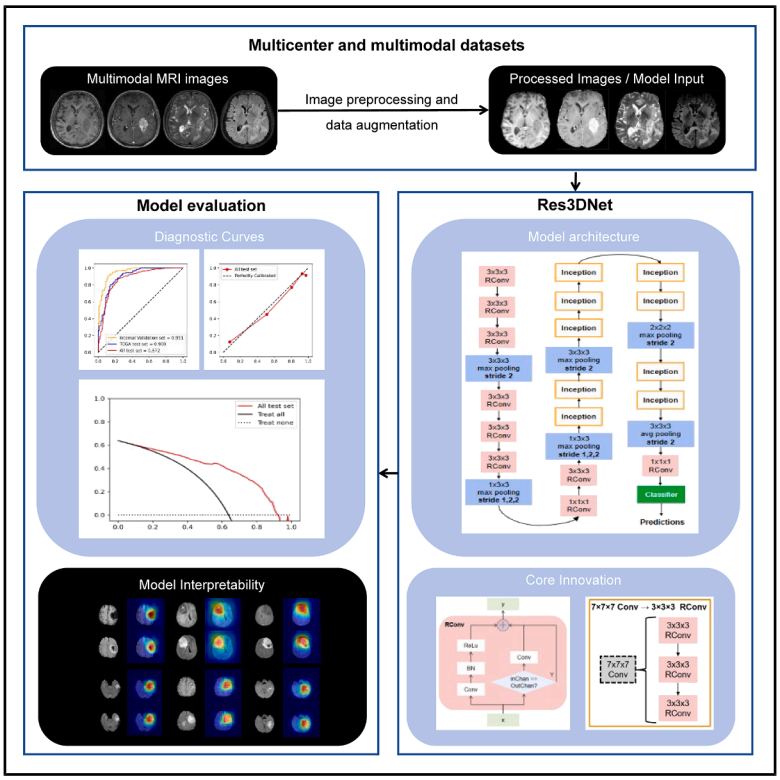
iScience

Article

Fully automated Res3DNet model to predict IDH mutation of gliomas from whole-brain MRI free of tumor segmentation

Graphical abstract



Highlights

- Res3DNet model improved deep learning model's essential design
- Res3DNet enhanced radiologists' efficiency for assessing IDH mutation status in gliomas
- Res3DNet is evaluated in multiple centers with robust generalizability

Authors

Zhengyang Zhu, Yuying Liu, Yang Song, ..., Xin Zhang, Wu-Jun Li, Bing Zhang

Correspondence

seuyhq@163.com (H.Y.), zhangxin@njglyy.com (X.Z.), liwujun@nju.edu.cn (W.-J.L.), zhangbing_nanjing@nju.edu.cn (B.Z.)

In brief

Neurology; oncology; artificial intelligence

Zhu et al., 2026, iScience 29, 114373
January 16, 2026 © 2025 The Authors. Published by Elsevier Inc.
https://doi.org/10.1016/j.isci.2025.114373

CellPress

iScience

CellPress
OPEN ACCESS

Article

Fully automated Res3DNet model to predict IDH mutation of gliomas from whole-brain MRI free of tumor segmentation

Zhengyang Zhu,[1,8] Yuying Liu,[2,8] Yang Song,[1,8] Jianan Zhou,[1,8] Sixuan Chen,[1] Meiping Ye,[1] Zhiqiang Zhang,[3] Jiaming Lu,[1] Xin Li,[1] Xinru Xu,[1] Cong Long,[1] Linqing Fu,[1] Yajing Zhu,[1] Xu Yang,[1] Weiping Li,[1] Yijun Bai,[1] Shunshun Du,[1] Huiquan Yang,[1,*] Xin Zhang,[1,*] Wu-Jun Li,[2,4,5,*] and Bing Zhang[1,6,7,9,*]

[1]Department of Radiology, Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing 210008, China
[2]National Key Laboratory for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, Nanjing, China
[3]Department of Diagnostic Radiology, Affiliated Jinling Hospital, Medical School of Nanjing University, Nanjing 210002, China
[4]Center of Medical Big Data, Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing 210002, China
[5]National Institute of Healthcare Data Science at Nanjing University, Nanjing 210002, China
[6]Medical Imaging Center, Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing 210008, China
[7]Institute of Medical Imaging and Artificial Intelligence, Nanjing University, Nanjing 210008, China
[8]These authors contributed equally
[9]Lead contact
*Correspondence: seuyhq@163.com (H.Y.), zhangxin@njglyy.com (X.Z.), liwujun@nju.edu.cn (W.-J.L.), zhangbing_nanjing@nju.edu.cn (B.Z.)
https://doi.org/10.1016/j.isci.2025.114373

SUMMARY

Accurate prediction of IDH mutation status in gliomas is critical for guiding diagnosis, prognosis, and treatment planning. We enrolled 2,537 preoperative MRI of glioma patients (mean age 55.91 ± 14.79, 1,063 females) from 11 different datasets, consisting of 1,382 patients (mean age 58.26 ± 14.38, 548 females) in training set, 346 patients (mean age 57.43 ± 14.04, 141 females) in internal validation set, and 809 patients (mean age 53.92 ± 14.04, 374 females) in external test set, including 242 patients from The Cancer Genome Archive (TCGA) dataset. A fully automated Res3DNet model was established for isocitrate dehydrogenase (IDH) gene prediction. Four radiologists also read images from TCGA test dataset as a comparison with the deep learning model. Our Res3DNet model achieved AUCs of 0.946 (internal validation), 0.872 (external test), and 0.912 (TCGA test), with corresponding accuracies of 0.925, 0.806, and 0.840, respectively, outperforming ResNet model, I3D model, transformer model, and four radiologists.

INTRODUCTION

Adult diffuse gliomas are the most common type among the central nervous system (CNS) tumors with a propensity for invasive growth and a poor prognosis.[1–3] In the 5th edition 2021 WHO classification of CNS tumors, isocitrate dehydrogenase (IDH) is the critical molecular marker in glioma diagnosis and particularly important in distinguishing astrocytoma and oligodendroglioma from glioblastoma.[4] IDH mutant patients have longer progression-free survival and overall survival than IDH wild type patients, making the detection of IDH mutation essential for both prognosis evaluation and treatment decision.[5–7] Considering the therapeutic implications, accurate and preoperative detection of IDH mutations has become a vital aspect of clinical glioma management.

In the clinical settings, IDH mutation status remains unknown until Sanger sequencing for the resected tumor sample in the surgery, which is time-consuming and expensive.[8] Magnetic resonance imaging (MRI) is the cornerstone of preoperative diagnosis of glioma.[9,10] MRI can provide detailed anatomical characteristics of tumors and may reflect the genetic information within the tumor microenvironment.[11,12] MRI has the potential to serve as an alternative, noninvasive method to identify crucial molecular markers of gliomas, which is highly relevant when these genetic alterations are required rapidly to determine the eligibility of clinical trials, sometimes even prior to surgical resection when the trial design involves initiating drugs before surgery.

In recent years, artificial intelligence and deep learning has revolutionized medical imaging analysis, including in the aspects of glioma radio-genomics.[13–17] Deep learning models have demonstrated outstanding performance in automatically extraction of complex features from MRI, allowing for further analysis, classification and prediction of the underlying tumor genetic characteristics. Specifically, deep learning models trained on preoperative MRI images have been successfully utilized in predicting molecular alterations of gliomas, including IDH mutation evaluation.[18–21] However, most of

This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Table 1. Patient characteristics

	Training set (n = 1382)	Internal validation set (n = 346)	p value (training vs. internal validation)	External test set (n = 809)	p value (training vs. external test)
Age	58.26 ± 14.38	57.43 ± 14.04	0.308	53.92 ± 14.04	<0.001
Sex	–	–	0.716	–	0.002
Male	838 (60.32%)	205 (59.25%)	–	431 (53.54%)	–
Female	548 (39.68%)	141 (40.75%)	–	374 (46.46%)	–
IDH mutation	–	–	0.177	–	<0.001
wildtype	1119 (80.97%)	269 (77.75%)	–	516 (63.78%)	–
mutated	263 (19.03%)	77 (22.25%)	–	293 (36.22%)	–
WHO grade	–	–	0.613	–	<0.001
2	154 (11.24%)	45 (13.12%)	–	163 (20.15%)	–
3	52 (3.80%)	12 (3.50%)	–	95 (11.74%)	–
4	1164 (84.96%)	286 (83.38%)	–	551 (68.11%)	–

these researches enrolled limited number of patients, which limited the robustness and generalizability of deep learning models.

Another main challenge in the application of deep learning for glioma imaging evaluation is the need for accurate tumor segmentation.[22,23] Traditionally, tumor segmentation in glioma

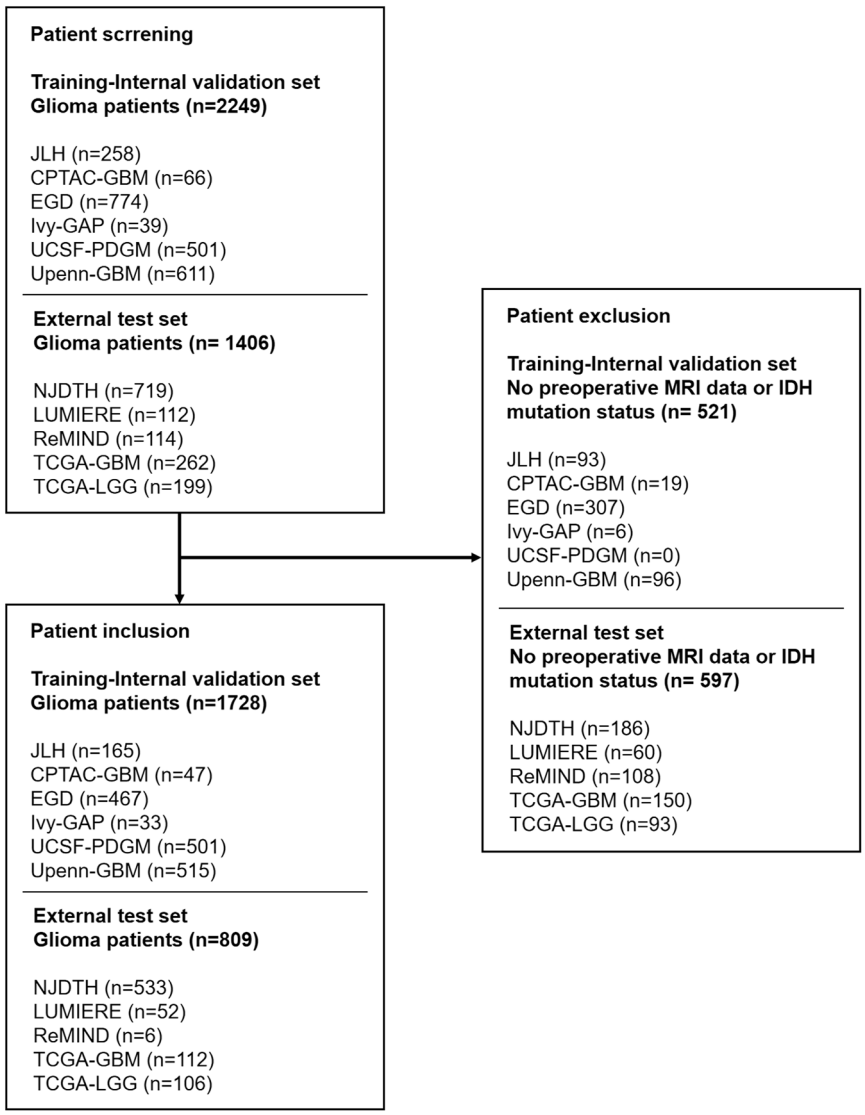


Figure 1. Inclusion-exclusion flowchart of glioma patients

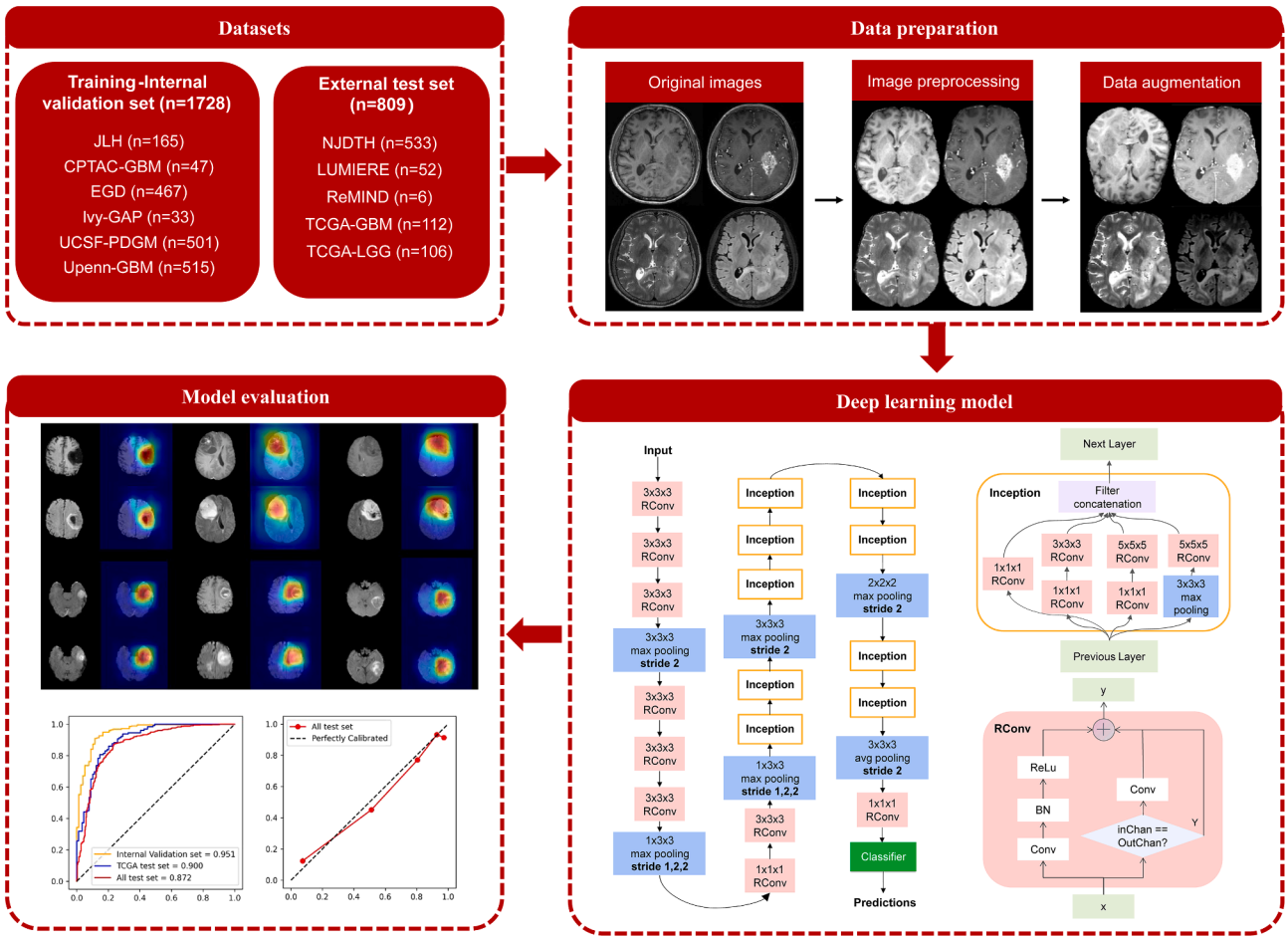


Figure 2. The study design of this research

imaging requires the intervention of experienced board-certificated neuroradiologists and can be subjected to personal variability, especially in these cases with irregular tumor boundaries. Manual tumor segmentation is time-consuming and impractical in a clinical setting and cannot be incorporated in daily clinical routine where rapid clinical decision is usually essential. To overcome these limitations, we created a new deep learning model Res3DNet that do not need explicit tumor segmentation. This model relied on the entire volume of the brain, encompassing not only tumor regions but also peri-tumor regions to capture more comprehensive information of tumor's microenvironment and its interaction of surrounding brain tissues. This method helped eliminate the need for tumor segmentation and facilitate the generalizability and robustness of the predictive model by allowing the model to learn from the whole-brain region without being biased by segmented tumor areas.

The aim of this study was to predict IDH mutation of glioma patients from preoperative MR images free of tumor segmentation using a novel and fully automated Res3DNet deep learning model. We hypothesize that the Res3DNet model can outperform other traditional deep learning structures such as ResNet, I3D and Transformer.

RESULTS

Baseline patients' characteristics

We enrolled a total of 2,537 patients in our study. The patients in the training-internal validation set originated from seven different datasets and the test set data were collected from four different datasets, respectively. 1,382 patients (mean age 58.26 ± 14.38, 548 females) were enrolled in the training set. 346 patients (mean age 57.43 ± 14.04, 141 females) were enrolled in the internal validation set and 809 patients (mean age 53.92 ± 14.04, 374 females) were enrolled in the test set. Table 1 illustrates a full overview of the baseline characteristics in the training set, internal validation set and external test set. Figure 1 provides the inclusion flowchart of patients. Figure S4 illustrates the IDH mutation distribution through training set, external test set and The Cancer Genome Archive (TCGA) test set. Figure 2 illustrates the study design of this research.

Performance of deep learning models in predicting IDH mutation in adult diffuse gliomas

On internal validation set, the ResNET model achieved an area under the curve (AUC) of 0.896 (95% CI: 0.853–0.935) and an

Table 2. Performance of ResNET model, I3D model, Res3DNet model, and radiologists in predicting IDH mutation status of gliomas

Dataset	Model/radiologist	AUC (95% CI)	AUPRC	Accuracy	Recall	Precision	Specificity	F1 score
Internal validation set	ResNET	0.896 (0.853–0.935)	0.962	0.846	0.856	0.940	0.815	0.896
	I3D	0.935 (0.898–0.965)	0.976	0.904	0.934	0.942	0.824	0.938
	Transformer	0.935 (0.913–0.956)	0.937	0.875	0.921	0.859	0.841	0.889
	Res3DNet	0.946 (0.915–0.973)	0.981	0.925	0.953	0.951	0.877	0.952
External test set	ResNET	0.782 (0.750–0.814)	0.861	0.722	0.821	0.763	0.547	0.790
	I3D	0.843 (0.814–0.871)	0.888	0.770	0.904	0.774	0.534	0.834
	Transformer	0.829 (0.858–0.942)	0.753	0.884	0.926	0.747	0.577	0.827
	Res3DNet	0.872 (0.845–0.897)	0.906	0.806	0.927	0.8006	0.706	0.859
TCGA test set	ResNET	0.869 (0.818–0.914)	0.901	0.771	0.941	0.742	0.524	0.829
	I3D	0.909 (0.866–0.947)	0.921	0.827	0.915	0.816	0.700	0.862
	Transformer	0.903 (0.800–0.927)	0.922	0.817	0.922	0.799	0.799	0.856
	Res3DNet	0.912 (0.871–0.947)	0.928	0.840	0.925	0.825	0.834	0.872
	Van der Voort et al.[17]	0.900 (0.850–0.950)	–	0.840	0.780	–	0.930	–
	Choi et al.[24]	0.860	0.810	0.788	–	–	–	–
	Radiologist1	0.760	–	0.798	0.892	0.814	0.628	0.851
	Radiologist2	0.705	–	0.745	0.841	0.781	0.570	0.810
	Radiologist3	0.698	–	0.753	0.885	0.768	0.512	0.822
	Radiologist4	0.799	–	0.819	0.866	0.855	0.733	0.861

AUC, area under the curve.

accuracy of 0.846; the I3D model achieved an AUC of 0.935 (95% CI: 0.898–0.965) and an accuracy of 0.904; the transformer model achieved an AUC of 0.935 (95% CI: 0.913–0.956) and an accuracy of 0.925; and the Res3DNet model achieved an AUC of 0.946 (95% CI: 0.915–0.973) and an accuracy of 0.925, as illustrated in Table 2 and Figure 3.

On external test set, the ResNET model achieved an AUC of 0.782 (95% CI: 0.750–0.814) and an accuracy of 0.722; the I3D model achieved an AUC of 0.843 (95% CI: 0.814–0.871) and an accuracy of 0.770; the transformer model achieved an AUC of 0.829 (95% CI: 0.858–0.942) and an accuracy of 0.753; and the Res3DNet model achieved an AUC of 0.872 (95% CI: 0.845–0.897) and an accuracy of 0.806, as illustrated in Table 2 and Figure 3.

On TCGA test set, the ResNET model achieved an AUC of 0.869 (95% CI: 0.818–0.914) and an accuracy of 0.771; the I3D model achieved an AUC of 0.909 (95% CI: 0.866–0.947) and an accuracy of 0.798; the transformer model achieved an AUC of 0.903 (95% CI: 0.800–0.857) and an accuracy of 0.817; and the Res3DNet model achieved an AUC of 0.912 (95% CI: 0.850–0.950) and an accuracy of 0.840; four board-certificated radiologists achieved an AUC of 0.760, 0.705, 0.698, and 0.799 separately with accuracy of 0.798, 0.745, 0.753, and 0.819 separately, as illustrated in Table 2 and Figure 3. All of the ResNET, I3D, transformer, and Res3DNet models outperformed neuroradiologists. The Res3DNet model performed better than previously reported methods.[17,24]

According to the results of Friedman test and DeLong test, the Res3DNet model performed better than the ResNET model, I3D model, and transformer model, as seen in Table 3. The performance of the Res3DNet model in predicting IDH mutation stratified by MRI Scanner Vendor on the external test set is illustrated in Table 4. For Siemens, General Electric, Philips, and United Imaging, Res3DNet achieved AUC of 0.919, 0.890, 0.889, and 0.855, separately. For non-enhancing glioma subgroup, T2-FLAIR mismatch sign showed specificity of 1.000, compared with specificity of 0.714) for the Res3DNet model, However, the Res3DNet exhibited higher recall than T2-FLAIR mismatch sign (0.773 vs. 0.439), seen in Table 5. This suggests that the Res3DNet model could serve as a valuable complementary tool to assist radiologists, particularly in cases where the T2-FLAIR mismatch sign is absent.

Results of GradCAM visualization

In IDH mutation status prediction task, our model exhibits high gradients in tumor regions and low gradients in other regions, validating the importance of focusing on tumor regions in the classification of the aforementioned tasks, as shown in Figure 4.

DISCUSSION

In this research, we enrolled a total of 2,537 patients with diffuse gliomas and developed an automated Res3DNet model to predict IDH mutation status without tumor segmentation. After developing the deep learning model based on seven different datasets spanning Asia, America, and Europe our automated Res3DNet model allowed for predicting IDH mutation effectively, with AUC of 0.872 and ACC of 0.806 on external test set. The Res3DNet performed satisfactorily in IDH gene mutation task, since radiologists cannot evaluate the status of molecular alterations by reading MRI images using naked eyes. In this study, four board-certificated neuro-radiologists predicted IDH mutation status with accuracy ranging from 0.748 to 0.819.

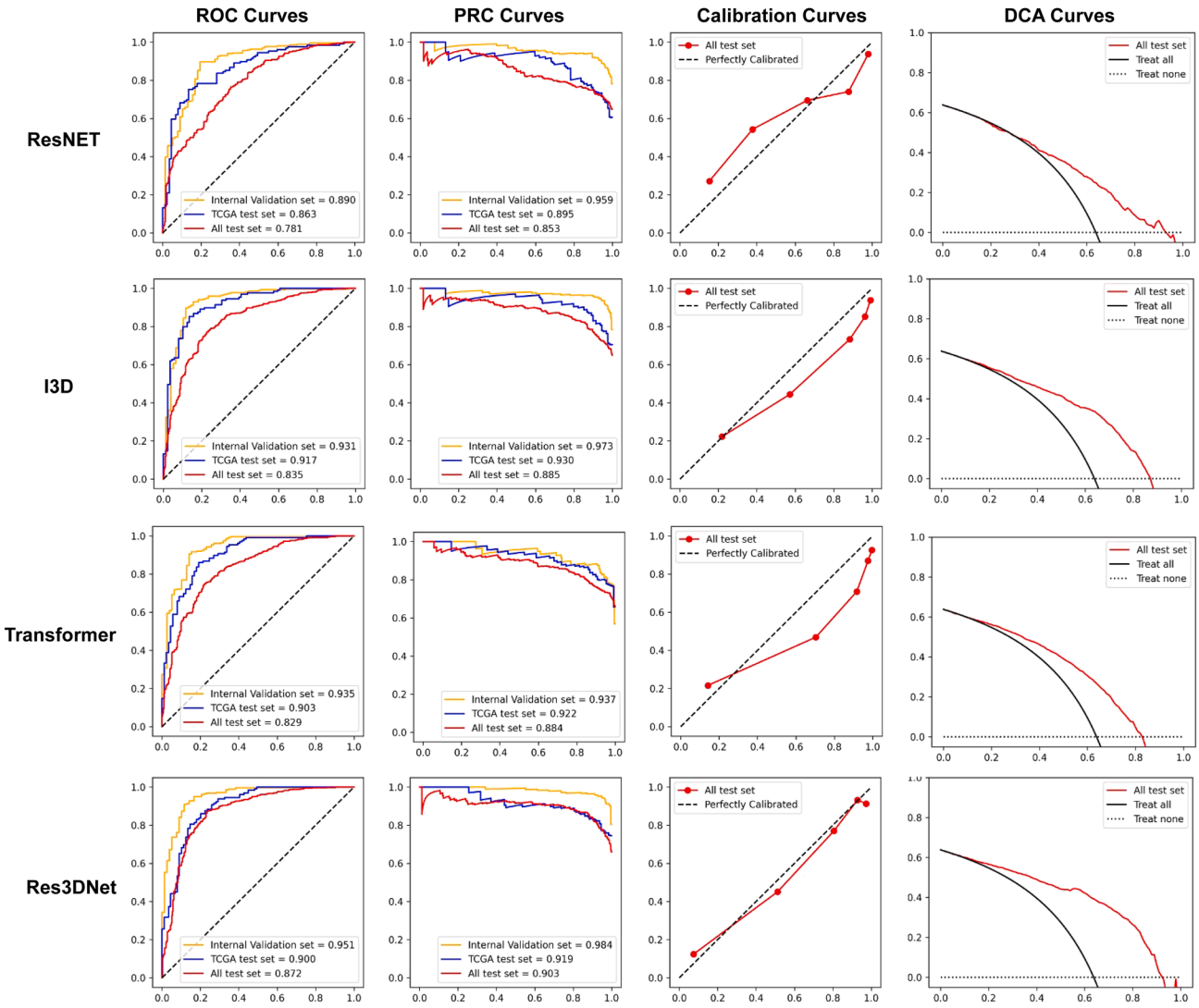


Figure 3. ROC curves, PRC curves, calibration curves, and DCA curves of ResNET, I3D, and Res3DNet model in predicting IDH mutation in gliomas

In real-world cohorts of adult diffuse gliomas, the IDH mutation rate is approximately 30%.[25] In our study, the IDH mutation rate in the training set was around 20%, primarily due to the composition of the UPenn-GBM dataset, which consists predominantly of IDH-wildtype glioblastomas. This skewed the overall mutation rate lower in the training cohort. However, the IDH mutation rate in the independent external test set closely reflects real-world distributions, helping to ensure the generalizability of our evaluation.

Table 3. Performance comparison of deep learning models for IDH mutation prediction in gliomas using external test set

Model comparison	*p* value for Friedman test	*p* value for Delong test
Res3DNet vs. ResNet	<0.001	<0.001
Res3DNet vs. I3D	<0.001	<0.001
Res3DNet vs. transformer	<0.001	<0.001

The slightly lower performance of the model on the external test set (AUC = 0.872) compared to the TCGA test set (AUC = 0.912) may be partly attributed to differences in data distribution between datasets. Specifically, the external test set had a younger patient population and a slightly higher IDH mutation rate, both of which may influence model performance. Age is a known clinical correlate of IDH mutation status, and a younger cohort may reduce the discriminative power of age-related imaging features learned by the model. Additionally, variation in mutation prevalence can affect the balance of classes during evaluation, potentially impacting metrics such as AUC and F1 score. These findings highlight the importance of validating deep learning models across diverse and demographically representative datasets to ensure robust generalizability in real-world clinical settings.

The Res3DNet model has demonstrated the potential to characterization of glioma and supporting clinical decision. Our

Table 4. Performance of the Res3DNet model in predicting IDH mutation stratified by MRI scanner vendor on the external test set

Vendors	AUC	AUPRC	Accuracy	Recall	Precision	Specificity	F1 score
Siemens	0.919	0.984	0.902	0.936	0.948	0.889	0.942
General electric	0.890	0.898	0.829	0.953	0.788	0.792	0.863
Philips	0.889	0.915	0.816	0.890	0.832	0.720	0.860
United imaging	0.855	0.879	0.757	0.942	0.736	0.462	0.826

AUC, area under the curve; AUPRC, area under the precision-recall curve.

model enhances the expressive capability of deep neural networks by introducing residual convolutional blocks, which effectively mitigates problems such as gradient vanishing and gradient explosion. Meanwhile, three 3 × 3×3 convolutional kernels are used instead of the large convolutional kernel, which reduces the semantic loss and improves the model accuracy. In addition, we used a hyperparameter alpha for mitigating the class imbalance problem in bi-classification effectively address the class imbalance problem and improve the performance of the model on imbalanced datasets. Together, these innovations constitute the core strengths of the Res3DNet model, enabling it to demonstrate superior performance in IDH mutation prediction tasks.

The prediction of molecular alterations is of great significance for the treatment of diffuse gliomas.[7,26] In WHO CNS5, IDH is one of the most important genetic markers in diffuse glioma classification.[4] Acquiring molecular alteration information preoperatively could guide surgical strategies and select potential candidates for molecular targeted therapies and neoadjuvant therapies, which could give patients with poor prognosis a chance to participate in clinical trial in time and get better treatment.

Our research expanded the work of several recent studies revealing associations between IDH mutation status and MRI manifestations. Nevertheless, most of previous researches needed accurate tumor segmentation to extract radiomic features or deep learning features, which required the intervention of neuro-oncology experts and could be subject to intra-observer variability.[27–30] Some studies incorporated segmentation and classification algorithms to establish a combined fully automated model. Choi and his colleagues proposed a fully automated hybrid approach to segment tumor area and predict IDH mutation status of gliomas via convolutional neural networks (CNNs) and radiomics, achieving AUC of 0.86 and ACC of 0.788 on TCGA test set.[24] Van der Voort developed a single multitask CNN model to automatically segment the hyperintense area on T2WI and predict several genetic and histological features of gliomas. On TCGA test set, the combined model obtained AUC of 0.90 and ACC of 0.84 in IDH mutation task.[17] Our Res3DNet model utilized whole-brain images and model input and achieved AUC of 0.917 and ACC of 0.844 on the same TCGA test set used in previous studies. Even though we didn't incorporate segmentation framework, the Res3DNet model performed well in IDH classification task compared with previous work using tumor segmentation mask. Avoiding tumor segmentation is a significant advantage of our approach, as it simplifies the clinical workflow and improves efficiency. Tumor segmentation is often time-consuming, labor-intensive, and highly dependent on radiologists' expertise and experience. By leveraging a deep learning model that directly processes whole-brain multimodal MRI data, our method eliminates the need for precise tumor delineation. This not only accelerates the prediction process, enabling near real-time results, but also reduces reliance on expert annotation, lowers deployment costs, and enhances accessibility, especially in resource-limited clinical settings. This segmentation-free design facilitates seamless integration into routine practice and represents a core strength of our study.

Grad-CAM visualizations suggest that our Res3DNet model primarily attends to tumor regions when predicting IDH mutation status. This is biologically plausible, as IDH-mutant gliomas often present as non-enhancing lesions with minimal edema, reflecting limited blood-brain barrier disruption. In contrast, IDH-wildtype gliomas are frequently characterized by contrast enhancement, necrosis, and cystic components on T1CE, indicating more aggressive behavior and disrupted vascular integrity.[31] On T2WI, IDH-wildtype gliomas tend to exhibit heterogeneous signal intensity and extensive edema, whereas IDH-mutant tumors often appear more homogeneous and well-demarcated.[32] These imaging features may relate to underlying pathophysiological differences. For IDH-mutant gliomas, the mutated IDH can cause the accumulation of the oncometabolite 2-hydroxyglutarate, which alters DNA and histone methylation, leading to epigenetic reprogramming.[33] This biological process can is related to reduced angiogenesis, decreased cell densities and less aggressive tumor phenotype.[34] The model's attention maps may help radiologists better understand how imaging correlates with molecular subtypes. The deep learning model can objectively quantify complex imaging feature patterns that are often subtle or imperceptible to the human eye. This capability likely contributes to its superior performance compared to expert radiologist visual assessment.

Our results imply that our fully automated Res3DNet model can be generalized across different MRI vendors, imaging protocols, and clinical cohorts, and it consistently yield better

Table 5. Performance of T2-FLAIR mismatch sign and Res3DNet model in predicting IDH mutation in non-enhancing gliomas

	AUC	AUPRC	Accuracy	Recall	Precision	Specificity	F1 score
T2-FLAIR mismatch sign	0.720	0.275	0.538	0.439	0.275	1.000	0.431
Res3DNet	0.838	0.604	0.762	0.773	0.400	0.714	0.513

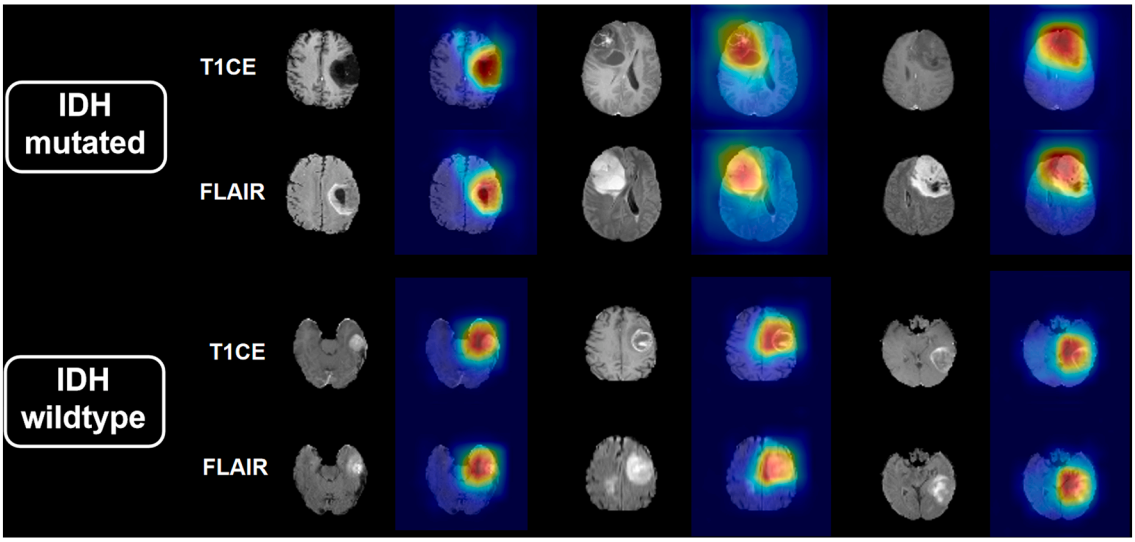


Figure 4. GradCAM visualization of IDH-mutated and IDH-wildtype cases

performance than the use of ResNET and I3D framework. Our proposed Res3DNet framework achieves an approximately 90%–95% reduction in analysis time compared to traditional segmentation-based methods. To be specific: the conventional approach requires expert radiologists to perform meticulous 3D tumor segmentation, a process that typically consumes 15–20 min per case. In stark contrast, our end-to-end deep learning model delivers the final IDH status prediction automatically in approximately 1 min after initial preprocessing, completely eliminating the need for manual intervention. In rural or source-limited areas, there may not be enough specialized experienced neuroradiologists to read complex cases like gliomas. By providing real-time IDH mutation predictions alongside MRI scans, our fully automated Res3DNet may serve as an effective tool to help radiologists read glioma images more effectively and improve diagnostic efficacy. Incorporating the Res3DNet model into clinical practice would provide neurosurgeons and neuro-oncologists with a decision support system to help interpret MRI scans more efficiently and make personalized surgical and treatment plans for glioma patients.

Limitations of the study

There are several limitations in this research. First of all, this research only focused on four conventional structural MRI sequences: T1WI, T2WI, FLAIR, and T1CE, without considering other advanced MRI sequences such as diffusion weighted images and dynamic contrast enhanced. Those advanced sequences can provide information about tumor cell density and microcirculation perfusion and may increase the model performance in IDH mutation prediction. The exclusion was primarily due to the limited availability and heterogeneity of these sequences across institutions. Given that our goal was to construct a clinically feasible model with the most widely available structural MRI sequences, using advanced multi-modality images may limit the feasibility of the models. Secondly, we only chose the IDH gene mutation as the classification task, whether the Res3DNet model can predict other molecular alterations requires further investigations. Moreover, this is a retrospective study. A prospective, multicenter clinical trial which involves diverse MRI protocol is needed to further evaluate the model performance and ensure consistency and generalizability.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Bing Zhang (zhangbing_nanjing@nju.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Data will be made available from the lead contact on request.
- All original codes have been deposited at Github (https://github.com/YLiu-nju/Res3DNet/), and are publicly accessible as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We would like to thank all glioma patients participating in this study.

This work was supported by the National Science and Technology Innovation 2030 – Major program of "Brain Science and Brain-Like Research" (2022ZD0211800); National Natural Science Foundation of China (82271965, 82330059, 12371377); Jiangsu Provincial Health Commission Medical Research Project (M2024057); and Fundings for Clinical Trials from the Affiliated Drum Tower Hospital, Medical School of Nanjing University (2022-LCYJ-MS-25).

AUTHOR CONTRIBUTIONS

Z.Zhu, Y.L., H.Y., X.Z., W.L., and B.Z. designed the experiments. Y.L. and W.L. developed the models. Z.Zhu and Y.L. created the figures. Z.Zhu, J.Z., S.C., and M.Y. conducted the reader study. Z.Zhang, Y.S., J.L., X.L., and X.X. collected the dataset. Z.Zhu and Y.L. performed the statistical analysis. X.Z., W.L., and B.Z. supervised the work. Z.Zhu and Y.L. wrote the manuscript.

Z.Zhu and Y.L. revised the paper. All authors read and approved the final version of the article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Study population
  - Ethical statement
- METHOD DETAILS
  - Image pre-processing
  - Deep neural network architecture
  - Radiologist-AI competing test
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2025.114373.

Received: April 10, 2025
Revised: August 9, 2025
Accepted: December 4, 2025
Published: December 9, 2025

REFERENCES

1. Pinson, H., Silversmit, G., Vanhauwaert, D., Vanschoenbeek, K., Okito, J.-P.K., De Vleeschouwer, S., Boterberg, T., and De Gendt, C. (2024). Epidemiology and survival of adult-type diffuse glioma in Belgium during the molecular era. Neuro Oncol. *26*, 191–202. https://doi.org/10.1093/neuonc/noad158.
2. Han, B., Zheng, R., Zeng, H., Wang, S., Sun, K., Chen, R., Li, L., Wei, W., and He, J. (2024). Cancer incidence and mortality in China, 2022. J. Natl. Cancer Cent. *4*, 47–53. https://doi.org/10.1016/j.jncc.2024.01.006.
3. van den Bent, M.J., Geurts, M., French, P.J., Smits, M., Capper, D., Bromberg, J.E.C., and Chang, S.M. (2023). Primary brain tumours in adults. Lancet *402*, 1564–1579. https://doi.org/10.1016/S0140-6736(23)01054-1.
4. Louis, D.N., Perry, A., Wesseling, P., Brat, D.J., Cree, I.A., Figarella-Branger, D., Hawkins, C., Ng, H.K., Pfister, S.M., Reifenberger, G., et al. (2021). The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro Oncol. *23*, 1231–1251. https://doi.org/10.1093/neuonc/noab106.
5. Price, M., Neff, C., Nagarajan, N., Kruchko, C., Waite, K.A., Cioffi, G., Cordeiro, B.B., Willmarth, N., Penas-Prado, M., Gilbert, M.R., et al. (2024). CBTRUS Statistical Report: American Brain Tumor Association & NCI Neuro-Oncology Branch Adolescent and Young Adult Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2016-2020. Neuro Oncol. *26*, iii1–iii53. https://doi.org/10.1093/neuonc/noae047.
6. Nakase, T., Guerra, G.A., Ostrom, Q.T., Ge, T., Melin, B.S., Wrensch, M., Wiencke, J.K., Jenkins, R.B., Eckel-Passow, J.E., Glioma International Case-Control Study GICC, et al. (2024). Genome-wide polygenic risk scores predict risk of glioma and molecular subtypes. Neuro Oncol. *26*, 1933–1944. https://doi.org/10.1093/neuonc/noae112.
7. Jiang, T., Nam, D.-H., Ram, Z., Poon, W.-S., Wang, J., Boldbaatar, D., Mao, Y., Ma, W., Mao, Q., You, Y., et al. (2021). Clinical practice guidelines for the management of adult diffuse gliomas. Cancer Lett. *499*, 60–72. https://doi.org/10.1016/j.canlet.2020.10.050.
8. Nelson, E.J., Gubbiotti, M.A., Carlin, A.M., Nasrallah, M.P., Van Deerlin, V.M., and Herlihy, S.E. (2023). Clinical Evaluation of IDH Mutation Status in Formalin-Fixed Paraffin-Embedded Tissue in Gliomas. Mol. Diagn. Ther. *27*, 371–381. https://doi.org/10.1007/s40291-022-00638-7.
9. Kang, K.M., Song, J., Choi, Y., Park, C., Park, J.E., Kim, H.S., Park, S.-H., Park, C.-K., and Choi, S.H. (2024). MRI Scoring Systems for Predicting Isocitrate Dehydrogenase Mutation and Chromosome 1p/19q Codeletion in Adult-type Diffuse Glioma Lacking Contrast Enhancement. Radiology *311*, e233120. https://doi.org/10.1148/radiol.233120.
10. Johnson, D.R., Giannini, C., Vaubel, R.A., Morris, J.M., Eckel, L.J., Kaufmann, T.J., and Guerin, J.B. (2023). A Radiologist's Guide to the 2021 WHO Central Nervous System Tumor Classification: Part I-Key Concepts and the Spectrum of Diffuse Gliomas. Radiology *306*, e229036. https://doi.org/10.1148/radiol.229036.
11. Zhou, J., Hou, Z., Tian, C., Zhu, Z., Ye, M., Chen, S., Yang, H., Zhang, X., and Zhang, B. (2024). Review of tracer kinetic models in evaluation of gliomas using dynamic contrast-enhanced imaging. Front. Oncol. *14*, 1380793. https://doi.org/10.3389/fonc.2024.1380793.
12. Yang, H., Zhu, Z., Long, C., Niu, F., Zhou, J., Chen, S., Ye, M., Peng, S., Zhang, X., Chen, Y., et al. (2024). Quantitative and Qualitative Parameters of DCE-MRI Predict CDKN2A/B Homozygous Deletion in Gliomas. Acad. Radiol. *31*, 3355–3365. https://doi.org/10.1016/j.acra.2024.02.017.
13. Zhu, Z., Shen, J., Liang, X., Zhou, J., Liang, J., Ni, L., Wang, H., Ye, M., Chen, S., Yang, H., et al. (2024). Radiomics for predicting grades, isocitrate dehydrogenase mutation, and oxygen 6-methylguanine-DNA methyltransferase promoter methylation of adult diffuse gliomas: combination of structural MRI, apparent diffusion coefficient, and susceptibility-weighted imaging. Quant. Imaging Med. Surg. *14*, 9276–9289. https://doi.org/10.21037/qims-24-1110.
14. Chen, S., Xu, Y., Ye, M., Li, Y., Sun, Y., Liang, J., Lu, J., Wang, Z., Zhu, Z., Zhang, X., and Zhang, B. (2022). Predicting MGMT Promoter Methylation in Diffuse Gliomas Using Deep Learning with Radiomics. J. Clin. Med. *11*, 3445. https://doi.org/10.3390/jcm11123445.
15. Yan, J., Zhang, B., Zhang, S., Cheng, J., Liu, X., Wang, W., Dong, Y., Zhang, L., Mo, X., Chen, Q., et al. (2021). Quantitative MRI-based radiomics for noninvasively predicting molecular subtypes and survival in glioma patients. NPJ Precis. Oncol. *5*, 72. https://doi.org/10.1038/s41698-021-00205-z.
16. Wu, X., Zhang, S., Zhang, Z., He, Z., Xu, Z., Wang, W., Jin, Z., You, J., Guo, Y., Zhang, L., et al. (2024). Biologically interpretable multi-task deep learning pipeline predicts molecular alterations, grade, and prognosis in glioma patients. NPJ Precis. Oncol. *8*, 181. https://doi.org/10.1038/s41698-024-00670-2.
17. van der Voort, S.R., Incekara, F., Wijnenga, M.M.J., Kapsas, G., Gahrmann, R., Schouten, J.W., Nandoe Tewarie, R., Lycklama, G.J., De Witt Hamer, P.C., Eijgelaar, R.S., et al. (2023). Combined molecular subtyping, grading, and segmentation of glioma using multi-task deep learning. Neuro Oncol. *25*, 279–289. https://doi.org/10.1093/neuonc/noac166.
18. Choi, Y.S., Bae, S., Chang, J.H., Kang, S.-G., Kim, S.H., Kim, J., Rim, T.H., Choi, S.H., Jain, R., and Lee, S.-K. (2021). Fully automated hybrid approach to predict theIDHmutation status of gliomas via deep learning and radiomics. Neuro Oncol. *23*, 304–313. https://doi.org/10.1093/neuonc/noaa177.
19. Cluceru, J., Interian, Y., Phillips, J.J., Molinaro, A.M., Luks, T.L., Alcaide-Leon, P., Olson, M.P., Nair, D., LaFontaine, M., Shai, A., et al. (2022). Improving the noninvasive classification of glioma genetic subtype with deep learning and diffusion-weighted imaging. Neuro Oncol. *24*, 639–652. https://doi.org/10.1093/neuonc/noab238.
20. Wei, Y., Chen, X., Zhu, L., Zhang, L., Schönlieb, C.-B., Price, S., and Li, C. (2023). Multi-Modal Learning for Predicting the Genotype of Glioma. IEEE Trans. Med. Imaging *42*, 3167–3178. https://doi.org/10.1109/tmi.2023.3244038.

21. Xu, Q., Xu, Q.Q., Shi, N., Dong, L.N., Zhu, H., and Xu, K. (2022). A multitask classification framework based on vision transformer for predicting molecular expressions of glioma. Eur. J. Radiol. *157*, 110560. https://doi.org/10.1016/j.ejrad.2022.110560.

22. Kihira, S., Mei, X., Mahmoudi, K., Liu, Z., Dogra, S., Belani, P., Tsankova, N., Hormigo, A., Fayad, Z.A., Doshi, A., and Nael, K. (2022). U-Net Based Segmentation and Characterization of Gliomas. Cancers *14*, 4457. https://doi.org/10.3390/cancers14184457.

23. Cheng, J., Liu, J., Kuang, H., and Wang, J. (2022). A Fully Automated Multimodal MRI-Based Multi-Task Learning for Glioma Segmentation and IDH Genotyping. IEEE Trans. Med. Imaging *41*, 1520–1532. https://doi.org/10.1109/tmi.2022.3142321.

24. Choi, Y.S., Bae, S., Chang, J.H., Kang, S.-G., Kim, S.H., Kim, J., Rim, T.H., Choi, S.H., Jain, R., and Lee, S.-K. (2021). Fully automated hybrid approach to predict the IDH mutation status of gliomas via deep learning and radiomics. Neuro Oncol. *23*, 304–313. https://doi.org/10.1093/neuonc/noaa177.

25. Yan, H., Parsons, D.W., Jin, G., McLendon, R., Rasheed, B.A., Yuan, W., Kos, I., Batinic-Haberle, I., Jones, S., Riggins, G.J., et al. (2009). IDH1 and IDH2 mutations in gliomas. N. Engl. J. Med. *360*, 765–773. https://doi.org/10.1056/NEJMoa0808710.

26. Ye, M., Cao, Z., Zhu, Z., Chen, S., Zhou, J., Yang, H., Li, X., Chen, Q., Luan, W., Li, M., et al. (2025). Integrating quantitative DCE-MRI parameters and radiomic features for improved IDH mutation prediction in gliomas. Front. Oncol. *15*, 1530144. https://doi.org/10.3389/fonc.2025.1530144.

27. Xie, S.-H., Lang, R., Li, B., Zhao, H., Wang, P., He, J.-L., Ma, X.-Y., Wu, Q., Wang, S.-Y., Zhang, H.-P., et al. (2023). Evaluation of diffuse glioma grade and proliferation activity by different diffusion-weighted-imaging models including diffusion kurtosis imaging (DKI) and mean apparent propagator (MAP) MRI. Neuroradiology *65*, 55–64. https://doi.org/10.1007/s00234-022-03000-0.

28. Guo, D., and Jiang, B. (2023). Noninvasively evaluating the grade and IDH mutation status of gliomas by using mono-exponential, bi-exponential diffusion-weighted imaging and three-dimensional pseudo-continuous arterial spin labeling. Eur. J. Radiol. *160*, 110721. https://doi.org/10.1016/j.ejrad.2023.110721.

29. Usuzaki, T., Inamori, R., Shizukuishi, T., Morishita, Y., Takagi, H., Ishikuro, M., Obara, T., and Takase, K. (2024). Predicting isocitrate dehydrogenase status among adult patients with diffuse glioma using patient characteristics, radiomic features, and magnetic resonance imaging: Multi-modal analysis by variable vision transformer. Magn. Reson. Imaging *111*, 266–276. https://doi.org/10.1016/j.mri.2024.05.012.

30. Foltyn-Dumitru, M., Schell, M., Rastogi, A., Sahm, F., Kessler, T., Wick, W., Bendszus, M., Brugnara, G., and Vollmuth, P. (2024). Impact of signal intensity normalization of MRI on the generalizability of radiomic-based prediction of molecular glioma subtypes. Eur. Radiol. *34*, 2782–2790. https://doi.org/10.1007/s00330-023-10034-2.

31. Hu, L.S., Smits, M., Kaufmann, T.J., Knutsson, L., Rapalino, O., Galldiks, N., Sundgren, P.C., and Cha, S. (2025). Advanced Imaging in the Diagnosis and Response Assessment of High-Grade Glioma: AJR Expert Panel Narrative Review. AJR Am. J. Roentgenol. *224*, e2330612. https://doi.org/10.2214/AJR.23.30612.

32. Yu, M., Ge, Y., Wang, Z., Zhang, Y., Hou, X., Chen, H., Chen, X., Ji, N., Li, X., and Shen, H. (2024). The diagnostic efficiency of integration of 2HG MRS and IVIM versus individual parameters for predicting IDH mutation status in gliomas in clinical scenarios: A retrospective study. J. Neuro Oncol. *167*, 305–313. https://doi.org/10.1007/s11060-024-04609-2.

33. Dang, L., White, D.W., Gross, S., Bennett, B.D., Bittinger, M.A., Driggers, E.M., Fantin, V.R., Jang, H.G., Jin, S., Keenan, M.C., et al. (2009). Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature *462*, 739–744. https://doi.org/10.1038/nature08617.

34. Jafari-Khouzani, K., Loebel, F., Bogner, W., Rapalino, O., Gonzalez, G.R., Gerstner, E., Chi, A.S., Batchelor, T.T., Rosen, B.R., Unkelbach, J., et al. (2016). Volumetric relationship between 2-hydroxyglutarate and FLAIR hyperintensity has potential implications for radiotherapy planning of mutant IDH glioma patients. Neuro Oncol. *18*, 1569–1578.

35. Puchalski, R.B., Shah, N., Miller, J., Dalley, R., Nomura, S.R., Yoon, J.-G., Smith, K.A., Lankerovich, M., Bertagnolli, D., Bickley, K., et al. (2018). An anatomic transcriptional atlas of human glioblastoma. Science *360*, 660–663. https://doi.org/10.1126/science.aaf2666.

36. Calabrese, E., Villanueva-Meyer, J.E., Rudie, J.D., Rauschecker, A.M., Baid, U., Bakas, S., Cha, S., Mongan, J.T., and Hess, C.P. (2022). The University of California San Francisco Preoperative Diffuse Glioma MRI Dataset. Radiol. Artif. Intell. *4*, e220058. https://doi.org/10.1148/ryai.220058.

37. Bakas, S., Sako, C., Akbari, H., Bilello, M., Sotiras, A., Shukla, G., Rudie, J.D., Santamaría, N.F., Kazerooni, A.F., Pati, S., et al. (2022). The University of Pennsylvania glioblastoma (UPenn-GBM) cohort: advanced MRI, clinical, genomics, & radiomics. Sci. Data *9*, 453. https://doi.org/10.1038/s41597-022-01560-7.

38. van der Voort, S.R., Incekara, F., Wijnenga, M.M.J., Kapsas, G., Gahrmann, R., Schouten, J.W., Dubbink, H.J., Vincent, A.J.P.E., van den Bent, M.J., French, P.J., et al. (2021). The Erasmus Glioma Database (EGD): Structural MRI scans, WHO 2016 subtypes, and segmentations of 774 patients with glioma. Data Brief *37*, 107191. https://doi.org/10.1016/j.dib.2021.107191.

39. Juvekar, P., Dorent, R., Kögl, F., Torio, E., Barr, C., Rigolo, L., Galvin, C., Jowkar, N., Kazi, A., Haouchine, N., et al. (2024). ReMIND: The Brain Resection Multimodal Imaging Database. Sci. Data *11*, 494. https://doi.org/10.1038/s41597-024-03295-z.

40. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J.S., Freymann, J.B., Farahani, K., and Davatzikos, C. (2017). Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. Sci. Data *4*, 170117. https://doi.org/10.1038/sdata.2017.117.

41. Suter, Y., Knecht, U., Valenzuela, W., Notter, M., Hewer, E., Schucht, P., Wiest, R., and Reyes, M. (2022). The LUMIERE dataset: Longitudinal Glioblastoma MRI with expert RANO evaluation. Sci. Data *9*, 768. https://doi.org/10.1038/s41597-022-01881-7.

42. Hossain, M.F., Alsharif, M.R., and Yamashita, K. (2010). Medical Image Enhancement Based on Nonlinear Technique and Logarithmic Transform Coefficient Histogram Matching (IEEE), pp. 58–62.

43. Carreira, J., and Zisserman, A. (2017). Quo Vadis, Action Recognition? a New Model and the Kinetics Dataset (IEEE), pp. 6299–6308.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
CPTAC-GBM	TCIA	https://www.cancerimagingarchive.net/collection/cptac-gbm/
EGD	Erasmus MC	https://xnat.bmia.nl/REST/projects/egd/
Ivy-GAP	Swedish Institute	https://ivygap.org/
UCSF-PDGM	UCSF	https://www.cancerimagingarchive.net/collection/ucsf-pdgm/
Upenn-GBM	Upenn	https://www.cancerimagingarchive.net/collection/upenn-gbm/
LUMIERE	Bern University	https://doi.org/10.6084/m9.figshare.c.5904905.v1/
ReMIND	TCIA	https://www.cancerimagingarchive.net/collection/remind/
TCGA-GBM	TCIA	https://www.cancerimagingarchive.net/collection/tcga-gbm/
TCGA-LGG	TCIA	https://www.cancerimagingarchive.net/collection/tcga-lgg/
NJDTH	This paper	N/A
JLH	This paper	N/A
Source Code	This paper	https://github.com/YLiu-nju/Res3DNet/
Software and algorithms		
Python (version 3.7.9)	Python software	https://www.python.org/
PyTorch (version 1.13.1)	PyTorch Software	https://pytorch.org/
Matplotlib (version 3.3.2)	Matplotlib Software	https://matplotlib.org/
Monai (version 0.7.0)	Monai Software	https://monai.org.cn/
SimpleITK (version 2.0.2)	SimpleITK Software	https://simpleitk.org/
SPSS (version 26.0)	SPSS Software	https://www.ibm.com/products/spss-statistics

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Study population

We conducted a retrospective analysis by gathering preoperative MRI images from patients with adult diffuse gliomas undergoing surgery or biopsy. The Training-Internal Validation set was collected from one in-house dataset Jinling Hospital (JLH) and five publicly available datasets. Four of the five public datasets were collected from The Cancer Imaging Archive (TCIA): The Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme (CPTAC-GBM) collection, Ivy Glioblastoma Atlas (Ivy-GAP) collection,[35] the University of California San Francisco Preoperative Diffuse Glioma MRI Dataset (UCSF-PDGM) collection,[36] and the University of Pennsylvania glioblastoma (UPenn-GBM) collection.[37] The fifth dataset was the Erasmus Glioma database (EGD) collection.[38] Patients were divided into the Training set and the Internal Validation set with a ratio of 8:2. The External Test set was collected from one in-house dataset Nanjing Drum Tower Hospital (NJDTH) and four publicly available datasets. three of the four public datasets were collected from TCIA: the ReMIND database collection,[39] The Cancer Genome Atlas Glioblastoma Multiforme (TCGA-GBM) collection and the Cancer Genome Atlas Low Grade Glioma (TCGA-LGG) collection.[40] The fourth dataset was the LUMIRER collection.[41] Both the Training-Internal validation and the External Test set ranges across America, Asia and Europe.

We enrolled a total of 2537 patients in our study. 1382 patients (mean age 58.26 ± 14.38, 548 females) were enrolled in the training set. 346 patients (mean age 57.43 ± 14.04, 141 females) were enrolled in the internal validation set and 809 patients (mean age 53.92 ± 14.04, 374 females) were enrolled in the test set.

Patients were enrolled if they were newly diagnosed and pathologically confirmed with a glioma and when preoperative T1-weighted images (T1CE), T2-weighted images (T2WI), fluid attenuated inversion recovery (FLAIR) and contrast-enhanced T1 weighted images (T1CE) were available. No further exclusion criteria were set, such as motion artifacts or degraded image quality.

Ethical statement

This research was approved by the Medical Ethical committee and the Institutional Review Board of Nanjing Drum Tower Hospital (2022-364-02) and Jinlin Hospital (2019NZGKJ-083). This research was registered on Clinical Trial (NCT05624736). The written informed consent was waived due to the retrospective nature of this study. The study was performed in accordance with Declaration of Helsinki.

METHOD DETAILS

Image pre-processing

In order to minimizing overfitting in our model, we implemented standard 3D medical image data augmentation strategies during training stage, including random flipping, random intensity adjustment, random histogram modification, random Gaussian noise injection and intensity standardization. Additionally, we applied 3D histogram matching for nonlinear enhancement.[42] Further details on our specific enhancement methodologies are available in supplemental information.

Deep neural network architecture

To address the challenge of class imbalance in predicting IDH mutation status, we developed a deep neural network architecture based on Residual convolutional blocks (RConv), specifically Res3DNet, illustrated in Figure 2. The Res3DNet architecture represents a development based on the I3D convolutional neural network in the realm of video action classification.[43] Our modifications include replacing the original convolution block with our proposed residual convolution block, replacing the original $7 \times 7 \times 7$ convolution with three $3 \times 3 \times 3$ convolution.

In addition, one approach is applied to tackle the class imbalance problem. Specifically, the hyper-parameter alpha, which is an oversampling method, was employed to address the class imbalance issue in binary classification. The core process includes: (i) identifying the number of samples in the minority class (which can be done manually or automatically); (ii) repeating the minority class samples α times to achieve training set balance. Unlike traditional oversampling techniques, we dynamically adjust the number of repetitions through alpha. This process is only applied to the training set, while the test set remains unchanged.

All details of the structure of our model (Figure S1) and the using of hyper-parameter alpha are given in supplemental information (Figures S2 and S3).

Radiologist-AI competing test

TCGA-GBM and TCGA-LGG datasets were designated as test set for Radiologist-AI competing because these two datasets were most wildly used in machine learning based Glioma imaging genotyping research. Four board-certified neuroradiologists independently reviewed the MRI scans and predicted the IDH mutation status of each patient. Two of the readers had 2 years of experience in neuro-oncology imaging interpretation, while the other two had 5 years of experience. During the assessment, the radiologists had access to all MRI modalities, including T1WI, T2WI, FLAIR, and T1CE. They were blinded to all clinical information, including patient age, sex, and histopathological results, to ensure unbiased evaluation. TCGA test set does not overlap with training data.

To assess model generalizability, a subgroup analysis was performed on the external test set based on MRI scanner vendors (Siemens, General Electric, Philips, and United Imaging). Performance metrics were calculated separately for each vendor using the Res3DNet model.

To compare the performance of the Res3DNet model and the T2-FLAIR mismatch sign in predicting IDH mutation status in non-enhancing gliomas, a subgroup analysis was conducted on the non-enhancing cases in the external test set. The presence of the T2-FLAIR mismatch sign was assessed by two neuroradiologists in consensus.

QUANTIFICATION AND STATISTICAL ANALYSIS

The performance of the deep learning models was evaluated by assessing the Aera under the receiver operative characteristic curve (AUC), area under the precision-recall curve (AUPRC), Accuracy, Recall, Precision, Specificity and F1 score.

Student's *t* test was performed to determine the differences of age distribution between the training cohort and internal validation cohort. Patients' age was reported as mean ± standard deviation. Chi-squared test was performed to evaluate the differences in sex, IDH mutation status between training-internal validation set and external validation set. The Friedman test and DeLong test were also undertaken to assess the AUC of multiple deep learning models. A p-value <0.05 was considered statistically significant in this research. Statistical analysis was conducted using SPSS software (version 26.0; IBM).