



OPEN ACCESS

EDITED BY

Christian F. Freyschlag,
Innsbruck Medical University, Austria

REVIEWED BY

Santiago Cepeda,
Hospital Universitario Rio Hortega, Spain
Andreas Holzinger,
Medical University Graz, Austria

*CORRESPONDENCE

Wenbin Li
✉ liwenbin@cmmu.edu.cn

RECEIVED 01 December 2023

ACCEPTED 19 January 2024

PUBLISHED 07 February 2024

CITATION

He Y, Duan L, Dong G, Chen F and
Li W (2024) Computational pathology-based
weakly supervised prediction model for
MGMT promoter methylation status in
glioblastoma.
Front. Neurol. 15:1345687.
doi: 10.3389/fneur.2024.1345687

COPYRIGHT

© 2024 He, Duan, Dong, Chen and Li. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Computational pathology-based weakly supervised prediction model for MGMT promoter methylation status in glioblastoma

Yongqi He¹, Ling Duan¹, Gehong Dong², Feng Chen¹ and Wenbin Li^{1*}

¹Department of Neuro-Oncology, Cancer Center, Beijing Tiantan Hospital, Capital Medical University, Beijing, China, ²Department of Pathology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China

Introduction: The methylation status of oxygen 6-methylguanine-DNA methyltransferase (MGMT) is closely related to the treatment and prognosis of glioblastoma. However, there are currently some challenges in detecting the methylation status of MGMT promoters. The hematoxylin and eosin (H&E)-stained histopathological slides have always been the gold standard for tumor diagnosis.

Methods: In this study, based on the TCGA database and H&E-stained Whole slide images (WSI) of Beijing Tiantan Hospital, we constructed a weakly supervised prediction model of MGMT promoter methylation status in glioblastoma by using two Transformer structure models.

Results: The accuracy scores of this model in the TCGA dataset and our independent dataset were 0.79 (AUC=0.86) and 0.76 (AUC=0.83), respectively.

Conclusion: The model demonstrates effective prediction of MGMT promoter methylation status in glioblastoma and exhibits some degree of generalization capability. At the same time, our study also shows that adding Patches automatic screening module to the computational pathology research framework of glioma can significantly improve the model effect.

KEYWORDS

computational pathology, glioblastoma, deep learning, MGMT, diagnostic

1 Introduction

Glioma is the most prevalent primary malignant tumor in the central nervous system. Glioblastoma (GBM) is the most common glioma subtype and an important cause of morbidity and mortality. It progresses rapidly and has the worst prognosis with a 5-year survival rate of less than 7% (1). For individuals newly diagnosed with GBM, the current standard treatment remains gross total resection followed by a combination of radiation therapy and temozolomide (TMZ) (2). O6-methylguanine-DNA methyltransferase (MGMT) is a DNA repair enzyme that reverses the DNA damage caused by alkylating agents, resulting in tumor resistance to TMZ and nitrosourea-based systemic therapy. Epigenetic silencing of the MGMT gene by promoter methylation makes the tumor more sensitive to treatment with alkylating agents and has been associated with longer overall survival in patients with GBM who received TMZ chemotherapy (3). There are multiple ways to detect MGMT promoter methylation, including methylation-specific PCR, methylation-specific high-resolution

melting, and pyrosequencing. However, these methods are often time-consuming and labor-intensive, and not all patients have the conditions to undergo relevant examinations. Obtaining an accurate diagnosis swiftly remains a significant challenge (4).

Histopathological biopsy, particularly hematoxylin and eosin (H&E)-stained slides, is still a gold standard for tumor diagnosis (5). However, accurate diagnosis of tumors requires high professionalism from pathologists. The clinical experience and subjectivity of the pathologist can also affect the diagnostic results (6). With the advancement of digital imaging and computer technology in recent years, computational pathology based on whole slide images (WSI) for artificial intelligence (AI)-assisted analysis is rapidly developing (7). For example, computational pathology can predict tumor classification (8, 9), prognosis (10), and molecular mutations based on histopathological images, etc. (11, 12). However, as far as we know, research regarding MGMT promoter methylation in glioma using pathological images remains relatively limited. Li introduced a weakly supervised primary brain tumor classifier VIT-WSI based on digital pathology slides. Within this framework, the MGMT promoter methylation status of gliomas was predicted with an accuracy of 0.7916 and an AUC of 0.845 (13). However, it is essential to note that the primary focus of the study was not on the prediction of MGMT promoter methylation. This study included only 71 patients with known MGMT methylation status, and 53 of those were determined by immunohistochemistry (IHC). Though IHC is a common and inexpensive assay used in clinical, the value of using IHC to determine MGMT status has been controversial. Due to the inconsistency between MGMT promoter methylation status and MGMT protein expression, the IHC results may not be accurate, thus affecting patients' treatment and prognosis (4). This indicates that it is of great clinical significance for us to develop a novel predictive model for the MGMT promoter methylation status. A study published in 2022 by Kim et al. (14) showed that MGMT promoter methylation was significantly associated with ATRX gene deletion in IDH wild-type glioblastoma. Moreover, ATRX is involved with the telomerase-independent alternative lengthening of telomeres (ALT) mechanism, and the ATRX gene is also involved in the regulation of the tumor microenvironment in glioma (15). Therefore, we hypothesized that MGMT promoter methylation in glioblastoma may cause subtle changes in tumor cell morphology through interactions with other genes, which can be captured by the neural network model to obtain corresponding predictions. Hence, we proposed a weakly supervised prediction model for MGMT promoter methylation status in glioblastoma based on computational pathology with H&E-stained histopathological slides. We used two models based on transformer architecture to achieve the end-to-end prediction of MGMT promoter methylation status and confirmed the generalization of the overall model using datasets from different sources. The contribution of this study is 2-fold: one is to confirm the possibility of predicting the methylation status of MGMT promoter in gliomas by WSI; second, an end-to-end model was established to simplify the cost of MGMT promoter status detection in glioma patients.

2 Materials and methods

2.1 Dataset

At present, the existence of an open database is very important. The complexity of cancer disease requires a large amount of data as a supporting basis, and the existence of an open database can accelerate related research and save researchers' resources (16). Therefore, this

retrospective study included histopathological slides of H&E-stained tissue samples from two cohorts of glioblastoma patients. The first cohort comprised patients from The Cancer Genome Atlas (TCGA) project, accessible through the National Institutes of Health Genome Data Sharing Portal (17–19) (TCGA cohort). The second cohort comprised patients with glioblastoma diagnosed in Beijing Tiantan Hospital from 2019 to 2023 (TianTan cohort). As cryopreservation destroys the original morphological structure of cells and tissues, all histopathological slides collected in this study were preserved by formalin-fixed paraffin embedding. Specific data inclusion and exclusion criteria are listed below:

Inclusion criteria were as follows: (1) precise diagnosis of glioblastoma by surgical pathology; (2) clear methylation status of the MGMT promoter by genetic testing; and (3) no other brain lesions. Exclusion criteria were as follows: (1) unclear images of histopathological slides; and (2) missing baseline information or unknown medical history.

All pathological images in the TCGA cohort were confirmed to be glioblastoma by secondary diagnosis by a pathologist with extensive experience in pathological diagnosis. All histopathological slides in the TianTan cohort were scanned using a Lecia Aperio CS2 scanner into WSIs of digital pathology in standard file format (the "svs" format). The scanning magnification was 20X. All data from the TianTan cohort were only used as an external independent test set to verify the generalization of the model and were not used in the training or hyperparameter adjustment process. Data in the TCGA cohort were divided into the training set and the test set according to certain standards: If multiple WSIs originated from the same patient, then all those WSIs from that patient were assigned to the training set. This approach ensured that slides from the same tumor sample were not used for both training and testing at the same time.

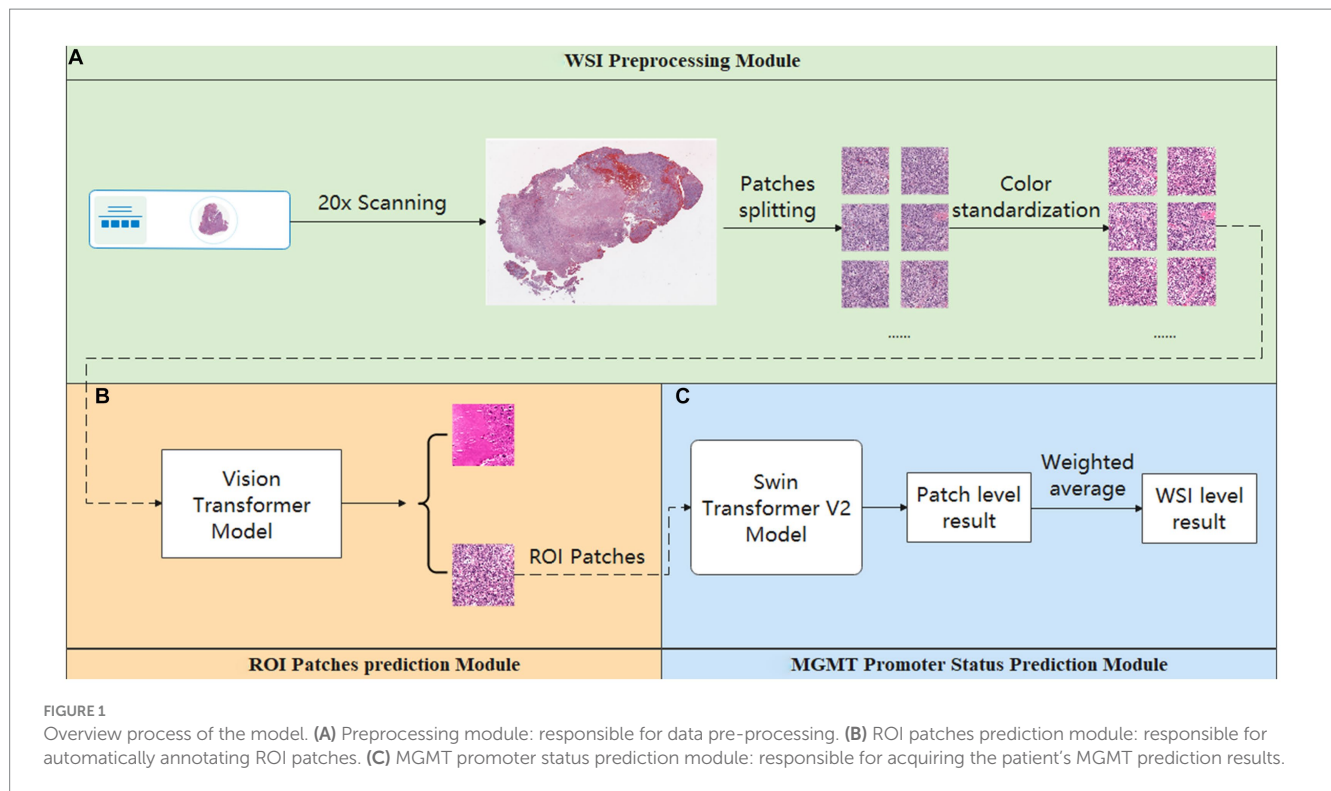
This study was approved by the Ethics Committee of Beijing Tiantan Hospital (ethical approval no: YW2022-025).

2.2 Overall model architecture

Our study achieved end-to-end prediction of MGMT promoter methylation status in glioma patients based on H&E-stained WSIs through three main modules. First, the scanned WSIs were divided into several patches using the preprocessing module. After that, the glioblastoma region of interest (ROI) patches prediction module was used to filter the sliced patches. Finally, the molecular status of patients was predicted by the MGMT promoter status prediction module. The main body model framework adopts the modular design of openMMLab (20). The details are described below. The complete process diagram is shown in Figure 1.

2.3 Preprocessing module

This module is crucial for pre-processing WSIs. It consisted of two parts. The first part is the patch segmentation. This part is responsible for splitting the complete WSI into patches. A complete WSI consists of billions of pixels through a pyramid structure. It is difficult to directly input such large image data into a neural network or machine learning model for training. Therefore, we used the sliding window method to divide the tissue image in WSI into several patches of 512 pixels × 512 pixels for subsequent data analysis and processing (21). The second part



is the color normalization. This part is dedicated to the color normalization of the patches. Color variation can arise due to several factors, including variations introduced by the operator during the H&E staining and differences in staining reagents. Such color variations can significantly interfere with subsequent modeling predictions. Thus, we used the Vahadane method to normalize the colors presented in H&E pathology images to eliminate the influence of staining differences, ensuring consistent and reliable data for further analysis (22).

2.4 ROI patches prediction module

This module is used to screen the patches obtained in the Preprocessing module. As a WSI may encompass diverse elements such as hemorrhage and normal tissue, potentially interfering with the final prediction, we employed the ROI patches prediction model to screen the patches generated in the preprocessing module to avoid the interference of confusing information as much as possible. The module utilized transfer learning, employing a Vision Transformer model pre-trained on the ImageNet dataset (23). We optimized the model's performance using the AdamW optimizer (24). To further enhance model convergence during training, we employed the CosineAnnealingLrUpdater strategy for optimal learning rate adjustments (25). The initial learning rate is set at 1.75×10^{-3} , with a minimum of 1.75×10^{-5} imposed during the learning rate adjustment process. We also employed linear preheating for the learning rate. Model training and prediction used the LabelSmoothLoss as the loss function (26).

In this process, the training data for the Vision Transformer model were derived from the annotations of pathologists. Additionally, pathologists screened patches to ensure that the training dataset exclusively includes ROI patches exhibiting typical tumor characteristics.

In addition, to investigate the potential impact of refined annotation on the results, we explored other two methods to filter the patches generated by the Preprocessing Module. This included relying on the pathologist's annotation for filtering (Tumor Only) and only removing the blank background (Whole Tissue). The implementation details for each method are described below:

1. Tumor only: For this method, two highly experienced pathologists annotated ROI regions. When the annotated ROI regions were inconsistent, a third more senior pathologist made the final judgment. When screening the patches, if a patch does not completely exist in the ROI area, we will discard it. Only patches fully contained within the ROI area will be retained.
2. Whole tissue: In this approach, while screening the patch, we first evaluate its overall color threshold. Given the RGB color distribution ranging from 0 to 255, where values closer to 255 represent colors closer to white. Therefore, patches with an overall RGB value greater than 220 were considered blank backgrounds and were discarded. The rest of the patches were retained for further analysis.

Through the above method, we obtained three groups of patches annotated in different ways. These three groups of patches were input into the final MGMT Predict Module to obtain diverse results and to compare the impact of these three annotation methods on the results.

2.5 MGMT promoter status prediction module

The ROI patches output by the second module were used as the input of this module to predict the final MGMT promoter methylation

status. The processing flow of this module involved considering each WSI as a package, which is then sliced into numerous ROI patches. Each ROI patch is treated as an instance, with the package's label being assigned to each patch within it. In this part, the main structure of our model adopts the Swin Transformer v2 structure (27). After position encoding of input data using PatchEmbed, the four stages are passed. Except for the first stage, each stage consists of a PatchMerging module and several Swinblock V2 modules. Then, the features are processed by the linear normalization layer and global average pooling layer and then predicted and classified by a fully connected layer. Each SwinBlockV2 module consists mainly of multi-head attention mechanisms with feedforward neural networks and linear normalization layers. In the four stages, the number of SwinBlockV2 is [2, 2, 18, 2], respectively, and the number of attention heads for each SwinBlockV2 is [3, 6, 12, 24], respectively. In this way, for each WSI, we could get the prediction probability of several patches and perform a weighted average of all the prediction results to obtain the final patient-level prediction. This ensured that each patch's predictions contributed appropriately to the patient-level result. The weighting formula was as follows:

$$\text{Weight}_i = \frac{|\text{pred_score}_i - 0.5|}{\sum_{i=1}^n |\text{pred_score}_i - 0.5|}$$

Weight_i represents the weight that the prediction result of the i -th patch contributes to the final weighted average, and pred_score_i signifies the prediction result of the i patch.

For patch-level results, weights were assigned based on the degree of deviation of their prediction probability from 0.5. This strategic weighting reduced the influence of results with prediction probabilities near 0.5 on the final average, thereby strengthening the reliability of model predictions.

2.6 Statistical analysis

We used SPSS 26 to analyze the differences in clinical baseline data of patients using a t -test and chi-square test. Python (3.7.9) was used for data processing and analysis and overall model construction. We employed the PyTorch framework to construct the neural network, and the hardware platform for model training was the NVIDIA A40 GPU with 48G of memory.

3 Results

3.1 Clinical baseline characteristics

According to the inclusion and exclusion criteria, we included 29 WSIs from 29 patients in the TianTan cohort and 632 WSIs from 199 patients in the TCGA cohort. A total of 57 patients in the TCGA database had only one WSI based on the previously described criteria, so 57 WSIs from these 57 patients were used as an independent test set for the TCGA cohort. The remaining 575 WSIs from 142 patients were used as the training set. According to the fifth edition of the WHO Classification of Tumors of the Central Nervous System (WHO CNS5) (5), we excluded patients with IDH mutation from the test set.

Therefore, 25 patients with corresponding 25 WSIs were finally included in the TianTan cohort.

There were no statistically significant differences between the two cohorts regarding age, sex, and MGMT promoter methylation status. Basic clinical characteristics are presented in Table 1.

3.2 Model training and evaluation results

Our primary evaluation metrics included accuracy and the area under the curve (AUC) value of the receiver operating characteristic (ROC) curve. These metrics were used to assess the performance of the models in both the glioblastoma ROI patches prediction module and the MGMT promoter status prediction module.

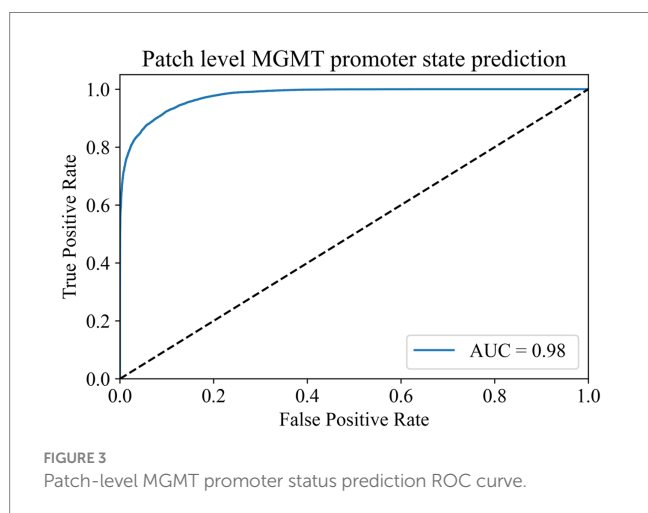
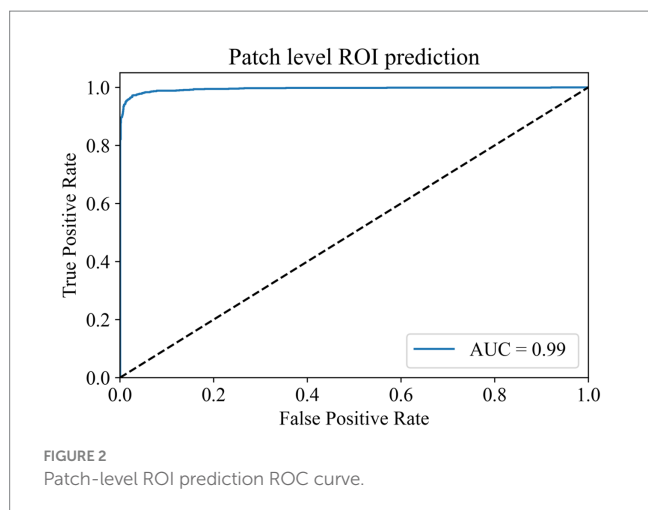
In the ROI patches recognition module, the training dataset comes from the 100 WSIs in the training set of the TCGA cohort described earlier. After segmentation and color normalization, 14,000 patches with typical characteristics were selected under the guidance of professional pathologists. There were 7,000 patches each in the ROI area and non-ROI area. The dataset was divided into training, validation, and test sets in a 6:2:2 ratio. During training, a batch size of 224 and 100 epochs were used for the parameter settings of the model. The model reached the highest accuracy of 97.88% in the validation set at the 80th epoch, and the model parameters at this time were saved for testing in the test set. The final accuracy of the ROI patches prediction module in the test set is 97.03%, and the AUC value is 0.9940 (Figure 2). The results show that the module has a superior ability to identify ROI patches, which can meet the screening needs of the subsequent modules.

In the MGMT promoter status prediction module, we first used the Swin Transformer V2 model for training. The ROI patches output by the ROI patches prediction module was used as the training data of this module. Whether the patient's MGMT promoter was methylated or not was taken as the label of the patches cut out from that patient. Therefore, we finally obtained 59,945 MGMT promoter methylated patches and 97,431 MGMT promoter unmethylated patches.

To address the issue of category imbalance, we employed random rotation and mirroring to expand both categories to 100,000, with a total of 200,000 patches for model training. The dataset was divided into training, validation, and test sets in a 6:2:2 ratio. We set the batch size to 148 for training parameters and the number of epochs to 500.

TABLE 1 Clinical and demographic characteristics of patients.

	TianTan ($n = 25$)	TCGA ($n = 199$)	p value
Age, mean \pm SD (years)	58.40 \pm 12.50	57.24 \pm 14.39	0.365
Sex			0.822
Male	15 (60%)	121 (61%)	
Female	10 (40%)	78 (39%)	
MGMT promoter status			0.167
Methylated	13 (52%)	75 (38%)	
Unmethylated	12 (48%)	124 (62%)	



Remarkably, during the 493rd epoch, the model achieved its highest accuracy of 91.38% in the validation set. We saved the model weights at this time for testing on the test set. In the end, the model exhibited an accuracy of 91.12% in the test set, with an AUC of 0.98 (Figure 3). At the same time, we compared it with the current commonly used Resnet50 model, and the comparison results are shown in Table 2. These results show the ability of the model to accurately distinguish whether the MGMT promoter is methylated at the patch level.

To enhance the model's interpretability, we utilized a class gradient activation heatmap to visualize the model's recognition of patches. In this visualization, regions with a redder color indicate a higher probability that those areas are methylated. As depicted in Figure 4A, represents a WSI with MGMT promoter methylation, while Figure 4B represents a WSI without MGMT promoter methylation. Each image's upper left corner displays the corresponding patch's gradient class activation heatmap. The results demonstrate that the models correctly identify their respective patches.

Hence, we aimed to extrapolate the patch-level results to the patient level. The patches obtained by three different methods were input into the MGMT promoter status prediction module for testing. According to the ROC curve of previous patch-level evaluation, we calculated that the best cutoff value for judging whether MGMT promoter methylation exists at the patch level was 0.52; that is, when

TABLE 2 Comparison between the Swin transformer V2 and Resnet50.

Model	Accuracy	Precision	Recall	F1 score	AUC
Swin_trans_V2	0.91	0.92	0.91	0.91	0.98
Resnet50	0.72	0.74	0.70	0.72	0.80

the probability of methylation of a patch MGMT promoter judged by the model was greater than 0.52, we think it has MGMT promoter methylation. We use this cutoff value for patient-level model evaluation. To obtain patient-level results, we aggregated the patch-level outcomes using the probability-weighted average method as previously described. The results are presented in Table 3.

In both TCGA and TianTan test sets, the approach combining the ROI patches prediction module with the MGMT promoter status prediction module demonstrates the highest AUC values. Moreover, the approach combining the ROI patches prediction Module with the MGMT promoter status in the Tiantan cohort and the TCGA cohort T achieved the same highest balance accuracy. The ROC curves for these three methods in the test set are depicted in Figure 5. These comprehensive results demonstrate that the overall process we designed can effectively determine the presence or absence of MGMT promoter methylation in patients through the H&E-stained histopathological slides.

4 Discussion

In this study, we introduced a weakly supervised deep learning model based on H&E-stained histopathological slides to determine the MGMT promoter methylation status of glioblastoma. The methylation status of MGMT gene promoter has a significant impact on histological typing and diagnosis as well as predicting patient survival and response to treatment. Although there is much evidence that it plays an important role in prognosis and treatment, its routine implementation in clinical practice has been challenging (3, 28). The best method and optimal cutoff definitions for MGMT status determination remain under debate. Thus, there is an urgent need to explore a simpler and faster way to obtain the status of the MGMT promoter.

Current research on AI-assisted prediction of MGMT promoter methylation in glioma mainly focuses on radiomics. However, the efficacy of MRI imaging measurement indicators to predict the methylation status of the MGMT promoter remains controversial. For instance, Choi et al. (29) reported that there was no significant difference between the apparent diffusion coefficient (ADC) values and MGMT promoter methylation status. This finding is consistent with the latest research results of Ma et al. (30), which suggest that specific MRI measures such as ADC may not reliably distinguish from MGMT promoter methylation status. In contrast, a study conducted by Han et al. in 2018 reported different findings. They observed a significant increase in ADC values in preoperative MRI scans of primary glioblastoma with MGMT promoter methylation. Furthermore, they utilized this difference as a classification predictor, achieving an AUC value of 0.86 (31). Indeed, machine learning and deep learning approaches have gained significant attention in recent years for predicting MGMT promoter methylation status in

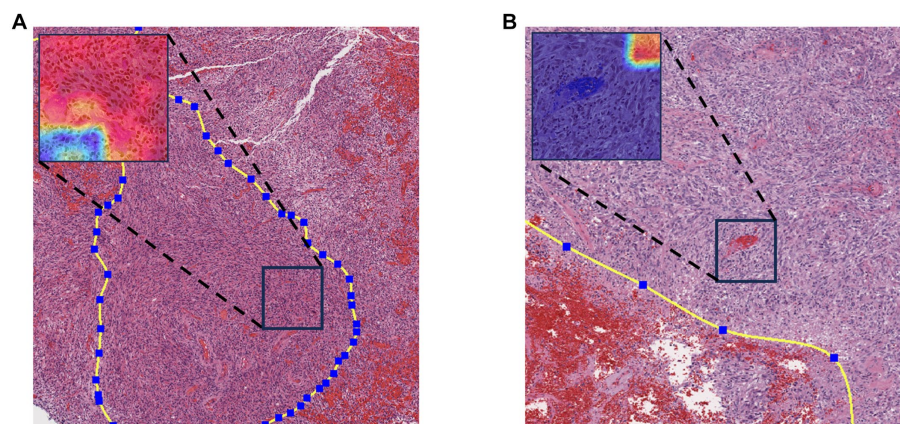


FIGURE 4
Gradient-weighted class activation mapping. Redder colors in the graph indicate a greater likelihood of methylation. **(A)** WSI with MGMT promoter methylation. **(B)** MGMT promoter unmethylated WSI.

TABLE 3 Patient-level test results.

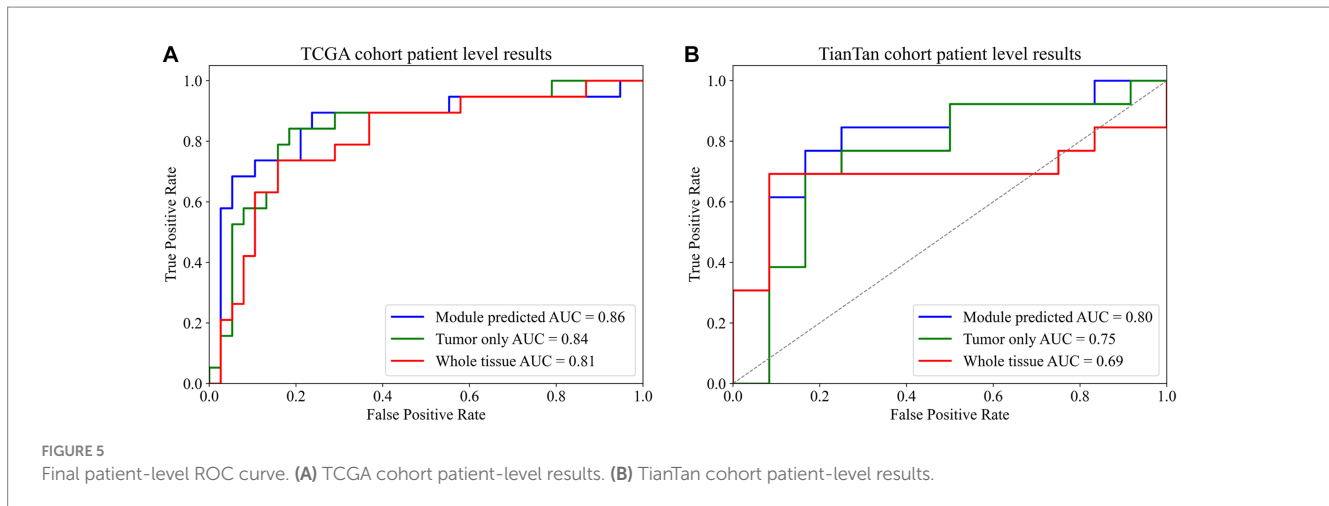
	Tumor only	Whole tissue	Module predict
TCGA Test (n=57)			
Balance_Accuracy	0.79	0.75	0.79
AUC	0.81	0.84	0.86
Recall	0.84	0.74	0.74
Specificity	0.74	0.76	0.84
Precision	0.62	0.61	0.70
F1 score	0.71	0.67	0.72
TianTan Test (n=25)			
Balance_Accuracy	0.76	0.69	0.76
AUC	0.75	0.69	0.83
Recall	0.69	0.46	0.69
Specificity	0.83	0.92	0.83
Precision	0.82	0.86	0.82
F1 score	0.75	0.60	0.75

glioblastoma. Pasquini et al. (32) employed a machine learning approach to forecast the methylation status of the MGMT promoter in glioblastoma patients using MRI, achieving an accuracy of 70.8% and an AUC value of 0.688. Chen et al. (33) extracted MRI-related features from diffuse gliomas and utilized a ResNet network for training. The model achieved an impressive AUC value of 0.90 and an accuracy of 91.0% in the test set (33). Therefore, we believe that more innovative and effective methods and approaches for MGMT promoter methylation prediction models are imperative. Our investigation employed two neural networks with transformer-like architectures to predict the MGMT promoter methylation status in glioblastoma. Our findings align with Li's study and prove the feasibility of utilizing histopathological images for MGMT promoter methylation status prediction (13). At the same time, we noticed that some recent studies used HE-stained pathological sections to predict MGMT promoter methylation. For example, Mili et al. (34) proposed

a MobileNetV2 model that introduced a spatial attention correlation mechanism to predict grade II and III gliomas, but their results only stayed at the patch level. No predictions were made for the patients. In addition, the study by Krebs et al. (35) also proposed the method of comparing self-supervised learning (SSL) and dual-stream multi-instance learning (DSMIL) to predict the methylation state of MGMT promoter by using HE-stained slices and achieved good results. Our research further suggests that adding an automatic data screening module in the process of model training and prediction may achieve better results.

In future research, WSI, as a different kind of data from MRI, may exhibit more accurate predictions of the molecular status of brain tumors. The large amount of information carried by WSI is more suitable for the use of deep learning models. Furthermore, the inherently two-dimensional characteristics of WSI offer more opportunities for widely interdisciplinary collaboration with the latest image classification algorithms in the computer field, such as transformer-based architectures. Additionally, the proposal of advanced deep learning techniques such as multiple instances learning and graph neural networks has further promoted the development of deep learning based on WSI (36, 37). Especially when WSI and MRI are combined, it may produce remarkable prediction effects in building brain tumor classification prediction models. However, it is also worth noting that the results obtained from simple MRI data and pathological section images are not enough to be directly applied to clinical applications (38). In the future, we must incorporate more clinical and even genetic characteristics of patients to build a multi-dimensional predictive model. To improve the reliability of our results. In addition, the disadvantage of the current deep learning model is that it is easy to become a black box. We do not know what features in the pathological sections make the model final, which is a problem for clinicians. To solve the problem of explainability of models, the development of explainable computational pathology combined with explainable AI in recent years may be an important direction for future research (39).

Indeed, this study has some limitations, including the relatively limited samples from Beijing TianTan Hospital and the utilization of data from a single source. In the future, we will try to incorporate more sources of data for analysis and combine MRI data to build



multimodal models to further enhance the accuracy and robustness of predictions.

5 Conclusion

Our weakly supervised prediction model based on H&E-stained histopathological slides proves to be an effective approach for predicting the MGMT promoter methylation status in glioblastoma. Additionally, accurate annotation of the tumor region can significantly enhance the predictive performance of the model.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Institutional Review Board of Beijing Tiantan Hospital, Capital Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin because this is a retrospective study and only the patient's previous pathological image information was used.

References

- Ostrom QT, Price M, Neff C, Cioffi G, Waite KA, Kruchko C, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2015-2019. *Neuro-Oncology*. (2022) 24:v1-v95. doi: 10.1093/neuonc/noac202
- Weller M, van den Bent M, Preusser M, Le Rhun E, Tonn JC, Minniti G, et al. EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood. *Nat Rev Clin Oncol*. (2021) 18:170-86. doi: 10.1038/s41571-020-00447-z
- Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med*. (2005) 352:997-1003. doi: 10.1056/NEJMoa043331
- Butler M, Pongor L, Su YT, Xi L, Raffeld M, Quezado M, et al. MGMT status as a clinical biomarker in glioblastoma. *Trends Cancer*. (2020) 6:380-91. doi: 10.1016/j.trecan.2020.02.010
- Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO classification of tumors of the central nervous system: a summary. *Neuro-Oncology*. (2021) 23:1231-51. doi: 10.1093/neuonc/noab106
- Elmore JG, Longton GM, Carney PA, Geller BM, Onega T, Tosteson AN, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *J Am Med Assoc*. (2015) 313:1122-32. doi: 10.1001/jama.2015.1405

Author contributions

YH: Conceptualization, Formal analysis, Methodology, Project administration, Writing – original draft. LD: Data curation, Investigation, Writing – review & editing. GD: Resources, Writing – review & editing. FC: Investigation, Writing – review & editing. WL: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by Clinical Major Specialty Projects of Beijing (2-1-2-038).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

7. Niazi M, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol.* (2019) 20:e253–61. doi: 10.1016/S1470-2045(19)30154-8
8. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med.* (2018) 24:1559–67. doi: 10.1038/s41591-018-0177-5
9. Jin L, Shi F, Chun Q, Chen H, Ma Y, Wu S, et al. Artificial intelligence neuropathologist for glioma classification using deep learning on hematoxylin and eosin-stained slide images and molecular markers. *Neuro-Oncology.* (2021) 23:44–52. doi: 10.1093/neuonc/noaa163
10. Luo C, Yang J, Liu Z, Jing D. Predicting the recurrence and overall survival of patients with glioma based on histopathological images using deep learning. *Front Neurol.* (2023) 14:1100933. doi: 10.3389/fneur.2023.1100933
11. Jiang S, Zanazzi GJ, Hassanpour S. Predicting prognosis and IDH mutation status for patients with lower-grade gliomas using whole slide images. *Sci Rep.* (2021) 11:16849. doi: 10.1038/s41598-021-95948-x
12. Wang D, Liu C, Wang X, Liu X, Lan C, Zhao P, et al. Automated machine-learning framework integrating histopathological and radiological information for predicting IDH1 mutation status in glioma. *Front Bioinform.* (2021) 1:718697. doi: 10.3389/fbiof.2021.718697
13. Li Z, Cong Y, Chen X, Qi J, Sun J, Yan T, et al. Vision transformer-based weakly supervised histopathological image analysis of primary brain tumors. *iScience.* (2023) 26:105872. doi: 10.1016/j.isci.2022.105872
14. Kim M, Yoo J, Chang JH, Kim SH. Association of MGMT gene promoter methylation with clinicopathological parameters in patients with wild-type IDH glioblastoma. *Anticancer Res.* (2022) 42:335–41. doi: 10.21873/anticancer.15490
15. Babikir H, Wang L, Shamardani K, Catalan F, Sudhir S, Aghi MK, et al. ATRX regulates glial identity and the tumor microenvironment in IDH-mutant glioma. *Genome Biol.* (2021) 22:311. doi: 10.1186/s13059-021-02535-4
16. Jean-Quartier C, Jeanquartier F, Holzinger A. Open data for differential network analysis in glioma. *Int J Mol Sci.* (2020) 21:547. doi: 10.3390/ijms21020547
17. Brennan CW, Verhaak RG, McKenna A, Campos B, Noushmehr H, Salama SR, et al. The somatic genomic landscape of glioblastoma. *Cell.* (2013) 155:462–77. doi: 10.1016/j.cell.2013.09.034
18. Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell.* (2010) 17:98–110. doi: 10.1016/j.ccr.2009.12.020
19. Yogananda C, Shah BR, Nalawade SS, Murugesan GK, Yu FF, Pinho MC, et al. MRI-based deep-learning method for determining glioma MGMT promoter methylation status. *Am J Neuroradiol.* (2021) 42:845–52. doi: 10.3174/ajnr.A7029
20. MMPreTrain Contributors (2023). OpenMMLab's pre-training toolbox and benchmark (version 0.15.0) [Computer software]. Available at: <https://github.com/open-mmlab/mmpretrain>
21. Salvi M, Acharya UR, Molinari F, Meiburger KM. The impact of pre- and post-image processing techniques on deep learning frameworks: a comprehensive review for digital pathology image analysis. *Comput Biol Med.* (2021) 128:104129. doi: 10.1016/j.combiomed.2020.104129
22. Vahadane A, Peng T, Sethi A, Albarqouni S, Wang L, Baust M, et al. Structure-preserving color normalization and sparse stain separation for histological images. *IEEE Trans Med Imaging.* (2016) 35:1962–71. doi: 10.1109/TMI.2016.2529665
23. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., et al. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv [Preprint]. doi: 10.48550/arXiv.2010.11929
24. Loshchilov I., Hutter F. (2017). Decoupled weight decay regularization. arXiv [Preprint]. doi: 10.48550/arXiv.1711.05101
25. Loshchilov I., Hutter F. (2016). SGDR: Stochastic gradient descent with warm restarts. arXiv [Preprint]. doi: 10.48550/arXiv.1608.03983
26. Müller R., Kornblith S., Hinton G. (2019). When does label smoothing help? arXiv [Preprint]. doi: 10.48550/arXiv.1906.02629
27. Liu Z, Hu H, Lin Y, Yao Z, Xie Z, Wei Y, et al. (2022). "Swin transformer V2: scaling up capacity and resolution" in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), LA, USA: New Orleans. 11999–12009.
28. Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJ, Janzer RC, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* (2009) 10:459–66. doi: 10.1016/S1470-2045(09)70025-7
29. Choi YS, Ahn SS, Kim DW, Chang JH, Kang SG, Kim EH, et al. Incremental prognostic value of ADC histogram analysis over MGMT promoter methylation status in patients with glioblastoma. *Radiology.* (2016) 281:175–84. doi: 10.1148/radiol.2016151913
30. Ma X, Cheng K, Cheng G, Li C, Lyu J, Lan Y, et al. Apparent diffusion coefficient as imaging biomarker for identifying IDH mutation, 1p19q Codeletion, and MGMT promoter methylation status in patients with glioma. *J Magn Reson Imaging.* (2023) 58:732–8. doi: 10.1002/jmri.28589
31. Han Y, Yan LF, Wang XB, Sun YZ, Zhang X, Liu ZC, 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. doi: 10.1186/s12885-018-4114-2
32. Pasquini L, Napolitano A, Lucignani M, Tagliente E, Dellepiane F, Rossi-Espagnet MC, et al. AI and high-grade glioma for diagnosis and outcome prediction: do all machine learning models perform equally well? *Front Oncol.* (2021) 11:601425. doi: 10.3389/fonc.2021.601425
33. Chen S, Xu Y, Ye M, Li Y, Sun Y, Liang J, et al. Predicting MGMT promoter methylation in diffuse gliomas using deep learning with Radiomics. *J Clin Med.* (2022) 11:3445. doi: 10.3390/jcm11123445
34. Mili M., Abdallah A. B., Otero J. J., Kerkeni A., Bedoui M. H. (2023). "Revolutionizing brain cancer diagnosis: automated prediction of MGMT methylation status using histological images" in 2023 International Conference on Cyberworlds (CW). Sousse, Tunisia, 149–156.
35. Krebs O, Agarwal S, Tiwari P (2023). "Self-supervised deep learning to predict molecular markers from routine histopathology slides for high-grade glioma tumors" in Proceedings Volume 12471, Medical Imaging 2023: Digital and Computational Pathology; 1247102.
36. Ilse M., Tomczak J. M., Welling M. (2018). Attention-based deep multiple instance learning. arXiv [Preprint]. doi: 10.48550/arXiv.1802.04712
37. Shaban M, Awan R, Fraz MM, Azam A, Tsang YW, Snead D, et al. Context-aware convolutional neural network for grading of colorectal Cancer histology images. *IEEE Trans Med Imaging.* (2020) 39:2395–405. doi: 10.1109/TMI.2020.2971006
38. Holzinger A, Haibe-Kains B, Jurisica I. Why imaging data alone is not enough: AI-based integration of imaging, omics, and clinical data. *Eur J Nucl Med Mol Imaging.* (2019) 46:2722–30. doi: 10.1007/s00259-019-04382-9
39. Plass M, Kargl M, Kiehl TR, Regitnig P, Geissler C, Evans T, et al. Explainability and causability in digital pathology. *J Pathol Clin Res.* (2023) 9:251–60. doi: 10.1002/cjp2.322