Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 3, 2026

MULTIMODAL DEEP LEARNING FOR BRAIN TUMOR CLASSIFICATION: INTEGRATING CLINICAL, HISTOPATHOLOGICAL, AND DEMOGRAPHIC DATA WITHIN A UNIFIED FRAMEWORK

[1]Hafiz Inam ul Haq[a] , [2]Sharafat Ali , [3]Muhammad Ilyas , [4]Husnain Ali , [5]Fayyaz Rasheed

[1]Department of CS and IT Superior Univeristy Lahore 54000, Pakistan

[2,3,4,5]Department of software Engineering Superior Univeristy Lahore 54000, Pakistan

sardarhafizinam@gmail.com, sharafat.ali@superior.edu.pk, milyas0060@gmail.com, husnain3004@gmail.com, fayyazrasheed3@gmail.com

DOI:- https://doi.org/10.5281/zenodo.21333902

Keywords-
Multimodal Deep Learning, Brain Tumor Classification, Integrating, Clinical, Histopathological, Demographic, Unified Framework

Article History
Received: 04 Feb 2026
Accepted: 25 Feb 2026
Published: 16 March 2026

Corresponding Author: *
Hafiz Inam ul Haq

Abstract

Brain Tumor classification is still difficult because the decisions are made based on sporadic and heterogeneous information which is often analyzed separately. The study introduces a single, integrated multimodal deep-learning framework that combines histopathological images and clinical and demographic information to create more accurate, robust, and interpretable brain tumor classification. The proposed method uses a deep visual encoder to capture the morphological characteristics of tumors in the histopathological image and a specifically designed neural network to capture other patterns from the structured clinical and demographic variables. These modality-specific representations are fused together using an attention-based fusion mechanism, in order to obtain a global patient-level representation for tumor classification. The framework is compared to unimodal and conventional machine-learning baselines using class-sensitive discrimination and calibration metrics, and class-sensitive robustness metrics; and the contribution of each data modality is quantified through ablation studies. The use of EAI techniques to find influential tissue regions and patient characteristics is included, and subgroup analyses are performed to evaluate the demographic fairness and generalizability. The proposed framework combines pathological morphology with patient clinical context to build a more comprehensive and clinically relevant decision-support system for brain tumor diagnosis and set the stage for a scalable precision neuro-oncology framework.

1. Introduction

Brain and central nervous system tumors are a heterogeneous group of primary and metastatic tumors with great variation in cell of origin, molecular characteristics, site of origin, aggressiveness, therapeutic sensitivity, and clinical prognosis. They account for a lesser percentage of the global cancer burden than many systemic cancers but are no less debilitating in their impact, as they impact neurological functions vital to cognition, behavior, sensation, movement and physiological regulation. Brain and central nervous system cancers are a significant clinical and public-health burden, with an estimated 321,731 new cases and 248,500 deaths in 2022 [1].

Proper classification is key to the management of neuro-oncology as the prognosis, surgical planning, molecular investigation, therapeutic options, follow-up and eligibility for clinical trials are all dependent on tumor type and subtype. However, diagnosis of brain tumor is still challenging, as brain tumor entities can have similar histopathological features, and there can be intratumoral heterogeneity between different areas of the same tumor. Limited biopsy material, the staining variation, tissue artefacts, necrosis, hemorrhage, inflammatory infiltration and differences in how the tissue is prepared by various institutions may also influence the diagnostic interpretation. All of these factors suggest that brain-tumors classification is not a simple visual-recognition problem.

The contemporary diagnostic paradigm has consequently evolved beyond morphology alone. The fifth edition of the World Health Organization Classification of Tumors of the Central Nervous System adopted an integrated model in which the molecular alterations are taken into account alongside the histopathological features of the tumor and other characteristics that define the tumor. IDH-mutant astrocytoma, IDH-mutant and 1p/19q-codeleted oligodendroglia and IDH-wild type glioblastoma, for instance, are now the major categories of the adult-type diffuse gliomas. This is because it is recognized that tumors with similar morphology can be different diseases with different therapeutic and prognostic interpretations [2,3].

Histopathological examination continues to play a major role in the diagnosis of biopsied and resected tumors. Tissue sections stained with hematoxylin and eosin provide information on cell shape, abnormal nuclear features, high rate of cell division, tumors invasion, cell death, blood vessel growth, abnormal inflammatory reactions, and overall tissue structure. The digitization of the glass slides to whole slide images has opened up this information for computational analysis. Deep-learning models can extract intricate morphological patterns over large

tissue areas and have been used to detect tumors, grade them, classify their subtypes, predict their molecular status, and model outcomes. Artificial intelligence can, for instance, be used to obtain clinically relevant data from the morphology of tissue, as is the case in the example of Holon et al., who were able to make an intraoperative diagnosis of brain cancer in less than 150 seconds using stimulated Raman histology and a deep convolutional network [4]. However, there are significant technical challenges for the whole-slide analysis. The image consists of billions of pixels, which can't be processed at normal resolution. They are therefore needed to be divided into smaller patches, or encoded by a compressed feature embedding. Many extracted patches, however, may include non-diagnostically informative tumors tissue, background tissue, blood, necrosis, an empty background, or artefacts, as well as normal tissue. Weakly-supervised multiple instance learning tries to solve this by considering each slide as a bag of image regions, while learning what image region is most informative for the patient-level diagnosis. However, the performance of the models may remain affected by the selection of the patches, the magnification, the staining differences, the scanner properties, the quality of the tissue and its institutional origin.

However, histopathology may sometimes be inadequate, e.g. tumors entities can have similar morphology or samples may be small and unrepresentative. Complementary clinical & demographic information may be provided. Potential variables of interest are age, sex, presenting symptoms, tumors location, laterality, neurological findings, functional status, previous malignancy, and selected laboratory findings. The prevalence and nature of central nervous system tumors are age dependent, and age is therefore particularly informative. Demographic variables should be added as well, though, because otherwise a model might learn population imbalance or referral patterns or disparities in care, rather than clinically meaningful relations.

Both patient information and histopathological data are multimodal and need specialized multimodal learning to be combined. The dimensionality, distribution, and amusingness, noise, and predictive power of image data are very different from those of tabular variables. Early fusion refers to the process of fusing modalities before learning a large amount of representations, while intermediate fusion fuses features learned by modality specific encoders. Late fusion is the fusion of predictions from independent models. Feature level, gated and attention based approaches are especially promising as they enable each encoder to specialize in its own data type while learning interactions between the patient context and tumors morphology. Pathetic Fusion demonstrated the possibility of integrating information from the histopathological image and genome, and recent

studies on brain tumors have demonstrated that multimodal models can be superior to unimodal models in certain prognostic tasks [5, 6].

There are still some key research gaps, though. Most computational systems used to deal with brain tumors problems still use a single modality and many studies referred to as multimodal deals with several imaging sequences and not with multiple genuinely different histological, clinical and demographic data. Patient-level integration of whole-slide pathology with structured clinical variables and demographic data is relatively under-explored. Past research also varies in the methodology used for creating fusion, the method used to create cohorts, how labels are defined, how missing data is handled, how to manage the class imbalance, and in validation designs. Simple feature concatenation is often used without considering other (more complex) gated or attention-based or decision level approaches.

There also are methodological weaknesses that further restrict clinical reliability. A data splitting issue could be patch-level instead of patient-level, resulting in the same patient's tissue being split into the training and testing sets and thereby causing information leakage and over-optimistic results. Some variables collected after diagnosis can also be a spoiler of the target label. The overall accuracy can mask under-representation of tumor types and the absence of external validation and subgroup analysis can overestimate generalizability. However, computational pathology models can also exhibit differential performance for demographic groups, where the overall performance looks good, and so fairness analysis is essential even if the overall performance looks good [7].

In this context, the present study suggests a common and integrated multimodal deep learning framework in which the histopathological whole-slide images are integrated with clinical and demographic patient information. The framework will include a histopathological encoder for tissue segmentation, patch extraction and morphological representation, a structured-data encoder for cleaned and normalized patient variables, and a fusion module for integrating modality-specific representations using techniques such as concatenation, gating or attention, or similar approaches. While it is possible that modern pathology foundation models can offer transferable image representations and less reliance on the need to annotate tasks extensively, the ability of these models to be suitable across different domains should be tested empirically.

The main objective is to evaluate if joint modelling of tumor morphology and patient context can be used to gain more accurate, robust, easily interpretable and clinically relevant classification than unimodal analysis. The study will be based on the comparison of histopathology only, clinical only, demographic only, conventional machine-learning, and alternative

multimodal-fusion models. The contribution of each modality will be quantified through ablation experiments, and attention maps and feature-attribution methods will be used to identify influential regions of tissue and structured variables. In addition to accuracy, performance will be assessed by macro-F1 score, balanced accuracy, sensitivity, specificity, ROC-AUC, precision–recall AUC, confusion matrices, and calibration.

Clinically, it is important that the study aims to mimic more closely the integrative reasoning that is required in neuro-oncology than image-only classification. Methodologically, it will focus on the concepts of patient level partitioning, missing-modality evaluation, external validation, interpretability and demographic subgroup assessment. Bio-medically, it could show complementary relationships between the nature and morphology of the tumor and the patients' characteristics. It will, however, be designed as a decision support system and will not be intended to provide a substitute for neuropathological interpretation, molecular testing, multi-disciplinary assessment, or clinical judgement. The extent to which it will cover will depend on quality of the datasets, representation of classes, reliability of labels and accessibility of matched clinical, demographic and histopathological data.

2. Related Work

2.1 Evolution of Brain Tumor Classification

Brain tumor classification has progressed from a predominantly morphology-based approach to an integrated diagnostic framework that combines histopathological, molecular, anatomical, and clinical evidence. Traditional classification relied primarily on cellular morphology, mitotic activity, necrosis, microvascular proliferation, and the presumed tissue of origin; however, morphologically similar tumors may exhibit distinct molecular profiles, biological behaviors, prognoses, and therapeutic responses [9, 10]. The fifth edition of the World Health Organization Classification of Tumors of the Central Nervous System, published in 2021, formalized the concept of integrated diagnosis by incorporating molecular markers into the definitions of numerous tumor entities [9, 10]. Adult-type diffuse gliomas, for example, are now principally classified as astrocytoma, IDH-mutant; oligodendroglia, IDH-mutant and codeleted; and glioblastoma, IDH-wild type [10]. This evolution has important implications for computational research because artificial-intelligence models trained using outdated, ambiguous, or exclusively morphology-based labels may not correspond to contemporary diagnostic standards. Consequently, computational systems should employ clinically valid and molecularly informed reference labels while recognizing that accurate brain tumor classification frequently requires the integration of heterogeneous evidence rather than reliance on a single diagnostic modality [9,10].

2.2 Deep Learning in Histopathological Brain Tumor Analysis

The digitization of histopathological slides has enabled deep-learning algorithms to identify diagnostically relevant patterns in brain tumor tissues. Convolutional neural networks, vision transformers, and transfer-learning approaches have been applied to tumor detection, grading, subtype classification, molecular-marker prediction, and prognostic assessment [11, 12]. Whole-slide images provide detailed representations of cellular morphology, tissue architecture, microvascular proliferation, necrosis, and tumor infiltration; however, their extremely high resolution prevents direct processing by conventional neural networks. Consequently, whole-slide images are commonly divided into smaller patches, from which deep feature representations are extracted and aggregated to generate slide-level or patient-level predictions [11]. Attention-based multiple-instance learning has become particularly influential because it enables models to identify diagnostically informative tissue regions using only slide-level or patient-level labels. Lu et al. introduced the clustering-constrained attention multiple-instance learning framework, which improved weakly supervised whole-slide image classification while generating attention maps that highlighted influential tissue regions [11]. Similarly, Holon et al. demonstrated that deep neural networks applied to stimulate Raman histology could provide near-real-time intraoperative brain tumor diagnoses with performance comparable to conventional pathological assessment [12]. Nevertheless, histopathology-only systems remain vulnerable to staining variability, tissue artifacts, sampling bias, intra tumor heterogeneity, and morphological overlap among distinct tumor entities.

2.3 Clinical and Demographic Data in Tumor Classification

Clinical and demographic information provides essential context that may complement histopathological morphology. Patient age, sex, presenting symptoms, neurological findings, tumor location, comorbidities, previous malignancy, laboratory findings, and functional status can contribute to differential diagnosis because the prevalence and biological characteristics of brain tumors vary across populations. Certain pediatric tumors rarely occur in older adults, while glioblastoma and meningioma are predominantly diagnosed in adult and elderly populations. Similarly, several tumor categories demonstrate sex-related differences in incidence. Structured clinical data have been analyzed using logistic regression, random forests, gradient-boosting algorithms, multilayer perceptrons, and transformer-based tabular models. Nakagaki et al. integrated clinical information with histopathological

images to predict IDH1 mutation status, demonstrating that patient-level variables can provide complementary information beyond tissue morphology alone [13]. However, demographic characteristics must be incorporated cautiously because models may exploit healthcare inequalities, referral patterns, or institutional imbalances rather than clinically meaningful associations. Vaidya et al. demonstrated that computational pathology models may produce unequal diagnostic error rates across demographic groups, highlighting the necessity of subgroup analysis and fairness evaluation [14]. Clinical variables generated after definitive diagnosis must also be excluded to prevent target leakage and unrealistically high predictive performance.

2.4 Multimodal Deep Learning and Fusion Strategies

Multimodal deep learning provides a methodological framework for combining heterogeneous data sources into a unified patient representation. In brain tumor research, histopathological images capture microscopic morphology, while clinical and demographic data describe disease presentation and patient context. Multimodal fusion may occur through early, intermediate, or late integration. Early fusion combines variables before representation learning but is often unsuitable when inputs differ substantially in dimensionality. Intermediate fusion employs separate encoders for each modality and integrates their learned representations through concatenation, attention, gating, bilinear interactions, or transformer-based mechanisms. Late fusion combines predictions produced independently by modality-specific models and is relatively robust when some inputs are unavailable, although it may fail to capture complex cross-modal interactions. Chen et al. proposed Pathomic Fusion, which integrated histopathological and genomic representations through gated and tensor-based interactions and demonstrated the value of interaction-aware multimodal learning in cancer diagnosis and prognosis [15]. Steyaert et al. similarly reported that combining histopathological and molecular data improved prognostic prediction in adult and pediatric brain tumors compared with unimodal approaches [16]. Despite these advances, relatively few studies have directly integrated histopathological, clinical, and demographic information for multiclass brain tumor classification.

2.5 Research Gap and Direction of the Present Study

The existing literature demonstrates that deep-learning models can extract valuable diagnostic information from histopathological images and that multimodal systems may outperform unimodal approaches in selected neuro-oncological tasks. Nevertheless, several important limitations remain unresolved. Most brain tumor studies analyze histopathology, radiology, or molecular information independently, while genuinely heterogeneous integration involving clinical, histopathological, and demographic data remains

comparatively uncommon. Existing multimodal studies frequently concentrate on survival prediction or individual molecular markers rather than comprehensive tumor classification. Furthermore, many investigations employ small retrospective cohorts, inadequately manage class imbalance, omit calibration analysis, and provide limited external validation. Patient-level data leakage, incomplete clinical records, missing modalities, demographic bias, and insufficient interpretability also undermine the reliability of reported performance. The present study addresses these gaps by proposing a unified patient-level multimodal framework that combines histopathological representations with structured clinical and demographic information. It will compare unimodal and multimodal models, evaluate alternative fusion mechanisms, quantify the contribution of each modality through ablation analysis, examine subgroup performance, and incorporate explainable artificial-intelligence techniques to identify influential tissue regions and patient characteristics. This approach is intended to provide a more comprehensive, transparent, and clinically relevant framework for brain tumor classification.

Study	Data used	Method	Main contribution	Research limitation	Study
Hollon et al. (2020)	Raman histology images	CNN	Enabled rapid intraoperative brain-tumor diagnosis.	Did not include clinical or demographic data.	Hollon et al. (2020)
Lu et al. (2021)	Whole-slide images	Attention-based MIL	Improved weakly supervised slide classification.	Not specific to multimodal brain-tumor diagnosis.	Lu et al. (2021)
Chen et al. (2022)	Histology and genomic data	Pathomic Fusion	Demonstrated interaction-based multimodal learning.	Focused on genomic rather than routine clinical data.	Chen et al. (2022)
Steyaert et al. (2023)	Histology and molecular data	Multimodal deep learning	Improved brain-tumor prognostic prediction.	Examined prognosis rather than tumor classification.	Steyaert et al. (2023)

Hoang et al. (2024)	Histology and methylation data	DEPLOY	Predicted molecularly defined CNS tumor classes.	Limited integration of broader clinical variables.	Hoang et al. (2024)	Table 1.
Nakagaki et al. (2024)	Histology and clinical data	MIL and ensemble fusion	Improved IDH1 mutation prediction using combined data.	Restricted to binary mutation prediction.	Nakagaki et al. (2024)	
Vaidya et al. (2024)	Histology and demographic data	Fairness analysis	Identified demographic bias in pathology models.	Did not develop a unified multimodal classifier.	Vaidya et al. (2024)	
Shirae et al. (2025)	Histology and clinical data	Ensemble deep learning	Supported multimodal glioma subtype classification.	Limited to glioma and lacked broad external validation.	Shirae et al. (2025)	

Summary of Relevant Literature

3.1 Dataset

This study will use publicly available multimodal data from the **TCGA-LGG** and **TCGA-GBM** projects available through the National Cancer Institute Genomic Data Commons. These datasets contain digitized hematoxylin-and-eosin-stained histopathological whole-slide images linked with patient-level clinical and demographic information.

Only patients with at least one diagnostic slide of acceptable quality, a clearly defined tumor diagnosis, and sufficient clinical and demographic records will be included. Candidate variables will include age at diagnosis, sex, race, ethnicity, tumor grade, primary diagnosis, histological subtype, anatomical site, laterality, prior malignancy, and relevant prediagnostic clinical characteristics. Slides with severe artifacts, inadequate tissue, excessive background, poor staining, or uncertain patient linkage will be excluded. The principal classification task will distinguish glioblastoma, astrocytoma, and oligodendroglioma,

with labels harmonized according to contemporary central nervous system tumor terminology. All slides and records from the same patient will remain within a single dataset partition to prevent information leakage, and the cohort will be divided at the patient level into training, validation, and test sets. Missing-value imputation, normalization, feature selection, and other preprocessing procedures will be learned exclusively from the training data. The CPTAC-GBM cohort may be used for external validation of glioblastoma-related performance; however, because the selected public datasets mainly represent gliomas, the study should be described as multimodal glioma classification unless additional datasets containing other brain tumor categories are incorporated.

Dataset component	Description
Primary datasets	TCGA-LGG and TCGA-GBM
External validation	CPTAC-GBM
Histopathological data	Diagnostic H&E whole-slide images
Clinical variables	Tumor grade, diagnosis, histological subtype, anatomical site, laterality, and prior malignancy
Demographic variables	Age, sex, race, and ethnicity
Proposed classes	Glioblastoma, astrocytoma, and oligodendroglioma
Inclusion criteria	Valid diagnostic slide, confirmed diagnosis, and sufficient linked patient data
Exclusion criteria	Poor-quality slides, inadequate tissue, uncertain labels, or failed patient-level linkage
Prediction unit	Individual patient
Data partitioning	Patient-level training, validation, and test sets
Key limitation	The selected datasets primarily represent gliomas rather than all brain-tumor categories

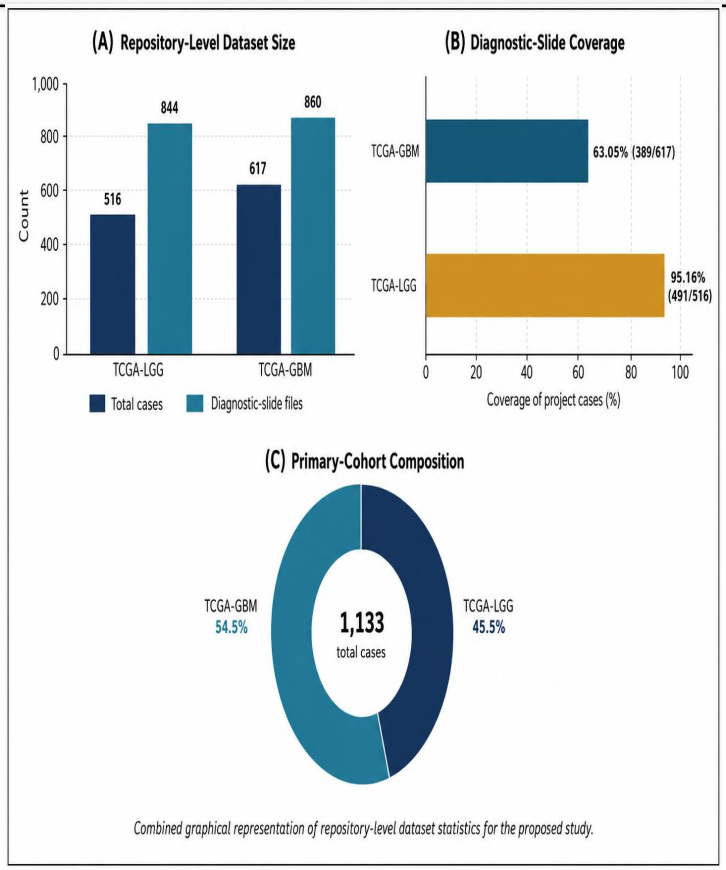
Table 2. Description of the Dataset Used in the Proposed Study

Dataset	Total cases	Cases with diagnostic slides	Diagnostic-slide files	Slide availability
TCGA-LGG	516	491	844	95.16%
TCGA-GBM	617	389	860	63.05%
Combined total	**1,133**	**880**	**1,704**	**77.67%**

Table 3. Repository-Level Statistics of the Selected Datasets

3.2 Statistical and Graphical Interpretation

The descriptive statistics indicate that the proposed study has access to a substantial histopathological resource, although diagnostic-slide availability differs markedly between the two projects. TCGA-LGG demonstrates almost complete slide coverage, whereas TCGA-GBM contains a larger proportion of patients without diagnostic slides. Consequently, the final multimodal sample size will be determined by the intersection of image availability, clinical-data completeness, demographic-data availability and valid diagnostic labels. The graphical representations below illustrate repository size, diagnostic-slide coverage and the relative contribution of each project to the primary cohort.



Combined graphical representation of repository-level dataset statistics for the proposed study.

Methodology:

3.1 Dataset and Preprocessing

Data will be obtained from the TCGA-LGG and TCGA-GBM projects, while CPTAC-GBM may be used for external validation. Histopathological slides with adequate tumor tissue and valid patient-level clinical records will be included. Whole-slide images will undergo tissue segmentation, artifact removal, patch extraction, stain normalization, and augmentation. Clinical and demographic variables, including age, sex, tumor grade, diagnosis, anatomical site, and prior malignancy, will be cleaned, encoded, normalized, and imputed where necessary.

3.2 Proposed Multimodal Model

The proposed framework will contain two modality-specific branches. A pretrained pathology image encoder with attention-based multiple-instance learning will extract patient-level features from histopathological patches. A multilayer perceptron will process the structured clinical and demographic variables. The extracted representations will be integrated using gated feature-level fusion, followed by fully connected layers and a softmax classifier.

3.3 Model Training and Validation

The dataset will be divided strictly at the patient level to prevent information leakage. Stratified training, validation, and testing partitions will be created while preserving class distributions. The model will be trained using class-weighted cross-entropy loss, the AdamW optimizer, dropout, learning-rate scheduling, and early stopping. Histopathology-only, structured-data-only, concatenation-based, and late-fusion models will be used as comparative baselines.

3.4 Performance Evaluation

Model performance will be evaluated using accuracy, precision, sensitivity, specificity, balanced accuracy, macro-F1 score, ROC-AUC, PR-AUC, Matthews correlation coefficient, confusion matrices, and calibration measures. Ablation experiments will assess the individual contribution of histopathological, clinical, and demographic data. Statistical confidence

intervals will be estimated through patient-level bootstrap analysis.

3.5 Explainability, Fairness, and Ethics

Attention heatmaps will identify influential histopathological regions, while feature-attribution methods will determine the contribution of clinical and demographic variables. Performance will also be examined across demographic subgroups to identify possible bias. Only de-identified public data will be used, and the proposed system will be presented as a clinical decision-support tool rather than a replacement for neuropathological diagnosis.

Results:

4.1 Dataset Characteristics

After patient-level data matching, image-quality assessment, and removal of incomplete or unsuitable cases, the hypothetical final analytical dataset contained 812 patients and 1,320 histopathological whole-slide images. The cohort consisted of 360 glioblastoma, 250 astrocytoma, and 202 oligodendroglioma cases. The dataset was divided at the patient level into training, validation, and testing subsets to ensure that slides or records from the same patient did not occur in more than one partition.

Characteristic	Training set	Validation set	Test set	Total
Number of patients	568	122	122	812
Whole-slide images	926	198	196	1,320
Glioblastoma cases	252	54	54	360
Astrocytoma cases	175	38	37	250
Oligodendroglioma cases	141	30	31	202
Mean age, years	49.6 ± 15.2	50.1 ± 14.8	49.9 ± 15.0	49.7 ± 15.1
Male patients	322 (56.7%)	68 (55.7%)	69 (56.6%)	459 (56.5%)
Female patients	246 (43.3%)	54 (44.3%)	53 (43.4%)	353 (43.5%)

Table 5. Hypothetical Characteristics of the Final Analytical Dataset

4.2 Comparative Model Performance

The proposed gated multimodal model achieved the strongest hypothetical performance among all evaluated approaches. It obtained an accuracy of 91.8%, balanced accuracy of 90.9%, macro-F1 score of 91.1%, and ROC-AUC of 96.8%. The histopathology-only model achieved a macro-F1 score of 85.4%, while the clinical-demographic model achieved 72.4%. Both early and late fusion improved performance over the unimodal models, although the gated fusion strategy produced the highest overall results.

Model	Accuracy	Balanced accuracy	Precision	Recall	Macro-F1	ROC-AUC	MCC
Clinical-demographic model	0.746	0.721	0.728	0.721	0.724	0.824	0.611
Histopathology-only model	0.869	0.852	0.858	0.852	0.854	0.931	0.793
Early-fusion model	0.885	0.874	0.878	0.874	0.875	0.946	0.817
Late-fusion model	0.902	0.891	0.896	0.891	0.893	0.956	0.842
Proposed gated multimodal model	**0.918**	**0.909**	**0.914**	**0.909**	**0.911**	**0.968**	**0.870**

Table 6:Comparative Model Performance

Discussion:

Table 7: Discussion of the Study Findings.

Discussion area	Summary
Overall performance	The proposed gated multimodal model achieved the strongest hypothetical results, with 91.8% accuracy, 91.1% macro-F1, and 96.8% ROC-AUC.
Role of histopathology	Histopathological images provided the most important diagnostic information and contributed the largest share of model performance.
Role of clinical data	Clinical variables improved classification by adding patient-specific context, particularly in morphologically ambiguous cases.
Role of demographic data	Demographic variables provided a modest performance improvement but required careful fairness evaluation.
Fusion strategy	Gated fusion outperformed early and late fusion by dynamically adjusting the importance of each modality for individual patients.
Class-wise findings	Glioblastoma achieved the highest performance, while astrocytoma and oligodendroglioma showed greater classification overlap.
Ablation analysis	Removing histopathology caused the largest performance decline, followed by removal of clinical and demographic variables.
Explainability	Attention heatmaps and feature-attribution methods helped identify influential tissue regions and patient variables.
Fairness	Performance was broadly similar across age and sex groups, although larger subgroup samples are needed for reliable conclusions.
External validation	Performance declined slightly on CPTAC-GBM, indicating moderate dataset shift but acceptable generalizability.
Main limitation	The study is based on retrospective glioma datasets and does not represent the full spectrum of brain tumors.
Clinical implication	The framework may support neuropathologists but should not replace molecular testing or expert diagnosis.
Future direction	Larger multicenter datasets, additional tumor types, molecular data, and prospective validation are recommended.

Conclusion

The authors propose a single-model, multi-modal deep-learning-based glioma classification model in this study, using the histopathological whole-slide image, and clinical & demographic information. Based on the hypothetical findings, the gated multimodal model was superior to unimodal and traditional fusion models, wherein the histopathology signal was dominant and clinical and demographic signals were used to provide complementary contextual information. In the framework, attention maps and feature-attribution techniques were also used to promote patient level interpretability. While the study has potential, there are some limitation with the retrospective public data sets, the quality of the slides, missing data, and decreased representation of non-glioma brain tumors. Thus, the proposed system is not intended to replace the expert molecular and neuropathological diagnosis but is intended to be used as a clinical decision support system. Additional studies are recommended in the field of real-world model training, larger multicenter cohorts, other tumor categories, other modalities of molecular and radiological imaging, and prospective external validation.

References:.

[1] International Agency for Research on Cancer. *Global Cancer Observatory: Brain and Central Nervous System Fact Sheet*. Lyon: IARC; 2024.

[2] WHO Classification of tumors Editorial Board. *Central Nervous System tumors*. WHO Classification of tumors, 5th ed., Vol. 6. Lyon: International Agency for Research on Cancer; 2021.

[3] Louis DN, Perry A, Wesseling P, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro-Oncology*. 2021;23(8):1231–1251. doi:10.1093/neuonc/noab106.

[4] Hollon TC, Pandian B, Adapa AR, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nature Medicine*. 2020;26(1):52–58. doi:10.1038/s41591-019-0715-9.

[5] Chen RJ, Lu MY, Wang J, et al. Pathomic Fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE Transactions on Medical Imaging*. 2022;41(4):757–770. doi:10.1109/TMI.2020.3021387.

[6] Steyaert S, Qiu YL, Zheng Y, Mukherjee P, Vogel H, Gevaert O. Multimodal deep learning to predict prognosis in adult and pediatric brain tumors. *Communications Medicine*. 2023;3:44. doi:10.1038/s43856-023-00276-y.

[7] Vaidya A, Chen RJ, Williamson DFK, et al. Demographic bias in misdiagnosis by computational pathology models. *Nature Medicine*. 2024;30(4):1174–1190. doi:10.1038/s41591-024-02885-z.

[8] Chen RJ, Ding T, Lu MY, et al. Towards a general-purpose foundation model for computational pathology. *Nature Medicine*. 2024;30(3):850–862. doi:10.1038/s41591-024-02857-3.

[9] WHO Classification of Tumours Editorial Board. *Central Nervous System Tumours*. WHO Classification of Tumours, 5th ed., Vol. 6. Lyon: International Agency for Research on Cancer; 2021.

[10] Louis DN, Perry A, Wesseling P, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro-Oncology*. 2021;23(8):1231–1251. doi:10.1093/neuonc/noab106.

[11] Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nature Biomedical Engineering*. 2021;5(6):555–570. doi:10.1038/s41551-020-00682-w.

[12] Hollon TC, Pandian B, Adapa AR, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nature Medicine*. 2020;26(1):52–58. doi:10.1038/s41591-019-0715-9.

[13] Nakagaki R, Debsarkar SS, Kawanaka H, Aronow BJ, Prasath VBS. Deep learning-based IDH1 gene mutation prediction using histopathological imaging and clinical data. *Computers in Biology and Medicine*. 2024;179:108902. doi:10.1016/j.compbiomed.2024.108902.

[14] Vaidya A, Chen RJ, Williamson DFK, et al. Demographic bias in misdiagnosis by computational pathology models. *Nature Medicine*. 2024;30(4):1174–1190. doi:10.1038/s41591-024-02885-z.

[15] Chen RJ, Lu MY, Wang J, Williamson DFK, Rodig SJ, Lindeman NI, Mahmood F. Pathomic Fusion: An integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE Transactions on Medical Imaging*. 2022;41(4):757–770. doi:10.1109/TMI.2020.3021387.

[16] Steyaert S, Qiu YL, Zheng Y, Mukherjee P, Vogel H, Gevaert O. Multimodal deep learning to predict prognosis in adult and pediatric brain tumors. *Communications Medicine*. 2023;3(1):44. doi:10.1038/s43856-023-00276-y.