perfusion imaging are probably required [47,48]. Although both tasks were evaluated on the same 63 test patients, the reported number of slices differs between Tables 2 and 6 because the two tasks use different evaluation protocols. Segmentation

performance was computed only on slices containing tumor annotations, whereas molecular classification was evaluated on all eligible slices from patients with the corresponding molecular labels. Therefore, the reported N values represent the number of evaluated slices for each task rather than the number of patients. Our proposed model differs from existing methods in several significant ways. First, unlike two-stage pipelines such as those of Iqbal et al. [37], which combine 3D ResU-Net segmentation followed by 3D ResNet10 classification, M²-YOLO performs both tasks in a single inference step (3.8 ms per slice). Inference time was measured on a single NVIDIA GeForce RTX 3090 GPU (24 GB VRAM) equipped with an AMD Ryzen Threadripper PRO 3975WX processor (32 cores, 64 threads; up to 4.37 GHz) and 503 GB system memory, running Ubuntu Linux with CUDA 12.8. Inference was performed in FP32 precision with a batch size of 1 using input images resized to 640 × 640 pixels. The reported inference time corresponds to the forward pass of the complete M²-YOLO framework, including the segmentation and molecular classification heads, and excludes MRI preprocessing and post-processing. Regarding IDH prediction, the ROC AUC of 0.849 falls within the high range of published models (0.85–0.89). Although slightly inferior to the performance of the UMBIF foundation model (AUC = 0.896) reported by Li et al. [28], our model has the advantage of performing segmentation and inference simultaneously, which is more suitable for the constraints of the operating room. This ability is important for intraoperative guidance of specific biopsies and surgical decisions. For 1p/19q prediction, the AUC of 0.718 is moderate, but the recall (82%) and specificity (96%) values are especially remarkable. Such high specificity is crucial in clinical practice to avoid inappropriate treatments in patients without the deletion. An AUC of 0.520 for MGMT prediction confirms that anatomical MRI alone cannot predict MGMT status, consistent with the literature [37,49], and motivates further research on the integration of advanced sequences (ADC, perfusion, PET).

Table 7 provides a structured comparison of M²-YOLO against representative state-of-the-art methods. It should be noted that a direct numerical comparison remains methodologically limited, as these works are evaluated on different datasets and address different task scopes. The lower section of the table, therefore, reports an internal ablation study in which all competing backbones were trained and evaluated under identical conditions on the UTSW-Glioma dataset. These results indicate that YOLO26m-seg combined with SE-channel attention provides a favorable trade-off between segmentation accuracy, molecular classification performance, and inference speed among the evaluated configurations.

Table 7. Contextual Comparison of M²-YOLO with Published Glioma Analysis Methods Across Different Datasets and Task Scopes.

Method	Tasks	Dice (Seg.)	IDH AUC	MGMT AUC	1p/19q AUC	Dataset
BC-TSEA [13]	Seg.	0.958	—	—	—	BraTS
NeuroYOLO [21]	Det. + Cls.	—	—	—	—	Figshare
3D ResU-Net [37]	Seg. + MGMT	0.84	—	0.66	—	BraTS
CNN enc-dec [36]	Seg. + IDH	0.73	—	—	—	BraTS
VGG + UNet [38]	bin. Cls + Seg.	0.95–0.99	—	—	—	Multi
UMBIF [28]	mol. Cls	—	0.896	0.859	0.916	Multi-centre
YOLO26m-seg + ViT	Seg. + IDH/MGMT/1p19q	0.853/0.697	0.823	0.510	0.690	UTSW (3 scan.)
Ours (M²-YOLO)	Seg. + IDH/MGMT/1p19q	0.853/0.697	0.849	0.520 [†]	0.718	UTSW (3 scan.)

—: Not Reported. Bold indicates the proposed method. [†] Not statistically significant (AUC ≈ 0.5, chance level).

5. Discussion

We propose multi-task learning, which is promising for joint glioma segmentation and molecular classification. Recent studies have demonstrated the feasibility of jointly learning glioma segmentation and molecular characterization within unified multi-task

frameworks [50,51]. Unlike the jointly trained multi-task frameworks proposed by van der Voort et al. [50] and Wen et al. [51], M²-YOLO adopts a sequential two-stage optimization strategy, in which the segmentation backbone is first optimized and subsequently frozen while the molecular classification heads are trained. Although this strategy does not correspond to joint end-to-end optimization, it improves training stability. In our work, we fill these gaps by proposing M²-YOLO, a unified YOLO-based multi-task framework for simultaneous glioma instance segmentation and prediction of IDH, MGMT, and 1p/19q molecular biomarkers in a single forward inference pass. We chose the UTSW-Glioma dataset because it consists of 625 patients acquired from various scanner manufacturers (Siemens, GE, Philips) providing a realistic multi-site scenario with rich molecular annotations (IDH: 625 patients, MGMT: 281, 1p/19q: 150). Stratified splitting by IDH status and scanner manufacturer resulted in balanced distributions across all splits and reduced the validation-to-test AUC gap for IDH from 13.7 to 5.2 points compared to random splitting, confirming the importance of accounting for acquisition heterogeneity when evaluating on multi-site data.

In terms of segmentation, M²-YOLO achieved the highest Dice scores among the evaluated baseline models with a Dice ED of 0.8535 and Dice TC of 0.6968. These enhancements can be largely ascribed to the hierarchical multi-scale feature representation capability of the YOLO26m-seg backbone that fits well to local boundary details and global tumor context. Nonetheless, we note the difficulty created by variability in acquisition protocol, as evidenced by the Dice TC of 16.2 points separating GE and Philips scanners, indicating a necessity for multi-site training to yield robust clinical deployment.

The observed performance gap between scanner manufacturers is likely multifactorial. Differences in MRI acquisition protocols, including pulse-sequence parameters, image reconstruction algorithms, and image contrast, may affect the delineation of tumor boundaries and consequently segmentation performance. In addition, variations in scanner hardware and magnetic field strength may influence image quality and signal-to-noise ratio, thereby affecting the features learned by the network. Finally, differences in the distribution of tumor grades and molecular subtypes across manufacturers may also contribute to the observed variability. Although the present study was not designed to quantify the individual contribution of these factors, these findings highlight the challenges of deploying deep learning models across heterogeneous clinical imaging environments. This scanner-dependent variability suggests that the current model has not fully learned scanner-invariant representations. Future work will explore domain-adversarial training [42] or histogram-matching harmonization across manufacturers to close this gap, rather than relying solely on data-level normalization. Although M²-YOLO operates on individual axial (2D) MRI slices rather than full volumetric (3D) data, this design choice substantially reduces computational complexity and enables real-time inference (3.8 ms per slice), making the framework more suitable for time-sensitive clinical applications. In contrast, most existing multi-task frameworks rely on computationally intensive 3D encoder–decoder architectures that better exploit inter-slice anatomical context but incur considerably higher memory consumption and inference latency. Consequently, M²-YOLO represents a trade-off between computational efficiency and volumetric contextual information. Future work will investigate 2.5D and fully 3D extensions to better capture inter-slice anatomical continuity while preserving the computational advantages of the proposed framework. Regarding molecular classification, the SE channel attention module was selected after comparing several classification head architectures under identical training conditions on the UTSW-Glioma dataset. A ViT-based head achieved a higher validation AUC for IDH (0.801 vs. 0.724 for SE attention) but showed poor generalization to the test set, at more than 20× the computational cost. The SE attention module offers an efficient

channel recombination (512→32→512) that selectively amplifies discriminative feature representations while remaining well-suited for learning from limited and class-imbalanced labeled cohorts.

For IDH prediction, M²-YOLO achieved an AUC of 0.849, placing it within the upper range of published MRI-based models (0.85–0.89). Although somewhat lower than the UMBIF foundation model (AUC = 0.896, [28]) pre-trained on 51,029 scans without simultaneous segmentation output, M²-YOLO provides the clinical benefit of joint segmentation and classification in a single real-time pipeline. The AUC for 1p/19q codeletion was 0.718 with a recall of 82% and specificity of 96%, an important characteristic for reducing false-positive diagnoses in oligodendroglioma patients. MGMT promoter methylation prediction remained close to chance (AUC = 0.520), consistent with previous studies reporting that standard anatomical MRI sequences do not provide adequate signal to differentiate methylated from unmethylated MGMT [37,48]. Future work will consider the integration of advanced imaging modalities, including ADC maps, perfusion imaging, and amino-acid PET [47], which may better capture the biological processes underlying this biomarker. This hypothesis has not yet been experimentally validated in the present study, as the UTSW-Glioma dataset does not include diffusion- or perfusion-weighted sequences; its evaluation is left for future work pending access to a multiparametric cohort. A formal ablation study removing the MGMT classification head was not performed in the present work. Therefore, the precise influence of the MGMT task on the prediction of IDH mutation and 1p/19q codeletion remains to be quantified. This constitutes a limitation of the current study and will be investigated in future work. Finally, further independent international cohorts are needed to validate M²-YOLO and ensure its robustness and generalizability across different clinical settings and scanner manufacturers.

6. Conclusions

In this paper, we propose M²-YOLO, the first unified YOLO-based framework for concurrent glioma instance segmentation and multi-task molecular classification of IDH, MGMT and 1p/19q from preoperative multi-sequence MRI. To avoid the latency and error propagation in two-stage pipelines, M²-YOLO uses a shared YOLO26m-seg backbone with SE channel attention classification heads. The entire pipeline produces segmentation masks and molecular predictions within 3.8 ms. Evaluated on the multi-site UTSW-Glioma dataset (625 patients, 3 scanner manufacturers), the model achieves a Dice ED of 0.8535 and a Dice TC of 0.6968, achieving the highest Dice scores among the evaluated baseline models, while reaching an AUC-ROC of 0.849 for IDH and 0.718 for 1p/19q with a specificity of 96%. As shown in the literature, the prediction of MGMT status from only anatomical MRI remains highly ineffective (AUC=0.520) and will need functional sequences such as ADC or perfusion imaging to be further developed in future work. Integrating real-time segmentation of tumor subregions with non-invasive molecular profiling into one framework provides concrete opportunities for intraoperative guidance and precision treatment planning according to the 2021 WHO classification.

Author Contributions: Conceptualization, S.H.A. and M.A.A.; methodology, S.H.A. and M.A.A.; validation, S.H.A. and M.A.A.; formal analysis, S.H.A. and M.A.A.; writing—original draft preparation, S.H.A.; writing—review and editing, M.A.A.; funding acquisition, M.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), under funding reference number RGPIN-2024-05287, and by the AI in Health Research Chair at the Université de Moncton.

Institutional Review Board Statement: This study is a secondary analysis of the publicly available, de-identified UTSW-Glioma dataset [46]. No additional institutional review board approval was required.

Informed Consent Statement: Not applicable.

Data Availability Statement: This work uses a publicly available dataset; see reference [46] for data availability. More details about this dataset are available in Section 3.1.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Ostrom, Q.T.; Cioffi, G.; Waite, K.; Kruchko, C.; Barnholtz-Sloan, J.S. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2014–2018. *Neuro-Oncology* **2021**, *23*, iii1–iii105. [CrossRef] [PubMed]
- Louis, D.N.; Perry, A.; Wesseling, P.; Brat, D.J.; Cree, I.A.; Figarella-Branger, D.; Hawkins, C.; Ng, H.; Pfister, S.M.; Reifenberger, G.; et al. The 2021 WHO classification of tumors of the central nervous system: A summary. *Neuro-Oncology* **2021**, *23*, 1231–1251. [CrossRef] [PubMed]
- Patil, N.; Dube, S.; Mutua, F.; Kakumanu, S.; Shankar, J.J.; Sinha, N.; Pareek, V.; Sinha, N. Survival Impact of Isocitrate Dehydrogenase (IDH)-Wildtype Histological Versus Molecular Glioblastoma: A Propensity Score-Matched Analysis. *Cureus* **2025**, *17*, e88667. [CrossRef]
- Leiss, S.M.; Wiestler, B.; Hou, H.X.; Schmottermeyer, D.; Sackerer, V.; Diehl, C.; Peeken, J.; Borm, K.; Negwer, C.; Wagner, A.; et al. Comorbidity burden in elderly high-grade glioma patients: Impact on radiotherapy outcomes. *BMC Cancer* **2025**, *25*, 1496. [CrossRef] [PubMed]
- Hou, H.X.; Leiss, S.M.; Schmottermeyer, D.; Diehl, C.; Wiestler, B.; Peeken, J.; Borm, K.; Negwer, C.; Wagner, A.; Yakushev, I.; et al. Radiotherapy and survival in elderly grade 4 glioma patients: The prognostic value of onco-functional outcome. *Clin. Transl. Radiat. Oncol.* **2025**, *56*, 101085. [CrossRef] [PubMed]
- Kurian, K.M.; Wright, A.P.; Sari, F.; Briggs, M.; Goetz, C.; Quinlivan, R.; Bushby, K.; Gilbertson, R.; Huskens, N. Transforming molecular neuropathology for adult brain tumor patients in the UK: Insights on implementation, adoption, and patient access (2021–2024). *Neuro-Oncol. Pract.* **2026**, *13*, 465–477.
- Christodoulou, R.C.; Pitsillos, R.; Papageorgiou, P.S.; Petrou, V.; Vamvouras, G.; Rivera, L.; Papageorgiou, S.G.; Solomou, E.E.; Georgiou, M.F. Artificial intelligence in glioma diagnosis: A narrative review of radiomics and deep learning for tumor classification and molecular profiling across positron emission tomography and magnetic resonance imaging. *Eng* **2025**, *6*, 262. [CrossRef]
- Contreras, K.; Velez-Varela, P.E.; Casanova-Carvajal, O.; Alvarez, A.L.; Urbano-Bojorge, A.L. A review of artificial intelligence-based systems for non-invasive glioblastoma diagnosis. *Life* **2025**, *15*, 643. [CrossRef]
- Wu, Q.; Deng, X.; Shao, H.; Jiang, Y. Recent advances in brain tumor segmentation: A review of self-supervised learning, federated learning, and graph neural networks. *Eng. Appl. Artif. Intell.* **2026**, *173*, 114473. [CrossRef]
- Ultralytics. Instance Segmentation—Ultralytics YOLO Docs. 2025. Available online: <https://docs.ultralytics.com/tasks/segment#models> (accessed on 15 June 2026).
- Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Daejeon, Republic of Korea, 23–27 September 2015*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
- Isensee, F.; Jaeger, P.F.; Kohl, S.A.; Petersen, J.; Maier-Hein, K.H. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **2021**, *18*, 203–211. [PubMed]
- Jadhav, A.; Rasool, A.; Gyanchandani, M. Bias Corrected Twin Squeeze-and-Excitation Attention Enhanced UNet for brain tumor segmentation. *Biomed. Signal Process. Control* **2026**, *115*, 109390. [CrossRef]
- Bhagyalaxmi, K.; Dwarakanath, B. CDCG-UNet: Chaotic optimization assisted brain tumor segmentation based on dilated channel gate attention U-Net model. *Neuroinformatics* **2025**, *23*, 12. [CrossRef] [PubMed]
- Aslam, M.; Maqbool, M.S.; Aoun, M.; Aslam, N.; Razzaq, A.M.; Ali, S. High-performance and efficient brain tumor segmentation for enhanced clinical analysis. *Spectr. Eng. Sci.* **2026**, *4*, 195–210. [CrossRef]
- 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, Las Vegas, NV, USA, 27–30 June 2016*; pp. 770–778.
- Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017*; pp. 4700–4708.
- Oktay, O.; Schlemper, J.; Folgoc, L.L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N.Y.; Kainz, B.; et al. Attention u-net: Learning where to look for the pancreas. *arXiv* **2018**, arXiv:1804.03999.

19. Chen, J.; Lu, Y.; Yu, Q.; Luo, X.; Adeli, E.; Wang, Y.; Lu, L.; Yuille, A.L.; Zhou, Y. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv* **2021**, arXiv:2102.04306.
20. Hatamizadeh, A.; Nath, V.; Tang, Y.; Yang, D.; Roth, H.R.; Xu, D. Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In *Proceedings of the International MICCAI Brainlesion Workshop, Virtual Event, 27 September 2021*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 272–284.
21. Hasan, M.M.; Paul, B.K.; Uddin, M.S.; Talukder, M.H.; Mostafiz, R. NeuroYOLO: A Lightweight YOLOv10-MobileNetV3 Framework for Real-Time Brain Tumor Detection in MRI Scans. *Biomed. Mater. Devices* **2025**, *4*, 4592–4617. [[CrossRef](#)]
22. Chen, J.; Hu, Y.; Yang, T.; Sun, Z.; Xie, L.; Zhao, H. YOLO-LS: A novel deep learning framework for brain tumor segmentation in Magnetic Resonance Imaging. *Sci. Rep.* **2026**, *16*, 20151. [[CrossRef](#)] [[PubMed](#)]
23. Mustafa, R.; Jiskani, S.M.; Panhwar, M.A.; Narejo, S.; Ahmed, Z. YOLO-Scan: A Real-Time Deep Learning Framework for Brain Tumor Detection in MRI. *Mehran Univ. Res. J. Eng. Technol.* **2026**, *45*, 62–73. [[CrossRef](#)]
24. Eun, S.; Jee, S.H.; Kim, Y.C. Effects of Brain Tumor Size and Intersection-Over-Union Threshold on the Tumor Detection and Classification Performance in YOLO Models. *J. Med. Biol. Eng.* **2026**, *46*, 278–288. [[CrossRef](#)]
25. Zhou, H.; Vallières, M.; Bai, H.X.; Su, C.; Tang, H.; Oldridge, D.; Zhang, Z.; Xiao, B.; Liao, W.; Tao, Y.; et al. MRI features predict survival and molecular markers in diffuse lower-grade gliomas. *Neuro-Oncology* **2017**, *19*, 862–870. [[CrossRef](#)] [[PubMed](#)]
26. Lambin, P.; Leijenaar, R.T.; Deist, T.M.; Peerlings, J.; De Jong, E.E.; Van Timmeren, J.; Sanduleanu, S.; Larue, R.T.; Even, A.J.; Jochems, A.; et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 749–762. [[CrossRef](#)]
27. Li, Z.; Wang, Y.; Yu, J.; Guo, Y.; Cao, W. Deep learning based radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. *Sci. Rep.* **2017**, *7*, 5467. [[CrossRef](#)] [[PubMed](#)]
28. Li, J.; Liu, R.; Xing, Y.; Gao, X.; Yin, Q.; Su, Q. Foundation model based on routine magnetic resonance imaging for brain tumor molecular profiling and progression prediction. *JCO Precis. Oncol.* **2026**, *10*, e2500930. [[CrossRef](#)] [[PubMed](#)]
29. Nalentzi, K.; Gerogiannis, K.; Bougias, H.; Stogiannos, N.; Papavasileiou, P. Comparative analysis of transformer-based deep learning models for glioma and meningioma classification. *J. Med. Imaging Radiat. Sci.* **2025**, *56*, 102008. [[CrossRef](#)] [[PubMed](#)]
30. Reis, H.C.; Turk, V. Advanced brain tumor analysis: A novel strategy for segmentation and classification using modern computational methods. *Neural Comput. Appl.* **2025**, *37*, 4697–4731.
31. Ketabi, S.; Wagner, M.W.; Hawkins, C.; Tabori, U.; Ertl-Wagner, B.B.; Khalvati, F. Multimodal contrastive learning for enhanced explainability in pediatric brain tumor molecular diagnosis. *Sci. Rep.* **2025**, *15*, 10943. [[CrossRef](#)]
32. Brändl, B.; Steiger, M.; Kubelt, C.; Rohrandt, C.; Zhu, Z.; Evers, M.; Wang, G.; Schuldt, B.; Afflerbach, A.K.; Wong, D.; et al. Rapid brain tumor classification from sparse epigenomic data. *Nat. Med.* **2025**, *31*, 840–848. [[CrossRef](#)] [[PubMed](#)]
33. Cheng, J.; Liu, J.; Kuang, H.; Wang, J. A fully automated multimodal MRI-based multi-task learning for glioma segmentation and IDH genotyping. *IEEE Trans. Med. Imaging* **2022**, *41*, 1520–1532. [[CrossRef](#)]
34. Ma, M.; Xi, Y.; Li, R.; Liu, X.; Hu, Z.; Li, Y.; Wu, X.; Hu, W.; Kang, X.; Yin, H.; et al. MTAF-Net: Multi-task learning for automated glioma segmentation and IDH genotyping in multi-parametric MRI. *Biomed. Signal Process. Control* **2026**, *112*, 108382. [[CrossRef](#)]
35. Zhu, J.; Liao, Y.; Chen, Y.; Gao, F.; Li, R.; Yang, G.; Lai, X.; Chen, J. Multimodal MRI-Based Glioma Segmentation and MGMT Promoter Methylation Status Prediction Using Multitask Learning Architecture. *Int. J. Imaging Syst. Technol.* **2025**, *35*, e70173. [[CrossRef](#)]
36. Bjørkeli, E.B.; Esmaili, M. Multi-task glioma segmentation and IDH mutation and 1p19q codeletion classification via a deep learning model on multimodal MRI. *Meta-Radiology* **2025**, *3*, 100152. [[CrossRef](#)]
37. Iqbal, M.S.; Bajwa, U.I.; Raza, R.; Anwar, M.W. Integrated brain tumor segmentation and MGMT promoter methylation status classification from multimodal MRI data using deep learning. *Digit. Health* **2025**, *11*, 20552076251332018. [[CrossRef](#)] [[PubMed](#)]
38. Rhanoui, M.; Alaoui Belghiti, K.; Mikram, M. Multi-Task Deep Learning for Simultaneous Classification and Segmentation of Cancer Pathologies in Diverse Medical Imaging Modalities. *Onco* **2025**, *5*, 34. [[CrossRef](#)]
39. Shu, X.; Li, Z.; Tian, C.; Chang, X.; Yuan, D. An active learning model based on image similarity for skin lesion segmentation. *Neurocomputing* **2025**, *630*, 129690. [[CrossRef](#)]
40. Xiong, Y.; Shu, X.; Liu, Q.; Yuan, D. HCMNet: A hybrid cnn-mamba network for breast ultrasound segmentation for consumer assisted diagnosis. *IEEE Trans. Consum. Electron.* **2025**, *3*, 8045–8054. [[CrossRef](#)]
41. Shinohara, R.T.; Sweeney, E.M.; Goldsmith, J.; Shiee, N.; Mateen, F.J.; Calabresi, P.A.; Jarso, S.; Pham, D.L.; Reich, D.S.; Crainiceanu, C.M.; et al. Statistical normalization techniques for magnetic resonance imaging. *Neuroimage Clin.* **2014**, *6*, 9–19. [[CrossRef](#)] [[PubMed](#)]
42. Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; March, M.; Lempitsky, V. Domain-adversarial training of neural networks. *J. Mach. Learn. Res.* **2016**, *17*, 1–35.
43. Varsavsky, T.; Orbes-Arteaga, M.; Sudre, C.H.; Graham, M.S.; Nachev, P.; Cardoso, M.J. Test-time unsupervised domain adaptation. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, 4–8 October 2020*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 428–436.

44. Canisius, J.; Buchner, J.; Rosier, M.; Griessmair, M.; Peeken, J.; Kirschke, J.S.; Piraud, M.; Bakas, S.; Menze, B.; Wiestler, B.; et al. GlioMODA: Robust Glioma Segmentation in Clinical Routine. *medRxiv* **2025**. [[CrossRef](#)]
45. Pati, S.; Baid, U.; Edwards, B.; Sheller, M.J.; Foley, P.; Anthony Reina, G.; Thakur, S.; Sako, C.; Bilello, M.; Davatzikos, C.; et al. The federated tumor segmentation (FeTS) tool: An open-source solution to further solid tumor research. *Phys. Med. Biol.* **2022**, *67*, 204002. [[CrossRef](#)]
46. Reddy, D.D.; Saadat, N.; Holcomb, J.M.; Wagner, B.C.; Truong, N.C.; Bowerman, J.; Hatanpaa, K.J.; Patel, T.R.; Pinho, M.C.; Yu, F.; et al. The University of Texas Southwestern Glioma Dataset-MRI, Molecular Markers and Segmentations. *Sci. Data* **2026**, *13*, 934. [[CrossRef](#)] [[PubMed](#)]
47. Kim, B.H.; Lee, H.; Choi, K.S.; Nam, J.G.; Park, C.K.; Park, S.H.; Chung, J.W.; Choi, S.H. Validation of MRI-based models to predict MGMT promoter methylation in gliomas: BraTS 2021 radiogenomics challenge. *Cancers* **2022**, *14*, 4827. [[CrossRef](#)] [[PubMed](#)]
48. Han, Y.; Yan, L.F.; Wang, X.B.; Sun, Y.Z.; Zhang, X.; Liu, Z.C.; Nan, H.Y.; Hu, Y.C.; Yang, Y.; Zhang, J.; et al. Structural and advanced imaging in predicting MGMT promoter methylation of primary glioblastoma: A region of interest based analysis. *BMC Cancer* **2018**, *18*, 215. [[CrossRef](#)] [[PubMed](#)]
49. Verger, A.; Couturier, O. Rôle de la TEP en neuro-oncologie: Apport des traceurs FDG, aux acides aminés et des récepteurs de la somatostatine dans la prise en charge des tumeurs cérébrales. *Méd. Nucl.* **2025**, *49*, 325–329. [[CrossRef](#)]
50. van der Voort, S.R.; Incekara, F.; Wijnenga, M.M.; Kapsas, G.; Gahrman, R.; Schouten, J.W.; Nandoe Tewarie, R.; Lycklama, G.J.; De Witt Hamer, P.C.; Eijgelaar, R.S.; et al. Combined molecular subtyping, grading, and segmentation of glioma using multi-task deep learning. *Neuro-Oncology* **2023**, *25*, 279–289. [[PubMed](#)]
51. Wen, S.; Yang, S.; Xu, L.; Yang, Y.; Qi, Y.; Hu, P.; Chen, Q.; Zhang, D. A one-stage multi-task network for molecular subtyping, grading, and segmentation of glioma. *Biomed. Signal Process. Control* **2025**, *108*, 107923. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.