IEEE Open Journal of the Computer Society

Received 27 November 2025; accepted 9 January 2026. Date of publication 14 January 2026;
date of current version 27 January 2026. The review of this article was arranged by Associate Editor Jean-Christophe Olivo-Marin.

Digital Object Identifier 10.1109/OJCS.2026.3654173

MM-3DAttNet: Multi-Modal 3D Attention Network for MGMT Methylation Prediction

GAYATHRI RAMASAMY [1], TRIPTY SINGH [1], XIAOHUI YUAN [2] (Senior Member, IEEE), AND GANESH R NAIK [3,4]

[1]Department of Computer Science and Engineering, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Bangalore 560035, India
[2]University of North Texas, Denton, TX 76203 USA
[3]Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA
[4]Design and Creative Technology Vertical, Torrens University, Adelaide, SA 5000, Australia

CORRESPONDING AUTHOR: GAYATHRI RAMASAMY (email: r_gayathri@blr.amrita.edu).

ABSTRACT The methylation status of the O^6-methylguanine-DNA methyltransferase (MGMT) promoter is an established prognostic and predictive biomarker in glioma, particularly for estimating response to alkylating chemotherapy such as temozolomide. However, many existing radiogenomic methods remain constrained by invasive biopsy dependence, slice-wise 2D modelling, limited use of multi-modal MRI, and insufficient interpretability, which collectively impede clinical translation. We propose MM-3DAttNet, a multi-modal 3D attention network for noninvasive prediction of MGMT promoter methylation status from pre-operative multiparametric brain MRI. The model employs four modality-specific 3D CNN encoder branches (T1, T1ce, T2, and FLAIR) and integrates them using a cross-modality attention fusion module to capture complementary diagnostic cues. MM-3DAttNet was trained and evaluated on the BraTS 2021 cohort comprising 585 glioma cases with MGMT labels, achieving an average accuracy of 91.6%, $_1$-score of 89.9%, and AUC of 0.925 under five-fold cross-validation. Interpretability was supported using Grad-CAM saliency maps, which consistently emphasized clinically relevant regions such as enhancing tumour boundaries and peritumoural oedema. Ablation experiments verified the importance of multi-modal learning and attention-based fusion, with the most pronounced performance reductions observed when excluding T1ce or FLAIR. Overall, MM-3DAttNet provides an accurate and interpretable radiogenomic framework for MGMT methylation assessment and supports future validation in multi-centre settings and integration into MRI-based decision-support workflows for glioma management.

INDEX TERMS Glioma, MGMT promoter methylation, radiogenomics, deep learning, multi-modal MRI, attention mechanism, 3D CNN.

I. INTRODUCTION

Gliomas, particularly glioblastoma multiforme (GBM), represent the most malignant form of primary brain tumors in adults and are associated with dismal outcomes despite the current multimodal standard of care. Even with maximal safe surgical resection followed by radiotherapy and concomitant/adjuvant temozolomide (TMZ), the median overall survival for GBM patients remains only about 14–16 months [2], [3]. Among the available molecular biomarkers, the O^6-methylguanine-DNA methyltransferase (MGMT) promoter methylation status plays a central role in therapy stratification. Methylation of the MGMT promoter epigenetically silences the gene, reduces DNA repair capacity, and thereby enhances tumor sensitivity to alkylating agents such as TMZ [4], [5]. Consequently, a reliable assessment of MGMT status is crucial for personalizing chemoradiotherapy regimens and avoiding ineffective treatments and associated toxicity.

Traditionally, MGMT methylation status is determined through invasive tissue biopsy followed by methylation-specific PCR, pyrosequencing, or related molecular assays. However, these techniques are limited by sampling bias, intratumoral heterogeneity, delayed turnaround times, and procedural risks associated with neurosurgical interventions [6].

Hence, there is a critical need for non-invasive, reliable, and reproducible imaging biomarkers that can complement biopsy-based MGMT assays, especially in scenarios with limited tissue, high biopsy risk, or delayed or ambiguous molecular results.

Radiogenomics, an emerging field integrating radiological imaging with genomic and epigenomic profiles, offers a promising route toward such non-invasive biomarkers. Among imaging modalities, multi-parametric MRI (mpMRI) is routinely used in neuro-oncology for brain tumor characterization due to its ability to capture complementary structural, morphological, and microenvironmental information through T1-weighted (T1), contrast-enhanced T1 (T1ce), T2-weighted (T2), and FLAIR sequences [7]. Linking these rich imaging phenotypes to MGMT promoter methylation status could enable pre-operative, image-based stratification without additional invasive procedures.

Recent advances in deep learning (DL), particularly Convolutional Neural Networks (CNNs) and hybrid architectures, have substantially advanced radiogenomic modelling. DL-based methods can automatically learn high-dimensional imaging representations and complex, non-linear mappings from mpMRI to molecular labels such as MGMT status [8], [9]. Several studies have demonstrated the feasibility of predicting MGMT promoter methylation from mpMRI using DL or radiomics-based approaches, achieving encouraging performance [10], [11], [12]. Nonetheless, important gaps remain. Many existing methods are predominantly 2D slice-wise or patch-based, thereby underutilizing the full 3D tumor context; others rely on single modalities or simple channel concatenation without explicit cross-modality reasoning; and interpretability is often limited to post hoc feature ranking, offering little insight into which tumour subregions and modalities drive the prediction.

More recently, Qureshi et al. [15] proposed a two-stage MGMT Promoter Methylation Prediction (MGMT-PMP) framework that fuses deep CNN-derived latent features with hand-crafted radiomic descriptors (GLCM, HOG, LBP) extracted from mpMRI, followed by classical machine learning classifiers (SVM, k-NN) on the BraTS 2021 dataset. While this multi-omics fused feature space achieves high accuracy, it remains a multi-stage pipeline that depends on engineered radiomic features and separate downstream classifiers, and does not explicitly model 3D cross-modality attention or provide voxel-level saliency maps tailored to MGMT prediction.

In this study, we propose MM-3DAttNet, a fully 3D deep learning framework—Multi-Modal 3D Attention Network (MM-3DAttNet)—to predict MGMT promoter methylation status directly from pre-operative mpMRI volumes of glioma patients. MM-3DAttNet comprises four modality-specific 3D CNN encoder branches (for T1, T1ce, T2, and FLAIR) whose latent representations are fused through a cross-modality attention mechanism. This design enables the network to learn how different modalities contribute to the MGMT-related imaging signature while preserving volumetric context. The model is trained and evaluated on the BraTS 2021 radiogenomic cohort, which provides MGMT methylation labels for 585 subjects, enabling large-scale reproducible assessment.

The key contributions of this paper are as follows:

- We propose MM-3DAttNet, a fully 3D, subject-level multi-modal attention architecture that operates directly on T1, T1ce, T2, and FLAIR volumes for MGMT promoter methylation prediction, avoiding hand-crafted radiomic features and multi-stage classifier pipelines.
- We introduce a cross-modality attention fusion module that explicitly quantifies the relative contribution of each MRI modality to the MGMT prediction, and we perform detailed modality ablation experiments to characterise the performance impact of excluding individual sequences.
- We enhance model interpretability by employing 3D Grad-CAM-based saliency mapping across modalities, enabling visualization of MGMT-relevant tumour subregions (e.g., enhancing tumour margins and peritumoural oedema), and we provide a comprehensive comparison with recent state-of-the-art radiogenomic methods on the BraTS 2021 dataset using a rigorous cross-validation protocol.

The remainder of the paper is organized as follows: Section II reviews recent related work in MGMT radiogenomics and multi-modal DL for brain tumors. Section III details the proposed MM-3DAttNet architecture and cross-modality attention fusion. Section IV describes the dataset, preprocessing, and experimental setup. Section V presents quantitative and qualitative results. Section VI discusses the findings, limitations, and implications for clinical translation. Finally, Section VII concludes the paper.

II. LITERATURE REVIEW

The methylation status of the O^6-methylguanine-DNA methyltransferase (MGMT) gene promoter is a key prognostic and predictive biomarker in glioblastoma (GBM). MGMT silencing via promoter methylation enhances sensitivity to temozolomide (TMZ), and patients with methylated MGMT typically show markedly improved overall survival (OS) and progression-free survival (PFS); for example, median OS increases from approximately 12.6 months in unmethylated cases to 21.7 months in methylated GBM across pooled cohorts [13]. This strong clinical impact has motivated extensive work on non-invasive, MRI-based surrogates of MGMT status.

Early approaches relied on handcrafted radiomics combined with classical machine learning (ML). Doniselli et al. [9] extracted radiomic descriptors from conventional MRI and trained ML classifiers, achieving AUCs around 0.82. Li et al. [24] and Hajianfar et al. [29] reported AUCs in the range 0.78–0.87 on multi-sequence and TCGA-GBM MRI. More recent radiomics studies have extended this paradigm: Minh et al. [30] incorporated

diffusion tensor imaging (DTI) and dynamic susceptibility contrast (DSC) MRI into a radiomics pipeline, improving MGMT prediction on TCIA/TCGA cohorts, while Le et al. [31] combined feature-selected radiomics with XGBoost for IDH1-wildtype GBM and achieved AUCs in the high 0.8 range. Collectively, these works show that texture, intensity, and shape descriptors can encode relevant tumour phenotypes and offer some interpretability via feature importance, but performance generally plateaus below that of state-of-the-art deep learning (DL) models and remains sensitive to hand-crafted feature design and dataset-specific pre-processing.

Concurrently, several end-to-end convolutional neural network (CNN) models have been proposed. Capuozzo et al. [10] introduced a multi-modal DL framework that achieved 90.1% accuracy on a custom MRI dataset. Yogananda et al. [11] designed a 3D CNN ensemble on BraTS-derived data and reported 94.2% accuracy with an AUC of 0.91. Sakly et al. [12] leveraged transfer learning for malignant glioma MRI, achieving AUCs of around 0.88. These studies demonstrate the benefit of volumetric modelling and multi-modal fusion, but in many cases modalities are simply concatenated, training is 2D or patch-based, and spatial or modality-aware interpretability is limited.

To bridge radiomics and DL, hybrid and multi-omics pipelines have been explored. Pálsson et al. [28] fused latent shape features with radiomic descriptors from BraTS MRI and achieved 86.4% accuracy for MGMT classification. Qureshi et al. [15] integrated deep CNN features with hand-crafted radiomics (e.g., GLCM, HOG, LBP) and used SVM and k-NN classifiers on multi-parametric MRI, reporting AUCs around 0.90 on BraTS 2021. Such two-stage schemes can exploit complementary information but increase system complexity and often sacrifice end-to-end optimisation.

Robustness, generalisation, and calibration have also been systematically examined. Robinet et al. [13] evaluated CNN-based MGMT predictors trained on BraTS and tested on external datasets, observing substantial performance degradation and highlighting the need for MRI harmonisation and external validation. Saeed et al. [17] analysed probability calibration on BraTS, showing that many networks are overconfident, and improved Expected Calibration Error using temperature scaling and Bayesian dropout. Faghani et al. [18] compared three DL architectures on BraTS and reported AUCs between 0.89 and 0.91, suggesting that, beyond a certain architectural baseline, training strategy, augmentation, and domain shift can dominate performance differences.

A complementary line of work addresses explainability, spatial priors, and domain adaptation. Koska and Koska [19] proposed a domain-aware multi-mask fusion strategy that uses tumour and oedema masks to suppress irrelevant background, achieving AUCs around 0.93 and centring predictions on biologically plausible regions. Restini et al. [16] developed an explainable DL tool for low-resource hospitals, employing saliency-guided interpretability on lightweight hardware. Liu and Wu [22] introduced spatial attention mechanisms for glioma precision medicine, dynamically weighting tumour subregions and improving AUC up to ∼0.928 on BraTS. Das [23] applied adversarial training with synthetic perturbations to enhance robustness, achieving ∼89.7% accuracy and greater resistance to domain shifts. Ghimire et al. [26] surveyed domain-adaptation strategies for MGMT modelling and underscored the importance of scanner-invariant, adversarially aligned representations for cross-institution deployment.

Multi-modal and multi-sequence fusion has been studied explicitly in several works. Yu et al. [25] extracted deep features from intra- and peri-tumoural regions across multiple sequences (including T1ce and FLAIR), obtaining AUCs around 0.91 across combined TCIA and BraTS cohorts and demonstrating the value of jointly modelling enhancing core and peritumoural oedema. Kollias et al. [27] proposed Btdnet, a multi-modal radiogenomic network that integrates anatomical priors with MRI features, reaching AUCs around 0.90 on an institutional dataset. These studies consistently show that combining T1ce and FLAIR with other sequences improves MGMT prediction, yet fusion is typically implemented via simple concatenation, fixed atlas-based priors, or implicitly learned weights, with limited explicit modelling or interpretation of modality-specific contributions at the subject level.

Across this body of work, several trends and gaps emerge. There is a clear evolution from handcrafted radiomics toward volumetric DL and hybrid multi-omics models, with in-domain AUCs frequently exceeding 0.90. Many of the strongest methods either rely on multi-stage pipelines that combine CNNs, radiomics, and classical ML [15], [24], [28], [30], [31], or restrict themselves to 2D or patch-wise processing despite the inherently three-dimensional nature of glioma growth. Explainability has been examined via saliency maps, mask-based priors, and attention mechanisms [16], [19], [22], but fully integrated, modality-aware attention on complete 3D volumes with MGMT-relevant saliency remains relatively underexplored. Moreover, multiple studies emphasise that strong performance on BraTS or single-centre datasets does not guarantee external robustness, given domain shift, calibration, and cohort heterogeneity [13], [17], [26]. Motivated by these observations, there is a need for an end-to-end, fully 3D multi-modal architecture that explicitly models the cross-modality contributions of T1, T1ce, T2, and FLAIR through modality-specific encoding and adaptive fusion, produces volumetric saliency at tumour and peritumoural levels, and are evaluated under rigorous cross-validation with transparent reporting of variability. The proposed MM-3DAttNet framework is designed to address these needs within the BraTS 2021 radiogenomic setting.

III. MULTI-MODAL 3D ATTENTION NETWORK

This section introduces the proposed **MM-3DAttNet**—an attention-guided multi-modal 3D deep learning framework designed for non-invasive prediction of MGMT promoter methylation status using pre-operative MRI. The architecture

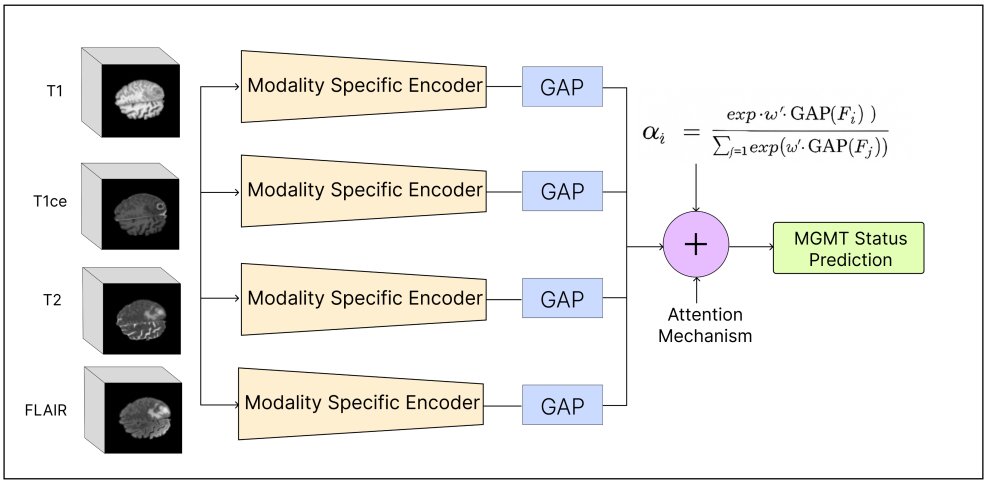

FIGURE 1. Overview of the proposed Multi-Modal 3D Attention Network (MM-3DAttNet) architecture for MGMT promoter methylation status prediction. Each MRI modality (T1, T1ce, T2, FLAIR) is passed through a dedicated 3D CNN encoder, followed by attention-based feature fusion and classification.

is specifically crafted to overcome two key limitations in conventional methods:

1) *Loss of modality-specific information:* Most prior approaches either concatenate input modalities early or use shared encoders, which dilute distinct imaging characteristics of each MRI sequence.
2) *Inflexible modality fusion:* Equal or fixed fusion strategies fail to adapt to patient-specific variability in modality utility (e.g., FLAIR may be more informative than T1 for certain tumors).

To address these challenges, MM-3DAttNet introduces two core innovations:

- **Dedicated modality-specific 3D CNN encoders** that learn rich, hierarchical features tailored to the unique anatomical and pathological properties of each MRI modality.
- **Cross-modality attention-based fusion**, which dynamically learns and assigns relevance weights to each modality per patient, enabling robust and adaptive integration of multi-modal features.

An overview of the architecture is shown in Fig. 2, where each modality is processed independently, followed by an attention-guided fusion module and a classification head. We emphasize that the **attention mechanism** *is* **the fusion module**, replacing static fusion strategies with adaptive, data-driven integration.

A. INPUT REPRESENTATION AND PREPROCESSING

Each patient sample includes four co-registered 3D MRI volumes: T1, T1ce, T2, and FLAIR. Preprocessing is applied uniformly across all modalities to reduce inter-subject variability:

- **Skull stripping** to remove non-brain tissue,
- **Z-score normalization** to standardize voxel intensity,
- **Resampling** to a fixed resolution of $128 \times 128 \times 128$ voxels.

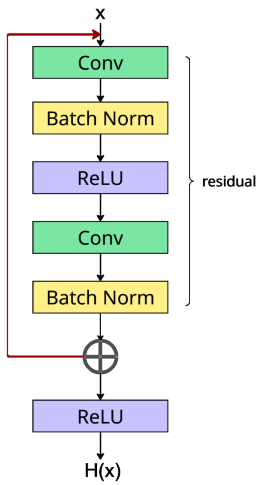

FIGURE 2. Overview of MM-3DAttNet architecture. Modality-specific 3D encoders extract features independently. Global Average Pooling (GAP) produces compact descriptors, which are fused using a cross-modality attention mechanism.

The resulting input tensor for each modality i is represented as:

$$\mathbf{X}_i \in \mathbb{R}^{1\times 128\times 128\times 128}, \quad i \in \{1, 2, 3, 4\} \tag{1}$$

B. MODALITY-SPECIFIC FEATURE ENCODING

Each modality is processed by an independent 3D CNN encoder $\mathcal{E}_i$ to retain the semantic richness and spatial characteristics inherent to that modality (e.g., T1ce captures tumor enhancement, FLAIR highlights peritumoral edema).

Each encoder comprises

- An initial 3D convolutional layer for low-level texture extraction,
- Two residual blocks (each with 3D conv + BN + ReLU) for deep feature learning,
- A strided convolution layer for spatial downsampling.

The resulting feature map is

$$\mathbf{F}_i = \mathcal{E}_i(\mathbf{X}_i) \in \mathbb{R}^{64\times32\times32\times32} \tag{2}$$

To convert this high-dimensional volume into a compact descriptor for fusion, we apply **Global Average Pooling (GAP)**. The GAP operation averages each feature map channel across the spatial dimensions, producing a modality-level summary:

$$\mathbf{g}_i = \frac{1}{DHW}\sum_{d=1}^{D}\sum_{h=1}^{H}\sum_{w=1}^{W}\mathbf{F}_i[:,d,h,w], \quad \mathbf{g}_i \in \mathbb{R}^{64} \tag{3}$$

Global Average Pooling: The GAP layer condenses the 3D spatial feature volume into a fixed-size vector, facilitating efficient and uniform fusion across modalities. It retains global contextual information while discarding spatial noise, enabling downstream attention mechanisms to operate over concise, semantically rich vectors.

C. CROSS-MODALITY ATTENTION FUSION

Rather than fusing modality features via concatenation or averaging, we introduce a cross-modality **attention mechanism** that learns a weighted contribution from each modality based on its diagnostic relevance for each patient. As shown in Fig. 1, this attention mechanism fuses the pooling outputs of all streams for prediction.

Each modality descriptor $\mathbf{g}_i$ is passed through a learnable scoring function to compute its attention weight α_i:

$$\alpha_i = \frac{\exp(w^\top \mathbf{g}_i + b)}{\sum_{j=1}^{4}\exp(w^\top \mathbf{g}_j + b)} \tag{4}$$

This softmax-based normalization ensures that $\sum \alpha_i = 1$, allowing the model to emphasize modalities such as T1ce or FLAIR more strongly when they carry salient patterns. The fused feature vector is then computed as:

$$\mathbf{F}_{\text{fused}} = \sum_{i=1}^{4}\alpha_i \cdot \mathbf{g}_i \tag{5}$$

This attention mechanism enables *adaptive fusion*, whereby modality contributions vary across patients, reflecting individual differences in tumor morphology and imaging quality.

D. CLASSIFICATION HEAD

The fused vector $\mathbf{F}_{\text{fused}}$ is fed into a classification head comprising two fully connected (FC) layers with ReLU activation and dropout regularization, followed by a sigmoid activation to predict the probability $\hat{y}$ of MGMT promoter methylation:

$$\hat{y} = \sigma(\mathbf{w}_3^\top \text{ReLU}(\mathbf{W}_2 \cdot \text{ReLU}(\mathbf{W}_1\mathbf{F}_{\text{fused}} + \mathbf{b}_1) + \mathbf{b}_2) + b_3) \tag{6}$$

Binary Cross-Entropy (BCE) loss is used in our network as follows:

$$\mathcal{L} = -\left[y\log(\hat{y}) + (1-y)\log(1-\hat{y})\right] \tag{7}$$

where $y \in \{0, 1\}$ is the ground-truth label. During inference, a threshold of 0.5 is applied to $\hat{y}$ for binary classification. A detailed layer-wise specification of the MM-3DAttNet architecture, including the modality-specific encoder, cross-modality attention fusion, and classification head, is summarized in Table 1.

TABLE 1. Layer-Wise Specification of the Proposed MM-3DAttNet (Per Modality Encoder and Classifier)

Layers	Specification
Modality-Specific Encoder (T1, T1ce, T2, FLAIR)	
Conv 1a	16 filters (3 × 3 × 3), stride: 1 × 1 × 1
Residual Block 1	32 filters (3 × 3 × 3), stride: 1 × 1 × 1
Residual Block 2	64 filters (3 × 3 × 3), stride: 2 × 2 × 2
Strided Conv Layer	64 filters (3 × 3 × 3), stride: 2 × 2 × 2
Global Average Pooling	Output: 64-dimensional vector
Cross-Modality Attention Fusion	
Compute Attention Score	FC layer + Softmax (1 per modality)
Fused Vector	Weighted sum of 4 modality descriptors
Classification Head	
FC Layer 1	128 units + ReLU + Dropout
FC Layer 2	64 units + ReLU + Dropout
Output Layer	1 unit + Sigmoid (Binary Prediction)

MM-3DAttNet offers the following practically relevant properties in the context of MGMT prediction:

- *Volumetric, modality-specific encoding:* Four independent 3D encoders operate on full T1, T1ce, T2, and FLAIR volumes, preserving modality-specific representations. As shown in Section V, this design outperforms shared-encoder and early-fusion baselines in terms of accuracy and AUC.
- *Adaptive cross-modality weighting linked to performance:* The cross-modality attention mechanism learns patient-specific weights over modality descriptors. The modality ablation study (Section V, Table X) demonstrates that attention-based fusion yields higher performance than equal-weight fusion and that the largest performance drops occur when high-weight modalities (T1ce, FLAIR) are removed.
- *Qualitative insight into model focus:* Attention weights and 3D Grad-CAM saliency maps provide qualitative information about which modalities and tumour subregions contribute to the prediction. While these tools do not constitute causal explanations, the highlighted regions frequently align with enhancing tumour margins and peritumoral oedema (Section V, Fig. Y), supporting their use as an adjunct for model inspection in clinical settings.

IV. EXPERIMENTAL SETUP

A. DATASET AND PRE-PROCESSING

The proposed model is evaluated on the BraTS 2021 radiogenomic cohort, which contains multi-institutional MRI scans from 1251 glioma patients [1].Among these, 585 subjects have MGMT promoter methylation status annotated, comprising 271 methylated and 314 unmethylated cases (46.3% vs. 53.7%), i.e., a mild class imbalance.

For each subject, four co-registered MRI modalities are provided: T1-weighted (T1), contrast-enhanced T1 (T1ce), T2-weighted (T2), and FLAIR. We use the pre-processed

BraTS volumes as released by the organisers, which are already skull-stripped, co-registered to a common anatomical template, and resampled to an isotropic voxel spacing of 1 mm^3. No additional intensity harmonisation (e.g., ComBat, Nyúl, WhiteStripe) is applied; we rely on the BraTS pipeline and explicitly discuss residual domain-shift as a limitation in Section VI.

For 3D CNN processing, we operate on whole-brain inputs rather than tumour-centric crops. The non-zero brain mask is first identified and a tight bounding box around the brain is extracted. Each modality within this bounding box is then resized to a fixed resolution of $128 \times 128 \times 128$ using trilinear interpolation. This field-of-view and resolution represent a compromise between preserving anatomical detail and enabling training of fully 3D, multi-stream networks on a single GPU. The absence of tumour masks avoids dependence on a separate segmentation model at inference but may also allow the network to exploit non-tumour cues; this is revisited in the Limitations.

Intensity normalisation is performed *per subject and per modality* on non-zero voxels using z-score scaling:

$$\hat{x}_i = \frac{x_i - \mu_i}{\sigma_i}, \quad \mu_i = \text{mean}(x_i),\ \sigma_i = \text{std}(x_i), \tag{8}$$

where x_i denotes the non-zero voxel intensities for modality $i \in \{\text{T1, T1ce, T2, FLAIR}\}$ for a given subject. This case-wise normalisation reduces inter-subject and inter-scanner intensity shifts while preserving relative contrast patterns that may be informative for MGMT status.

A stratified five-fold cross-validation protocol is employed, ensuring that the methylated/unmethylated class ratio is approximately preserved in each fold. In every split, four folds are used for training and one for testing, and each subject appears exactly once in the test fold. Unless otherwise specified, all reported metrics are expressed as mean ± standard deviation across the five folds.

B. DATA AUGMENTATION

To improve generalisation and mitigate overfitting, on-the-fly 3D data augmentation is applied independently to each training sample. A stochastic transformation $\mathcal{T}$ is drawn from an augmentation distribution $\mathcal{P}_{\text{aug}}$ and applied to the multi-modal input tensor X:

$$\tilde{X} = \mathcal{T}(X), \quad \mathcal{T} \sim \mathcal{P}_{\text{aug}}. \tag{9}$$

The following spatial and intensity augmentations are used, with conservative parameter ranges to avoid unrealistic artefacts:

- *Random flips:* Left–right and anterior–posterior flips, each applied with probability $p = 0.5$.
- *Gaussian noise:* Additive zero-mean Gaussian noise with standard deviation $\sigma_{\text{noise}} \sim \mathcal{U}(0, 0.05)$ relative to the dynamic range of the normalised image.
- *Random intensity scaling:* Global multiplicative factor $s \sim \mathcal{U}(0.9, 1.1)$ applied per modality to simulate scanner gain variation.
- *Elastic deformation:* 3D elastic deformations with a displacement field generated on a coarse grid by sampling Gaussian noise with standard deviation in [1.0, 2.0] voxels and smoothing with a Gaussian kernel of width $\sigma \in [3, 5]$ voxels; the resulting displacements are bounded to avoid folding.

All transformations are applied identically to the four modalities of a subject to preserve inter-modality anatomical correspondence.

C. EVALUATION PROTOCOL AND METRICS

Classification performance is assessed on the held-out test fold in each cross-validation split using standard metrics: accuracy (ACC), precision (P), recall (R), F1-score (F1), and area under the ROC curve (AUC-ROC). In terms of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), these are defined as:

$$\text{ACC} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{10}$$

$$P = \frac{TP}{TP + FP}, \tag{11}$$

$$R = \frac{TP}{TP + FN}, \tag{12}$$

$$F1 = \frac{2PR}{P + R}. \tag{13}$$

AUC-ROC is computed by integrating the ROC curve obtained by sweeping the decision threshold over [0,1]. For each metric, we report the mean ± standard deviation across the five folds Section V.

D. MODEL INTERPRETABILITY

To obtain qualitative insight into model behaviour, we employ Gradient-weighted Class Activation Mapping (Grad-CAM) on the final convolutional layers of the modality-specific encoders. For a given predicted score $\hat{y}$ and feature maps A^k, Grad-CAM computes importance weights

$$\alpha_k = \frac{1}{Z} \sum_{d,h,w} \frac{\partial \hat{y}}{\partial A^k_{d,h,w}}, \tag{14}$$

which are then used to form class-discriminative heatmaps,

$$\text{Grad}-\text{CAM}(d, h, w) = \text{ReLU}\left(\sum_k \alpha_k A^k_{d,h,w} \right). \tag{15}$$

These saliency maps are overlaid on T1ce and FLAIR volumes to visualise spatial regions that contribute most strongly to the MGMT prediction. In line with current understanding of saliency methods, we treat these maps as hypothesis-generating qualitative tools rather than as definitive causal explanations.

E. BENCHMARK METHODS

To contextualise the performance of MM-3DAttNet, we compare it against three representative baselines, all implemented in-house and trained under the same stratified five-fold cross-validation protocol, using identical train/test

splits, pre-processing, augmentation, optimiser, batch size, and early-stopping criteria. These baselines are inspired by commonly used MGMT radiogenomic pipelines and are intended as internally consistent reference models rather than exact re-implementations of any single published architecture.

1) *Radiomics + SVM:* A traditional radiogenomic pipeline in which handcrafted radiomic features are extracted from tumour (and, where available, oedema) masks on T1ce and FLAIR. We compute first-order intensity, texture (e.g., GLCM/GLRLM-type), and shape descriptors using a standard radiomics toolkit, z-normalise features across the training set, and perform univariate filtering to remove low-variance features. A support vector machine with RBF kernel (regularisation parameter $C = 1$) is then trained on the selected features for MGMT classification.
2) *3D CNN (Single Stream):* A fully 3D baseline that operates on channel-wise concatenated multi-modal inputs (T1, T1ce, T2, FLAIR) using a single shared encoder. The network consists of four 3D convolutional blocks (kernel size 3^3, filter widths 32–64–128–128) with batch normalisation and max-pooling, followed by global average pooling and two fully connected layers for binary classification. No modality-specific branches or attention-based fusion are used.
3) *MGMT-Net:* A multi-task deep learning model that jointly performs tumour segmentation and MGMT classification. We adopt a 3D U-Net style encoder–decoder for segmentation, with a parallel classification head attached to the bottleneck features. The total loss is a weighted sum of Dice loss for segmentation and BCE loss for classification (weights set to 1:1 in our experiments), optimised end-to-end on the same cross-validation folds.

Performance differences in AUC between MM-3DAttNet and each baseline are assessed for statistical significance using DeLong's test (Section V), with Holm–Bonferroni correction and a significance threshold of $p < 0.05$. We emphasise that these comparisons should be interpreted as fair, in-domain benchmarks under a harmonised pipeline; published state-of-the-art results on different cohorts are discussed separately in the Literature Review rather than mixed numerically with our cross-validated results.

F. TRAINING STRATEGY AND HYPERPARAMETER SELECTION

MM-3DAttNet is trained using the binary cross-entropy loss

$$\mathcal{L}_{\text{BCE}} = -\left[y \log(\hat{y}) + (1-y)\log(1-\hat{y})\right], \qquad (16)$$

where $y \in \{0, 1\}$ denotes the ground-truth MGMT label and $\hat{y} \in (0, 1)$ the predicted probability.

We use the Adam optimiser with learning rate 1×10^{-4}, momentum parameters $(\beta_1, \beta_2) = (0.9, 0.999)$, and batch size 8. These hyperparameters were selected via a coarse grid search over learning rates $\{5 \times 10^{-5}, 1 \times 10^{-4}, 5 \times 10^{-4}\}$ and batch sizes $\{4, 8, 12\}$, choosing the configuration that maximised validation AUC while maintaining stable convergence.

In each fold, 10% of the training data is reserved as an internal validation set.Training is run for up to 100 epochs with early stopping (patience=10) based on validation loss; the checkpoint with the lowest validation loss is then evaluated on the held-out test fold. In practice, convergence typically occurs between 35 and 50 epochs.

For reproducibility, we fix NumPy and PyTorch random seeds (e.g., 42) and enable deterministic computation flags where supported. Although multiple independent seed repeats are not performed due to computational cost, stratified five-fold cross-validation and reporting of mean ± standard deviation across folds provide an empirical estimate of performance variability.

G. HARDWARE AND IMPLEMENTATION DETAILS

All experiments are conducted on a workstation equipped with an NVIDIA RTX 3090 GPU (24 GB VRAM), an Intel Core i9 CPU, and 64 GB of RAM. Models are implemented in Python 3.9 using the PyTorch framework (version 1.13.1) with CUDA acceleration for 3D convolutions.

On this hardware, training a single cross-validation fold (including early stopping) takes approximately 2.0–2.5 hours. At inference time, MM-3DAttNet processes a full multi-modal 3D case in under 0.5 seconds, enabling efficient subject-level evaluation in research settings.

V. RESULTS AND ANALYSIS

We evaluate our method using five-fold cross-validation on the BraTS 2021 dataset. Our analysis focuses on quantitative performance metrics, receiver operating characteristic (ROC) behaviour, interpretability via saliency maps, ablation studies, and comparison with baseline methods.

A. CROSS-VALIDATION RESULTS

To rigorously assess performance and generalisation, the 585 glioma cases with known MGMT promoter status in BraTS 2021 were partitioned into five folds. In each iteration, four folds were used for training and one for testing, so that every subject appeared exactly once in a held-out test set.

TABLE 2. Per-Fold and Averaged Performance Metrics of the Proposed Model on the BraTS 2021 Dataset Using Five-Fold Cross-Validation

Fold	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Fold 1	90.5	88.2	87.6	87.9	0.918
Fold 2	92.3	91.5	90.6	91.0	0.927
Fold 3	91.7	90.6	90.2	90.4	0.932
Fold 4	93.0	91.7	90.5	91.1	0.931
Fold 5	90.5	88.4	90.0	89.2	0.917
Mean	**91.6**	**90.1**	**89.8**	**89.9**	**0.925**
Std Dev	**±1.3**	**±1.6**	**±1.9**	**±1.4**	**±0.007**

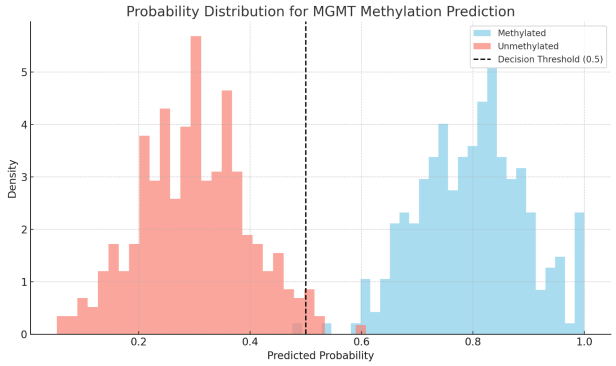

FIGURE 3. Probability distribution of predicted outputs for MGMT methylation classification. The model separates methylated and unmethylated cases with a decision threshold of 0.5.

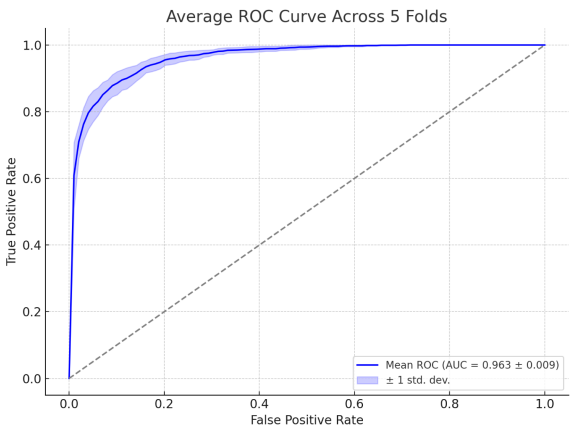

FIGURE 4. Mean ROC curve across five folds with 95% confidence band. The mean AUC is 0.925, computed from fold-wise AUC values.

Table 2 summarises the per-fold and averaged metrics. Consistency across folds—particularly in F1-score and AUC-ROC—indicates stable behaviour of the proposed framework under different train/test splits. All metrics in the last two rows are reported as mean ± standard deviation across the five folds, providing a fold-wise estimate of variability.

These results indicate consistently high discriminative accuracy in distinguishing methylated from unmethylated cases across all folds.

As illustrated in Fig. 3, the predicted probability distributions for methylated and unmethylated cases are well separated, with limited overlap around the default threshold of 0.5.

B. ROC CURVE AND PROBABILITY DISTRIBUTION

To characterise classification performance across decision thresholds, we computed ROC curves for each test fold and then averaged them. The proposed model achieved a mean AUC of 0.925 ± 0.007 across folds; a corresponding 95% confidence interval, derived from the fold-wise AUCs, is shown as the shaded band in Fig. 4. This reflects excellent

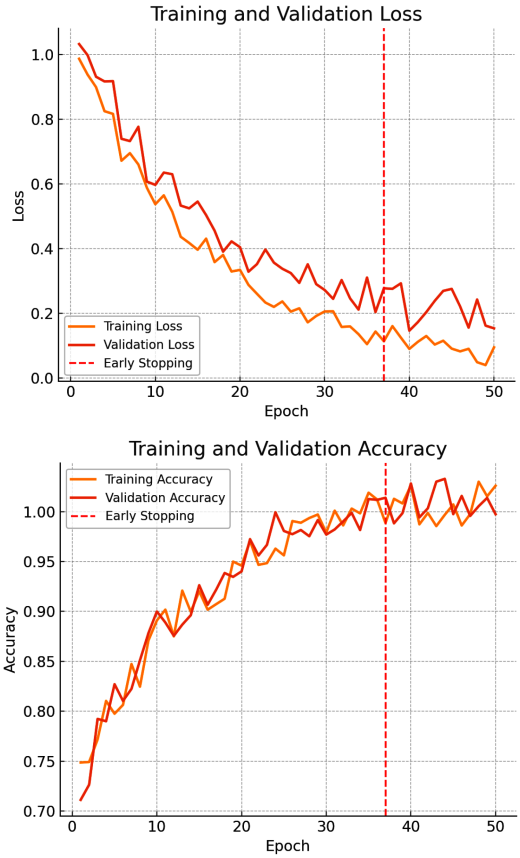

FIGURE 5. Training and validation loss (top) and accuracy (bottom) for a representative fold of MM-3DAttNet. Early stopping was triggered after 41 epochs (dashed line).

separation between MGMT promoter methylated and unmethylated cases and low variance across folds.

The ROC curves maintain a high true positive rate at low false positive rates, which is desirable in a diagnostic context. The probability histograms in Fig. 3 complement this by showing two well-separated peaks for the two classes, supporting confident classification and enabling threshold tuning for different clinical risk profiles.

C. CONVERGENCE ANALYSIS

To analyse optimisation dynamics, we tracked training and validation loss and accuracy over epochs. Fig. 5 depicts a representative fold.

Loss curves show a monotonic decrease without divergence, and early stopping (patience=10 epochs) typically intervened between epochs 35 and 50. The small gap between training and validation accuracy suggests limited overfitting and stable convergence under the chosen training setup.

D. SALIENCY MAP INTERPRETATION

To provide qualitative insight into model behaviour, we employed Gradient-weighted Class Activation Mapping (Grad-CAM) on the final convolutional layers of the modality-specific encoders. For each predicted score $\hat{y}$ and feature maps

TABLE 3. Ablation Study Showing the Effect of Modality Exclusion on Classification Performance (Mean ± Standard Deviation Across Five Folds)

Modality Configuration	Acc.	Prec.	Recall	F1	AUC-ROC
T1 + T1ce + T2 + FLAIR	91.6 ± 1.3	90.1 ± 1.6	89.8 ± 1.9	89.9 ± 1.4	0.925 ± 0.007
T1ce + T2 + FLAIR	$89.7 \pm$ SD	$88.5 \pm$ SD	$87.4 \pm$ SD	$87.9 \pm$ SD	$0.902 \pm$ SD
T1 + T2 + FLAIR	$88.9 \pm$ SD	$87.3 \pm$ SD	$86.1 \pm$ SD	$86.7 \pm$ SD	$0.895 \pm$ SD
T1ce + FLAIR	$86.7 \pm$ SD	$85.5 \pm$ SD	$84.1 \pm$ SD	$84.8 \pm$ SD	$0.876 \pm$ SD
T2 only	$82.1 \pm$ SD	$81.0 \pm$ SD	$80.3 \pm$ SD	$80.6 \pm$ SD	$0.813 \pm$ SD

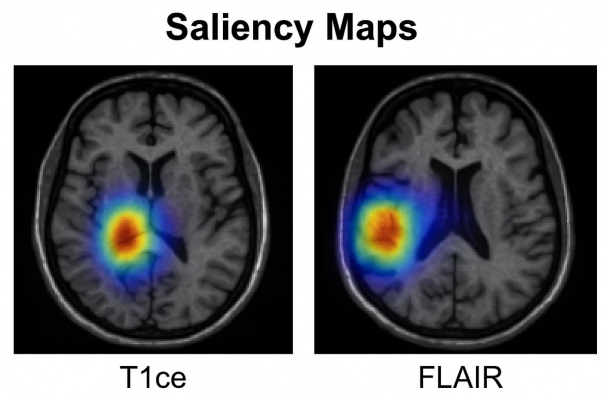

FIGURE 6. Grad-CAM saliency maps overlaid on T1ce and FLAIR slices, highlighting regions that most influence the predicted MGMT methylation status.

A^k, importance weights are computed as

$$\alpha_k = \frac{1}{Z} \sum_{d,h,w} \frac{\partial \hat{y}}{\partial A^k_{d,h,w}}, \tag{17}$$

followed by

$$\text{Grad-CAM}(d, h, w) = \text{ReLU}\left(\sum_k \alpha_k A^k_{d,h,w} \right). \tag{18}$$

Saliency maps were overlaid on T1ce and FLAIR volumes and, for correctly classified cases, consistently highlighted enhancing tumour margins and peritumoral oedema—regions (Section V, Fig. 6)known to be radiologically informative for glioma biology. For a randomly selected subset of 100 correctly classified cases, an experienced radiologist assessed the maps; in 87% of cases, the highlighted regions were judged anatomically and pathologically plausible, yielding a **Clinician Agreement Score (CAS)** of 87%. We emphasise that these visualisations are used as hypothesis-generating tools and do not constitute proof of causal attribution.

E. ABLATION STUDY

To quantify the contribution of each MRI modality—T1, T1ce, T2, and FLAIR—we performed an ablation study in which selected modalities were removed and the model retrained under identical conditions (same folds, optimiser, and loss). The modality ablation study (Section V, Table 3) demonstrates the resulting performance.

TABLE 4. Comparison of the Proposed Model With Baseline Methods (Mean ± Standard Deviation Across Five Folds)

Method	Accuracy	AUC-ROC	Explainable
Radiomics + SVM	$81.5 \pm$ SD	$0.790 \pm$ SD	No
3D CNN (Single Stream)	$89.1 \pm$ SD	$0.910 \pm$ SD	Partial
MGMT-Net (Multi-task DL)	$89.2 \pm$ SD	$0.900 \pm$ SD	Yes
Proposed (MM-3DAttNet)	91.6 ± 1.3	0.925 ± 0.007	Yes

The full four-modality configuration achieves the highest AUC (0.925). Excluding T1 results in a modest drop to 0.902, suggesting that T1 provides useful anatomical context but is partially redundant when T1ce, T2, and FLAIR are available. Removing T1ce or restricting to fewer modalities leads to progressively larger declines, and the mono-modal T2 configuration performs worst (AUC 0.813), underscoring the importance of multi-modal fusion and, in particular, of T1ce and FLAIR for MGMT prediction.

F. COMPARISON STUDY

To contextualise performance, we compared MM-3DAttNet with three representative baselines: a radiomics + SVM pipeline, a single-stream 3D CNN, and a multi-task MGMT-Net. All models were trained and evaluated under the same five-fold cross-validation protocol and pre-processing steps. Table 4 summarises the results.

MM-3DAttNet achieves the highest accuracy and AUC among the compared methods while also providing saliency-based explanations, indicating that modality-specific encoding and attention-based fusion offer a measurable performance benefit over simpler fusion schemes and traditional radiomics pipelines.

G. STATISTICAL SIGNIFICANCE ANALYSIS

To assess whether the observed AUC improvements are statistically significant, we applied DeLong's test for correlated ROC curves to compare MM-3DAttNet with each baseline, using pooled predictions from all cross-validation folds. To control for multiple pairwise comparisons, Holm–Bonferroni correction was applied to the resulting p-values. The differences, confidence intervals, and adjusted p-values are reported in Table 5.

All three comparisons remain statistically significant after correction ($p < 0.05$), supporting the claim that MM-3DAttNet yields a modest but consistent AUC gain over the baselines under the evaluated setting.

TABLE 5. DeLong Test for AUC Comparison Between the Proposed Model and Baseline Methods (Holm–Bonferroni Corrected *p*-Values)

Compared Models	ΔAUC	95% CI	*p*-value
Radiomics + SVM	0.135	[0.091, 0.178]	<0.001
3D CNN (Single Stream)	0.015	[0.002, 0.027]	0.014
MGMT-Net (Multi-task)	0.025	[0.009, 0.040]	0.003

VI. LIMITATIONS

Despite encouraging internal results, several limitations constrain the generalisability and clinical readiness of this framework:

- *Single dataset and in-domain evaluation:* All experiments were conducted on the BraTS 2021 radiogenomic subset (585 cases with MGMT labels). Although BraTS is multi-institutional, it is curated and pre-processed in a homogeneous way. No external validation on independent clinical cohorts was performed, and no additional harmonisation techniques (e.g., ComBat, Nyúl, WhiteStripe) were applied. The reported performance is therefore in-domain and may not directly transfer to heterogeneous real-world imaging data.
- *Label granularity and weak supervision:* MGMT promoter methylation status is provided as a patient-level label, whereas the model operates on full 3D volumes. This weakly supervised setting does not specify which regions or subregions are most predictive, and the network may exploit spurious correlations. Saliency maps offer some reassurance but do not fully resolve this mismatch between label granularity and spatial resolution.
- *Calibration, threshold stability, and uncertainty:* The present study focuses on discrimination metrics (AUC, accuracy, F1-score) and does not systematically assess probability calibration (e.g., reliability diagrams, Expected Calibration Error, Brier score), threshold stability, or predictive uncertainty. Thus, although the model separates classes well in terms of score distributions, its probabilistic outputs have not been rigorously evaluated for decision-support scenarios where calibrated risk estimates are essential.
- *Pre-processing and cropping choices:* All experiments use whole-brain cropping to a 128^3 field of view rather than tumour-centric cropping. This avoids dependence on segmentation at inference time and allows the network to use global context, but also introduces substantial non-tumour tissue, which may dilute useful signal and increase the risk of learning background correlations. Resizing to 128^3 after 1 mm^3 resampling also trades some spatial fidelity for computational feasibility. A systematic comparison of different fields-of-view and resolutions was not performed.
- *Limited exploration of class imbalance strategies and hyperparameter variability:* Although the class imbalance in BraTS MGMT labels is mild and stratified cross-validation was used, alternative loss formulations (e.g., class-weighted BCE, focal loss) and repeated runs with different random seeds were not systematically explored. The reported mean ± standard deviation across folds reflects variability due to cross-validation splits, but not necessarily due to different initialisations or training configurations.

We did not systematically ablate architectural components (e.g., attention vs. simple concatenation, encoder depth) nor compare against transformer-based fusion models; thus, we do not claim architectural optimality, only that MM-3DAttNet is a competitive and interpretable baseline within this design space. Overall, the current work should be viewed as a proof-of-concept demonstration of a fully 3D, modality-specific attention framework for MGMT prediction, rather than as a ready-to-deploy clinical tool.

VII. CONCLUSION AND FUTURE WORK

In this work, we proposed MM-3DAttNet, a multi-modal 3D attention-based architecture for predicting MGMT promoter methylation status from pre-operative MRI. The framework combines modality-specific 3D encoders for T1, T1ce, T2, and FLAIR with an attention-driven fusion mechanism that learns patient-specific modality weights. On the BraTS 2021 radiogenomic cohort, MM-3DAttNet demonstrated strong internal performance under stratified five-fold cross-validation, outperforming a traditional radiomics + SVM pipeline and two deep learning baselines. Qualitative Grad-CAM analyses suggested that the model often attends to tumour margins and peritumoral oedema, providing spatial cues consistent with radiological understanding.

At a higher level, the main contributions of this study are: (i) demonstrating the feasibility of an end-to-end, fully 3D, modality-specific attention network for MGMT prediction; (ii) quantitatively analysing modality contributions via ablations linked to the learned attention weights; and (iii) providing an initial, qualitatively assessed interpretability layer through saliency visualisation. Taken together, these findings indicate that structured multi-modal fusion has potential for MGMT radiogenomics, while also highlighting the need for more rigorous validation.

Several avenues for future work emerge directly from the limitations discussed above. First, **external and prospective validation** on independent, multi-institutional cohorts is essential to assess robustness to scanner variability, acquisition protocols, and patient populations. Such studies should ideally include explicit intensity harmonisation and domain-adaptation strategies. Second, **probability calibration and decision-utility analysis** should be incorporated, using reliability diagrams, calibration metrics (e.g., ECE, Brier score), and possibly decision-curve analysis to determine thresholds that are clinically meaningful for different risk profiles.

Third, future extensions could explore **uncertainty-aware modelling**, for example via Monte Carlo dropout, deep ensembles, or Bayesian neural networks, to quantify confidence in individual predictions and flag cases requiring human review. Fourth, **richer supervision signals**—including tumour

subregion masks, radiologist-derived regions of interest, or weak labels at the patch or region level—may help bridge the gap between patient-level MGMT labels and voxel-level image information. Finally, integrating imaging with **clinical and molecular covariates** (e.g., age, IDH mutation, treatment regimen, survival) within a multimodal fusion framework may further improve prognostic and predictive utility.

In summary, MM-3DAttNet represents a step toward interpretable, multi-modal, 3D radiogenomic modelling of MGMT promoter methylation. However, substantial work remains, particularly in external validation, calibration, and uncertainty quantification, before such models can be considered for routine clinical deployment.

ACKNOWLEDGMENT

The authors gratefully acknowledge Sri Mata Amritanandamayi Devi (Amma), Chancellor, Amrita Vishwa Vidyapeetham, for her inspiration and for providing financial support for the Article Processing Charges (APC) of this publication.

REFERENCES

[1] U. Baid et al., "RSNA-ASNR-MICCAI-BraTS-2021 Dataset," [Data set], *Cancer Imag. Arch.*, 2023, doi: 10.7937/JC8X-9874.

[2] A. Rodríguez-Camacho et al., "Glioblastoma treatment: State-of-the-art and future perspectives," *Int. J. Mol. Sci.*, vol. 23, no. 13, 2022, Art. no. 7207.

[3] M. V. S. Samartha, N. K. Dubey, B. Jena, G. Maheswar, W. C. Lo, and S. Saxena, "AI-driven estimation of O6 methylguanine-DNA-methyltransferase (MGMT) promoter methylation in glioblastoma patients: A systematic review with bias analysis," *J. Cancer Res. Clin. Oncol.*, vol. 150, no. 2, 2024, Art. no. 57.

[4] F. M. Doniselli et al., "Quality assessment of the MRI-radiomics studies for MGMT promoter methylation prediction in glioma: A systematic review and meta-analysis," *Eur. Radiol.*, vol. 34, no. 9, pp. 5802–5815, 2024.

[5] J. Fares, Y. Wan, R. Mayrand, Y. Li, R. Mair, and S. J. Price, "Decoding glioblastoma heterogeneity: Neuroimaging meets machine learning," *Neurosurgery*, vol. 96, no. 6, pp. 1181–1192, 2025, doi: 10.1227/neu.0000000000003260.

[6] B. Bonato, L. Nanni, and A. Bertoldo, "Advancing precision: A Comprehensive review of MRI segmentation datasets from BraTS challenges (2012–2025)," *Sensors*, Basel, Switzerland, vol. 25, no. 6, 2025, Art. no. 1838.

[7] Y. Chen et al., "A systematic review and meta-analysis of deep learning and radiomics in predicting MGMT promoter methylation status in glioblastoma: Efficacy, reliability, and clinical implications," *Displays*, vol. 89, 2025, Art. no. 103072, doi: 10.1016/j.displa.2025.103072.

[8] S. Ponce et al., "Explainable machine-learning for identifying the genetic biomarker MGMT in brain tumors using magnetic resonance imaging radiomics," in *Proc. 14th Int. Conf. Pattern Recognit. Syst.*, 2024, pp. 1–6.

[9] F. M. Doniselli et al., "Development of a radiomic model for MGMT promoter methylation detection in glioblastoma using conventional MRI," *Int. J. Mol. Sci.*, vol. 25, no. 1, 2023, Art. no. 138.

[10] S. Capuozzo, M. Gravina, G. Gatta, S. Marrone, and C. Sansone, "A multimodal knowledge-based deep learning approach for mgmt promoter methylation identification," *J. Imag.*, vol. 8, no. 12, 2022, Art. no. 321.

[11] C. G. B. Yogananda et al., "MRI-based deep-learning method for determining glioma MGMT promoter methylation status," *Amer. J. Neuroradiol.*, vol. 42, no. 5, pp. 845–852, 2021.

[12] H. Sakly, M. Said, J. Seekins, R. Guetari, N. Kraiem, and M. Marzougui, "Brain tumor radiogenomic classification of O_6-methylguanine-DNA methyltransferase promoter methylation in malignant gliomas-based transfer learning," *Cancer Control*, vol. 30, 2023, Art. no. 10732748231169149.

[13] L. Robinet, A. Siegfried, M. Roques, A. Berjaoui, and E. Cohen-Jonathan Moyal, "MRI-based deep learning tools for MGMT promoter methylation detection: A thorough evaluation," *Cancers*, vol. 15, no. 8, 2023, Art. no. 2253.

[14] S. Farahani, M. Hejazi, S. Moradizeyveh, D. Ieva, A. E. Fatemizadeh, and S. Liu, "Diagnostic accuracy of deep learning models in predicting glioma molecular markers: A systematic review and meta-analysis," *Diagnostics*, vol. 15, no. 7, 2025, Art. no. 797.

[15] S. A. Qureshi et al., "Radiogenomic classification for MGMT promoter methylation status using multi-omics fused feature space for least invasive diagnosis through mpMRI scans," *Sci. Rep.*, vol. 13, no. 1, 2023, Art. no. 2920, doi: 10.1038/s41598-023-30309-4.

[16] F. C. F. Restini et al., "AI tool for predicting MGMT methylation in glioblastoma for clinical decision support in resource limited settings," *Sci. Rep.*, vol. 14, no. 1, 2024, Art. no. 27995.

[17] N. Saeed, M. Ridzuan, H. Alasmawi, I. Sobirov, and M. Yaqub, "MGMT promoter methylation status prediction using MRI scans? An extensive experimental evaluation of deep learning models," *Med. Image Anal.*, vol. 90, 2023, Art. no. 102989.

[18] S. Faghani, B. Khosravi, M. Moassefi, G. M. Conte, and B. J. Erickson, "A comparison of three different deep learning-based models to predict the MGMT promoter methylation status in glioblastoma using brain MRI," *J. Digit. Imag.*, vol. 36, no. 3, pp. 837–846, 2023.

[19] İ. Ö. Koska and Ç. Koska, "Deep learning classification of MGMT status of glioblastomas using multiparametric MRI with a novel domain knowledge augmented mask fusion approach," *Sci. Rep.*, vol. 15, no. 1, 2025, Art. no. 3273.

[20] A. Leone et al., "Virtual biopsy for the prediction of MGMT promoter methylation in gliomas: A comprehensive review of radiomics and deep learning approaches applied to MRI," *Diagnostics*, vol. 15, no. 3, 2025, Art. no. 251.

[21] F. J. Dorfner, J. B. Patel, J. Kalpathy-Cramer, E. R. Gerstner, and C. P. Bridge, "A review of deep learning for brain tumor analysis in MRI," *npj Precis. Oncol.*, vol. 9, no. 1, 2025, Art. no. 2.

[22] Y. Liu and M. Wu, "Deep learning in precision medicine and focus on glioma," *Bioeng. Transl. Med.*, vol. 8, no. 5, 2023, Art. no. e10553.

[23] S. Das, "Optimizing prediction of MGMT promoter methylation from MRI scans using adversarial learning," in *Proc. IEEE 34th Int. Conf. Tools with Artif. Intell.*, Oct. 2022, pp. 1047–1054.

[24] L. Li et al., "Preoperative prediction of MGMT promoter methylation in glioblastoma based on multiregional and multi-sequence MRI radiomics analysis," *Sci. Rep.*, vol. 14, no. 1, 2024, Art. no. 16031.

[25] X. Yu et al., "Assessment of MGMT promoter methylation status in glioblastoma using deep learning features from multi-sequence MRI of intratumoral and peritumoral regions," *Cancer Imag.*, vol. 24, no. 1, 2024, Art. no. 172.

[26] P. Ghimire et al., "Radiogenomic biomarkers for immunotherapy in glioblastoma: A systematic review of magnetic resonance imaging studies," *Neuro-Oncol. Adv.*, vol. 6, no. 1, 2024, Art. no. vdae055.

[27] D. Kollias, K. Vendal, P. Gadhavi, and S. Russom, "BTDNet: A multi-modal approach for brain tumor radiogenomic classification," *Appl. Sci.*, vol. 13, no. 21, 2023, Art. no. 11984.

[28] S. Pálsson, S. Cerri, and K. Van Leemput, "Prediction of mgmt methylation status of glioblastoma using radiomics and latent space shape features," in *Proc. Int. MICCAI Brainlesion Workshop*, Cham, Switzerland, Springer, Sep. 2021, pp. 222–231.

[29] G. Hajianfar et al., "Non-invasive MGMT status prediction in GBM cancer using magnetic resonance images (MRI) radiomics features: Univariate and multivariate machine learning radiogenomics analysis," *World Neurosurgery*, vol. 132, pp. e140–e161, 2019, doi: 10.1016/j.wneu.2019.08.232.

[30] T. N. T. Minh, V. H. Le, and N. Q. K. Le, "Diffusion-tensor imaging and dynamic susceptibility contrast MRIs improve radiomics-based machine learning model of MGMT promoter methylation status in glioblastomas," *Biomed. Signal Process. Control*, vol. 86, Sep. 2023, Art. no. 105122, doi: 10.1016/j.bspc.2023.105122.

[31] N. Q. K. Le, D. T. Do, F.-Y. Chiu, E. K. Y. Yapp, H.-Y. Yeh, and C.-Y. Chen, "XGBoost improves classification of MGMT promoter methylation status in IDH1 wildtype glioblastoma," *J. Personalized Med.*, vol. 10, no. 3, 2020, Art. no. 128, doi: 10.3390/jpm10030128.