



OPEN ACCESS

EDITED BY

Michael Albert Thomas,
University of California, Los Angeles,
United States

REVIEWED BY

Ann-Christin Hau,
Laboratoire National de Santé (LNS),
Luxembourg
Kumar Pichumani,
Houston Methodist Research Institute,
United States

*CORRESPONDENCE

Yun Liu

✉ liuyun@njmu.edu.cn

Yongping You

✉ yypl9@njmu.edu.cn

Yun Yu

✉ yuyun@njmu.edu.cn

[†]These authors have contributed
equally to this work and share
first authorship

RECEIVED 13 January 2023

ACCEPTED 17 August 2023

PUBLISHED 30 August 2023

CITATION

Zhang H, Fan X, Zhang J, Wei Z, Feng W,
Hu Y, Ni J, Yao F, Zhou G, Wan C, Zhang X,
Wang J, Liu Y, You Y and Yu Y (2023)

Deep-learning and conventional
radiomics to predict *IDH* genotyping
status based on magnetic resonance
imaging data in adult diffuse glioma.

Front. Oncol. 13:1143688.

doi: 10.3389/fonc.2023.1143688

COPYRIGHT

© 2023 Zhang, Fan, Zhang, Wei, Feng, Hu,
Ni, Yao, Zhou, Wan, Zhang, Wang, Liu, You
and Yu. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). 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.

Deep-learning and conventional radiomics to predict *IDH* genotyping status based on magnetic resonance imaging data in adult diffuse glioma

Hongjian Zhang^{1†}, Xiao Fan^{2†}, Junxia Zhang², Zhiyuan Wei¹,
Wei Feng¹, Yifang Hu³, Jiaying Ni¹, Fushen Yao¹,
Gaoxin Zhou^{1,4}, Cheng Wan^{1,4}, Xin Zhang^{1,4}, Junjie Wang^{1,4},
Yun Liu^{1,4*}, Yongping You^{2*} and Yun Yu^{1,4*}

¹Department of Medical Informatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, Jiangsu, China, ²Department of Neurosurgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China, ³Department of Geriatric Endocrinology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China, ⁴Institute of Medical Informatics and Management, Nanjing Medical University, Nanjing, Jiangsu, China

Objectives: In adult diffuse glioma, preoperative detection of isocitrate dehydrogenase (*IDH*) status helps clinicians develop surgical strategies and evaluate patient prognosis. Here, we aim to identify an optimal machine-learning model for prediction of *IDH* genotyping by combining deep-learning (DL) signatures and conventional radiomics (CR) features as model predictors.

Methods: In this study, a total of 486 patients with adult diffuse gliomas were retrospectively collected from our medical center (n=268) and the public database (TCGA, n=218). All included patients were randomly divided into the training and validation sets by using nested 10-fold cross-validation. A total of 6,736 CR features were extracted from four MRI modalities in each patient, namely T1WI, T1CE, T2WI, and FLAIR. The LASSO algorithm was performed for CR feature selection. In each MRI modality, we applied a CNN+LSTM-based neural network to extract DL features and integrate these features into a DL signature after the fully connected layer with sigmoid activation. Eight classic machine-learning models were analyzed and compared in terms of their prediction performance and stability in *IDH* genotyping by combining the LASSO-selected CR features and integrated DL signatures as model predictors. In the validation sets, the prediction performance was evaluated by using accuracy and the area under the curve (AUC) of the receiver operating characteristics, while the model stability was analyzed by using the relative standard deviation of the AUC (RSD_{AUC}). Subgroup analyses of DL signatures and CR features were also individually conducted to explore their independent prediction values.

Results: Logistic regression (LR) achieved favorable prediction performance (AUC: 0.920 ± 0.043, accuracy: 0.843 ± 0.044), whereas support vector machine with the linear kernel (l-SVM) displayed low prediction performance

(AUC: 0.812 ± 0.052 , accuracy: 0.821 ± 0.050). With regard to stability, LR also showed high robustness against data perturbation (RSD_{AUC} : 4.7%). Subgroup analyses showed that DL signatures outperformed CR features (DL, AUC: 0.915 ± 0.054 , accuracy: 0.835 ± 0.061 , RSD_{AUC} : 5.9%; CR, AUC: 0.830 ± 0.066 , accuracy: 0.771 ± 0.051 , RSD_{AUC} : 8.0%), while DL and DL+CR achieved similar prediction results.

Conclusion: In *IDH* genotyping, LR is a promising machine-learning classification model. Compared with CR features, DL signatures exhibit markedly superior prediction values and discriminative capability.

KEYWORDS

IDH, glioma, deep-learning, radiomics, magnetic resonance imaging

1 Introduction

Adult diffuse gliomas are a group of primary malignant brain tumors and have a relatively high mortality rate (1). Despite the availability of a diverse array of treatments including tumor resection, radiotherapy, chemotherapy, and experimental targeted therapy, the prognosis for patients remains generally unfavorable (2, 3). According to the 2007 World Health Organization Central Nervous System (2007 WHO CNS) tumor classification (4), adult diffuse gliomas are classified based on tumor histology. However, a growing number of studies have shown that adult diffuse gliomas with different histological classifications may have similar biological behaviors and prognosis because of the same genetic changes (5). Therefore, the two newest guidelines from the WHO CNS and European Association of Neuro-Oncology (EANO), both published in 2021, underscore the significance of incorporating molecular biomarkers with both clinical and pathological values into the precision classification of adult diffuse gliomas, to promote the development of tumor precision treatment (6, 7).

One of the important molecular biomarkers for adult diffuse gliomas is the expression status of isocitrate dehydrogenase (*IDH*), which is now routinely incorporated into the clinical management of patients (7). According to the 2021 EANO guideline for adult diffuse gliomas, the presence of an *IDH* mutation can be diagnosed as an astrocytoma with WHO grade 2–4 or an oligodendroglioma with WHO grade 2–3. Conversely, in most cases, the absence of an *IDH* mutation indicates the diagnosis of a glioblastoma with WHO grade 4. Some studies have shown that adult patients with *IDH*-wildtype glioblastoma undergoing standard treatment typically have an average overall survival of 15–18 months (8), while the average overall survival can extend up to 14.7 years in cases of oligodendroglioma with *IDH* mutation and *1p/19q*-codeletion, but with appropriate treatment (9, 10). Furthermore, the presence of *IDH* mutation has emerged as a specific treatment target, leading to its exploration in various clinical trials involving peptide vaccination and small-molecule inhibitor approaches (11, 12). Therefore, the *IDH* expression status is highly relevant to the

glioma prognosis and keeps high clinical value for the classification of adult diffuse gliomas.

In clinical practice, *IDH* genotyping is carried out on biopsied tumor samples. The limitation of this approach is the relatively long detection period, invasiveness, and the sampling difficulty from certain brain areas. Conversely, magnetic resonance imaging (MRI) has been considered the most promising candidate to aid decision-making in clinical practice due to its non-invasive nature, fast and global detection ability, and high resolution for soft tissues (13, 14). Taking full advantage of the wealth of information obtained from preoperative MRIs could assist in filling the knowledge gaps in local tumor biopsy. Thus, MRI-based *IDH* genotyping appears to be an important preoperative assessment method to help guide the treatment and prognosis of glioma patients.

Radiomics (15), an imaging analysis method, advocates that the high-throughput quantitative handcrafted features extracted from medical images can be utilized to build machine-learning models to enable the preoperative evaluation of tumors. Several studies have investigated the potential of MRI-based radiomics analysis to noninvasively facilitate tumor grading, molecular subtyping, and prognosis evaluation in gliomas (16–19). However, in some aspects, conventional radiomics (CR) involves rigorous and complex analysis to inevitably require extracting and selecting handcrafted features that could introduce additional errors because of feature calculations (20). Furthermore, the limited and handcrafted radiomics features cannot adequately reflect tumor heterogeneities, which could limit the prediction abilities of machine-learning models (21).

Nowadays, advances in computational technology have promoted the development of deep-learning (DL) that has been widely used in tumor preoperative evaluation given its end-to-end prediction advantage and ability to simplify the analysis process (22, 23). In contrast to CR, deep neural networks possess remarkable representation capabilities, enabling the extraction of high-throughput discriminative features that can directly capture abundant tumor information. These deep neural networks will eliminate the need for additional feature extraction and selection

operations, simplifying the process while still yielding valuable insights. DL features can further reveal tumor heterogeneity and stimulate the prediction potential of machine-learning models. In glioma *IDH* genotyping, machine-learning models only based on CR features have been widely applied and analyzed (24, 25). However, little work has been done on evaluating the effectiveness of different machine-learning models by combining DL signatures and CR features as model predictors.

In the present study, we aimed to determine the optimal machine-learning model by making full use of the DL signatures and CR features. We built an image sequence model to extract DL signatures. Based on the extracted DL signatures and CR features from four conventional MRI modalities used in most medical centers, we evaluated and compared eight classic machine-learning models in terms of their stability and prediction performance in *IDH* genotyping. Furthermore, the subgroup analyses of DL signatures and CR features were also individually conducted to explore their independent prediction values.

2 Materials and methods

2.1 Patient enrollment

All the cases used in this retrospective study were de-identified and obtained from the public datasets (TCGA-GBM, TCGA-LGG), and our local medical center (The First Affiliated Hospital of Nanjing Medical University) between May 2015 and July 2020. Institutional Review Board approval from our medical center was obtained, but this was not required for the public datasets. Patients who met the following criteria were included in this study: (1) pathological diagnosis of primary diffuse glioma; (2) age ≥ 18 years; (3) known *IDH* status (detected by immunohistochemistry or Sanger sequencing); (4) no history of preoperative therapy, biopsy, or any treatment; and (5) four available preoperative MRI modalities, including T1-weighted, T2-weighted, gadolinium contrast-enhanced T1-weighted, and T2-weighted fluid-attenuated inversion recovery images (T1WI, T2WI, T1CE, and FLAIR, respectively).

2.2 MR imaging

The MR images are heterogeneous, because they are acquired using either 1.5T or 3.0T MRI scanners according to the different imaging protocols at each institution. The MRI scanners were provided by different MR vendors, including Philips, General Electric, and Siemens. The preoperative MR imaging protocols include the acquisition of T1WI (parameters vary from TR: 300–2000 ms, TE: 5–25 ms, FOV: 60–100 mm, slice thickness: 1.5–7.5 mm, and matrix: 256×256 or 512×512); T2WI (parameters vary from TR: 2000–10000 ms, TE: 80–150 ms, FOV: 70–110 mm, slice thickness: 1.5–8 mm, and matrix: 256×256 or 512×512); T1CE (parameters vary from TR: 200–1100 ms, TE: 4–20 ms, FOV: 70–100 mm, slice thickness: 1–8.5 mm, and matrix: 256×256 or 512×512); and FLAIR (parameters vary from TR: 5000–11000 ms,

TE: 80–200 ms, FOV: 70–110 mm, slice thickness: 2–7 mm, and matrix: 256×256 or 512×512). All MR imaging data including pixel matrixes and metadata were saved in the digital imaging and communications in medicine (DICOM) format.

2.3 Image preprocessing and tumor segmentation

We converted MR images with DICOM format into the neuroimaging informatics technology initiative (NIFTI) format by using the python package SimpleITK. Referring to the image preprocessing in the competition of brain tumor segmentation (BraTS2021) (<https://www.med.upenn.edu/cbica/brats2021>), images were co-registered to the same anatomical template (SRI24) (26), interpolated to a uniform isotropic resolution (1mm³) and skull-stripped by using FSL software (<https://fsl.fmrib.ox.ac.uk/fsl>).

Tumor segmentation is a crucial step for the following feature extraction and quantitative analysis. As is known in the BraTS competition, we can split gliomas into two subregions: tumor core (TC, comprising a contrast-enhancing area and necrotic portions, if any) and the whole tumor (WT, combining the tumor core and edema). TC describes the bulk of the tumor, which is what is typically resected, while WT describes the complete extent of the disease. In our study, we used the fully automated nnU-Net segmentation framework based on a convolutional neural network (CNN) to segment these two tumor subregions. The nnU-Net framework is the first plug-and-play tool for biomedical image segmentation (<https://github.com/MIC-DKFZ/nnUNet>). It has been widely validated in the BraTS competition and achieved superior performance in brain tumor segmentation (27). Inexperienced users can use nnU-Net out of the box for their custom 3D segmentation problem without the need for manual intervention.

2.4 DL signature extraction

Convolutional neural networks and recurrent neural networks (RNN) are different types of artificial neural networks that can perform representational learning on imaging data and provide different hierarchical feature representations at each network layer (28). It is precisely the stacking employment of multiple network layers with non-linear activation functions that make the feature representation complex and diverse. After passing through a series of chained CNN or RNN layers, posterior probability can be calculated from the representational features and used as a predictor for tumor preoperative evaluation. In our study, we combined CNN-based *eca_nfnet_l0* (29) and RNN-based long-short-term memory network (LSTM) (30) together as our DL model, extracting posterior probability as the DL signature for *IDH* genotyping (Figure 1).

Data preparation needs to be conducted before the imaging data is imported into the DL model. To avoid heterogeneity bias, various image signal intensities were transformed into standardized intensity ranges via *z*-score normalization ($Z = (x-\mu)/\sigma$), where μ

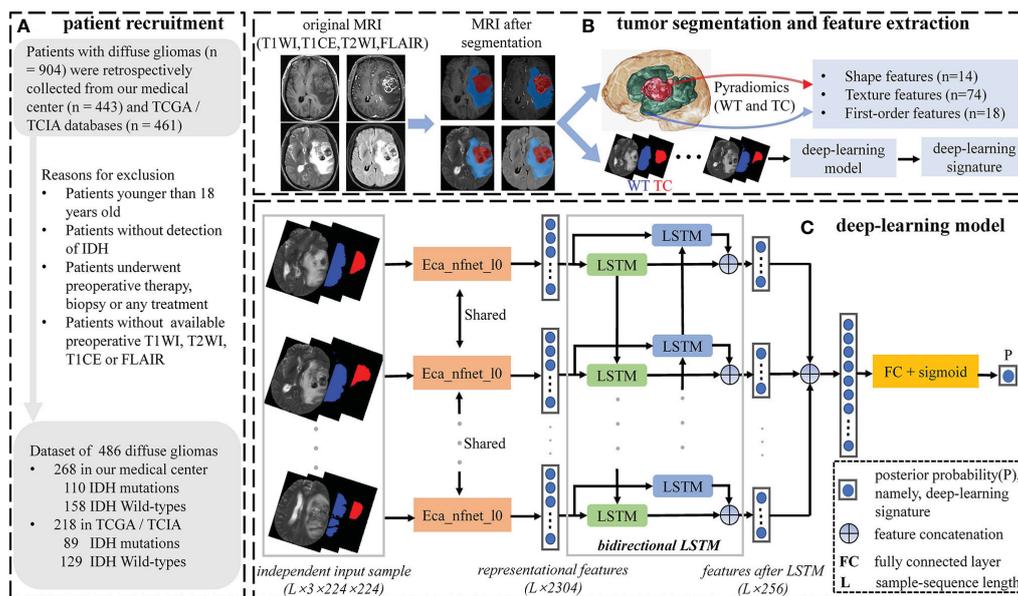


FIGURE 1

Workflow of patient recruitment, tumor segmentation and feature selection. (A) Patient recruitment process. (B) Tumor segmentation and feature extraction process. Tumor segmentation was performed with T1WI, T1CE, T2WI, and FLAIR images by using nnU-Net automated segmentation framework. Conventional radiomics features were extracted from the WT and TC volumes under every single-modality MRI, respectively. (C) The deep-learning model. Single-modality DL signature was extracted using the deep-learning model.

and σ are the mean and standard deviation of pixel values, respectively). Typically, glioma volumes can reflect abundant and complex tumor characteristics under different spatial dimensions. To utilize abundant tumor information from three spatial dimensions, we extracted the axial, coronal, and sagittal tumor slices from each 3D image. Referring to the segmented WT mask, we selected slices with the largest tumor area from each spatial dimension, as well as slices with tumor regions of 50th, 55th, 60th, 65th, 70th, 75th, 80th, 85th, 90th, and 95th percentiles. Subsequently, we integrated each selected slice with its two corresponding mask slices (WT and TC) to create a new 3-channel image. A total of 33 representative 3-channel images were created from each 3D image. Unlike conventional 2D-CNN models, our model takes the correlation between different tumor slices as well as the hidden tumor spatial information into account. When it comes to extracting features using LSTM, we need to feed an image sequence representing an independent input sample into the model. Thereupon, all 3-channel images needed to be arranged according to their respective slice ordinal numbers and integrated into an image sequence based on the order of axial, coronal, and sagittal sequences. Finally, in each 3D image, we obtained the independent input samples with a total number of $((33 - \text{length}) / \text{stride} + 1)$ from the 3-channel image sequence with a total length of 33 by specifying the sample-sequence length and moving stride (Supplementary Figure S1). We also performed random flip, rotation, and translation on the independent samples to enhance the robustness of the DL model, as well as resampled to $224 \times 224 \times 3$ (determined by the pretrained ImageNet dataset (31)).

Eca_nfnets_10 is a variant of the nfnets (normalization free net) model family, whose prediction performance achieved on the ImageNet dataset is better than that of the residual networks

(ResNet) (31). We utilized eca_nfnets_10 to obtain the representational features from each 3-channel image of the independent input sample. Bidirectional LSTM was then used to learn the intrinsic and mutual relationships of the representational features from different 3-channel images. Finally, the output features of LSTM were ensembled into a class posterior probability (namely, DL signature) through a fully connected layer with a sigmoid activation function. A single patient will have multiple independent input samples under every single-modality MRI. We considered the average of all posterior probabilities from different input samples as the patient's final DL signature. A total of four DL signatures were independently extracted from four single-modality MRIs.

2.5 Conventional radiomics feature extraction

To reduce heterogeneity bias between patients, MR images were normalized via z-score normalization for subsequent CR feature extraction, which was based entirely on an open-source python package pyradiomics that was established to provide a reference standard according to the image biomarker standardization initiative (IBSI) for radiomics analysis of medical imaging (32). Based on the segmented sub-volumes (WT and TC), we extracted shape, first-order, and texture features from original and derived images (wavelet decompositions via directional low-pass and high-pass filtering to yield eight derived images on each original 3D MR image).

Shape features describe the size and shape of the region of interest (ROI). These features are independent of the gray-level

intensity distribution in the ROI and are only calculated on 3D mask. First-order features take the properties of individual pixel values ignoring the spatial interaction between image pixels into account. While texture features describe this spatial interaction between every image pixel and their surrounding neighborhoods. Texture features can be extracted using five methods, including the gray-level co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM), gray-level size zone matrix (GLSZM), neighborhood gray-tone difference matrix (NGTDM), and gray-level dependence matrix (GLDM).

2.6 Feature selection and model development

2.6.1 Nested-CV

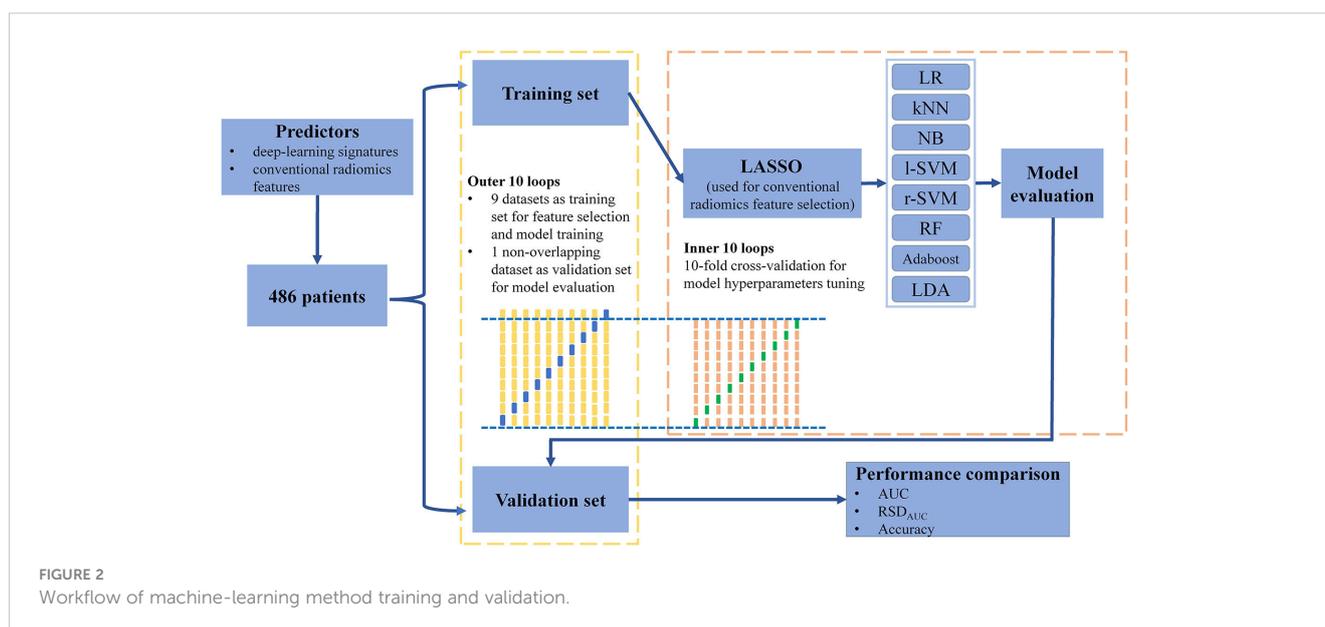
Nested cross-validation (nested-CV) is a search for hyperparameters by estimating the generalization error of the training model to obtain the best hyperparameters. It consists of the outer loops and inner loops. The inner loop refers to the cross-validation with the ability to search for the best hyperparameters to provide the best hyperparameters for the training model validated in the outer loop. The outer loop provides training data to the inner loop while retaining some extra data to validate the model trained in the inner loop. Compared to simple-CV, nested-CV can prevent information leakage of data to obtain a relatively low model scoring bias, especially in relatively small datasets. Nested-CV is also successfully employed in the machine-learning analysis of neuroimaging (33). In the present study, we utilized nested-10-fold-CV to perform feature selection and model hyperparameters tuning, in which nine non-overlapping datasets of each outer loop were trained in its inner loop and the remaining one non-overlapping dataset was the validation set of this outer loop (Figure 2 and Supplementary Table S1).

2.6.2 Conventional radiomics feature selection

When building a machine-learning classifier involving high-throughput features, feature selection provides a crucial step to reduce the risk of over-fitting, improve accuracy, and decrease training time. Before performing CR feature selection, all CR features are subjected to z-score normalization to eliminate the influence of different feature magnitudes. We utilized the least absolute shrinkage and selection operator penalty (LASSO) algorithm with L1 regularization to select CR features (34). A hyperparameter, designated as α , controls the extent of L1 regularization: the larger the value of α , the fewer the selected CR features. The optimal hyperparameter α , ranging from 0 to 0.1 in 0.05 increments, was selected by using the inner 10-fold cross-validation in the nine non-overlapping datasets of each outer loop (33). We then retained the CR features with a non-zero coefficient resulting from the optimal α for further analyses. In this way, 10 specific CR feature subsets were selected by using the LASSO model with optimal α .

2.6.3 Classifier building

After the CR feature selection, we utilized the mixture of DL signatures and CR features as predictors to construct machine-learning classifiers. Eight classical machine-learning classifiers were built and compared: logistic regression (LR), k-nearest neighbors (kNN), naive Bayes (NB), support vector machines with the linear kernel (l-SVM), support vector machines with radial basis function kernel (r-SVM), random forest (RF), adaptive boosting (Adaboost) and linear discriminant analysis (LDA) (Supplementary Table S2). In each outer loop, we tuned every classifier by using the inner 10-fold cross-validation and compared the area under the curve (AUC) value of the receiver operating characteristics to identify the classifier with optimal hyperparameters. Random grid searches were used for all hyperparameter tuning processes. The average prediction performance of classifiers was estimated in the validation sets of the outer loops by quantifying the accuracy and AUC values. Meanwhile, the stability of the machine-learning classifier was



quantified by using the relative standard deviation of the AUC value (RSD_{AUC}). $RSD_{AUC}\%$ is defined as:

$$RSD_{AUC}\% = \frac{\sigma_{AUC}}{\mu_{AUC}} \times 100\%$$

Where, σ_{AUC} and μ_{AUC} are the standard deviation and mean of the AUC value, respectively. A lower $RSD_{AUC}\%$ value represents the higher stability of the machine-learning classifier.

2.7 Implementation details

All the calculations and modeling were based on the pytorch, sklearn, and pycaret libraries in python as the backend. When building the DL model, we created consecutive image sequences as

the independent input samples by setting the sample-sequence length to 11 and moving stride to 2. In this way, we could obtain 12 independent samples from single patient under every single-modality data. We used the sigmoid linear unit activation function in each hidden layer and the binary cross-entropy as the objective function. The weights of the network were optimized via a gradient descent algorithm with a mini-batch size of 32 and a learning rate of 1e-6.

3 Results

3.1 Patient and tumor characteristics

Overall, we obtained 486 cases and the patient and tumor characteristics were statistically analyzed and shown in [Table 1](#).

TABLE 1 Patient and tumor characteristics of the entire cohort.

Parameters	Entire cohort	Public dataset		P_{intra}	Local population		P_{intra}	P_{inter}
		IDH expression status			IDH expression status			
		Mutation	Wildtype		Mutation	Wildtype		
Total Number	486	89	129		110	158		
Sex				0.24			0.79	0.32
Male	274(56%)	43(48%)	74(57%)		66(60%)	91(58%)		
Female	212(44%)	46(52%)	55(43%)		44(40%)	67(42%)		
Grade				<0.05			<0.05	0.83
LGG(WHO 2 or 3)	240(49%)	84(94%)	22(17%)		93(85%)	41(26%)		
HGG(WHO 4)	246(51%)	5(6%)	107(83%)		17(15%)	117(74%)		
Age (years)	51.7 ± 14.2	43.9 ± 13.9	58.3 ± 13.0	<0.05	44.3 ± 11.7	55.8 ± 12.6	<0.05	1.01
18 - 51	232(48%), 39.6 ± 8.8	62(70%), 36.6 ± 8.9	34(26%), 41.8 ± 9.0		82(75%), 39.2 ± 8.4	54(34%), 42.1 ± 8.1		
≥ 52	254(52%), 62.8 ± 7.8	27(30%), 60.7 ± 6.9	95(74%), 64.2 ± 8.3		28(25%), 59.3 ± 5.1	104(66%), 62.9 ± 7.7		
Histology				<0.05			<0.05	0.57
Astrocytoma	118(24%)	22(25%)	10(8%)		53(48%)	33(21%)		
Oligodendroglioma	92(19%)	40(45%)	8(6%)		39(36%)	5(3%)		
Oligoastrocytoma	39(8%)	22(25%)	4(3%)		10(9%)	3(2%)		
Glioblastoma	237(49%)	5(5%)	107(83%)		8(7%)	117(74%)		
1p/19q				<0.05			<0.05	0.92
Codeletion	66(14%)	21(24%)	0(0%)		37(34%)	8(5%)		
Nono-codeletion	240(49%)	40(45%)	22(17%)		58(53%)	120(76%)		
Unknown	180(37%)	28(31%)	107(83%)		15(13%)	30(19%)		
MGMT				0.09			0.35	0.14
Methylation	110(22%)	5(6%)	27(21%)		32(29%)	46(29%)		
Unmethylation	139(29%)	0(0%)	28(22%)		37(34%)	74(47%)		
Unknown	237(49%)	84(94%)	74(57%)		41(37%)	38(24%)		

Data is presented as the mean ± standard deviation or as numbers (with percentage in parentheses). Age is divided into two groups based on its mean value. P-values are calculated from unpaired t-test for continuous variables and chi-square test for categorical variables, to compare statistical differences of variables between IDH mutation and IDH wildtype (P_{intra}) as well as between public and local cohorts (P_{inter}). A $P < 0.05$ is considered to indicate statistically significant differences.

The workflow of the current study is displayed in Figures 1 and 2. The proportions of patients with *IDH* mutation and wildtype were 41% (199/486) and 59% (287/486), respectively. There was no significant difference in sex between patients with *IDH* mutation and wildtype ($p > 0.05$). The proportion of patients presenting with LGG was 89% (177/199) in *IDH* mutation patients, while the proportion of patients presenting with HGG was 78% (224/287) in *IDH* wildtype patients ($p < 0.05$). The mean (\pm standard deviation) age was 51.7 ± 14.2 years and its P -value was < 0.05 between patients with *IDH* mutation and *IDH* wildtype.

3.2 Conventional radiomics features and DL signatures

In each MR modality, we extracted 842 CR features from each tumor subregion, respectively, including 18×9 first-order, 23×9 GLCM, 16×9 GLRLM, 16×9 GLSZM, 5×9 NGTDM, 14×9 GLDM, and 14 shape features. A total of 6,736 CR features were extracted from the MRI data of every patient. After LASSO, 10 specific CR feature subsets were obtained, whose numbers ranged from 5 to 9 (Supplementary Table S3). Features that were selected in at least five of the 10 loops were considered the most valuable and stable CR features (33). Finally, we obtained seven valuable CR features, comprising texture, first-order, and shape features extracted from the original images and wavelet-transformed images. These seven CR features together with four DL signatures were then compared with the *IDH* mutation status by using the unpaired t -test, revealing that all the selected features were significantly different between patients with *IDH* mutation and wildtype ($p < 0.05$) (Supplementary Table S4).

3.3 Model performance

3.3.1 Classifiers using DL+CR features as predictors

We analyzed the prediction performance and stability of the eight classic machine-learning classifiers by using the DL+CR features as predictors (Table 2). In terms of the prediction

TABLE 2 Validation performance based on DL+CR features.

Classifier	Accuracy	AUC	RSD _{AUC}
LR	0.843 \pm 0.044	0.920 \pm 0.043	4.7%
kNN	0.848 \pm 0.065	0.904 \pm 0.081	9.0%
NB	0.850 \pm 0.052	0.916 \pm 0.063	6.9%
l-SVM	0.821 \pm 0.050	0.812 \pm 0.052	6.4%
r-SVM	0.856 \pm 0.056	0.911 \pm 0.073	8.0%
RF	0.842 \pm 0.051	0.917 \pm 0.052	5.7%
Adaboost	0.827 \pm 0.050	0.903 \pm 0.044	4.9%
LDA	0.839 \pm 0.048	0.918 \pm 0.045	4.9%
Average	0.841 \pm 0.051	0.900 \pm 0.065	7.2%

Numbers in bold font represent the best performance among different models.

performance in all classifiers, LR had the best AUC (0.920 ± 0.043), r-SVM had the best accuracy (0.856 ± 0.056), while l-SVM had the lowest prediction performance with AUC (0.812 ± 0.052) and accuracy (0.821 ± 0.050). In terms of stability, the most stable classifier was LR (RSD_{AUC}: 4.7%), followed by LDA (RSD_{AUC}: 4.9%) and Adaboost (RSD_{AUC}: 4.9%). kNN (RSD_{AUC}: 9.0%) and r-SVM (RSD_{AUC}: 8.0%) had the lowest stability among all the classifiers. We mainly referred to the AUC value to compare the prediction performance between different classifiers. Overall, LR with the best AUC and stability (AUC: 0.920 ± 0.043 , RSD_{AUC}: 4.7%) outperformed other machine-learning classifiers in the *IDH* genotyping prediction.

3.3.2 The optimal classifier using different feature subcategories as predictors

The DL signatures and CR features were individually analyzed to investigate their respective prediction potential in *IDH* genotyping. The average prediction performance of the DL signatures among the eight classifiers (AUC: 0.900 ± 0.058 , accuracy: 0.835 ± 0.056 , RSD_{AUC}: 6.4%) was better than that of the CR features (AUC: 0.799 ± 0.095 , accuracy: 0.754 ± 0.072 , RSD_{AUC}: 11.9%). In contrast, the average prediction performance of the DL+CR features had little improvement (AUC: 0.900 ± 0.065 , accuracy: 0.841 ± 0.051 , RSD_{AUC}: 7.2%) (Table 3 and Supplementary Table S5).

Since LR had the best prediction performance and stability in DL+CR, we additionally utilized it to compare the prediction potential of the different feature subcategories. As illustrated in Figure 3, LR in the DL achieved the favorable prediction results and stability (AUC: 0.915 ± 0.054 , accuracy: 0.835 ± 0.061 , RSD_{AUC}: 5.9%). In contrast, LR in the CR achieved the lowest prediction results and stability (AUC: 0.830 ± 0.066 , accuracy: 0.771 ± 0.051 , RSD_{AUC}: 8.0%). After combining the DL signatures and CR features, we found that the CR features did not obviously help to improve the prediction performance of LR (AUC: 0.920 ± 0.043 , accuracy: 0.843 ± 0.044 , RSD_{AUC}: 4.7%). Furthermore, we analyzed the prediction potential of single-modality DL signature in *IDH* genotyping by ROC analysis. The T1CE signature had the best AUC (AUC: 0.904 ± 0.044 , RSD_{AUC}: 4.9%), followed by the T2WI signature (AUC: 0.899 ± 0.062 , RSD_{AUC}: 6.9%), T1WI signature (AUC: 0.880 ± 0.062 , RSD_{AUC}: 7.0%) and FLAIR signature (AUC: 0.877 ± 0.088 , RSD_{AUC}: 10.0%). The t -SNE visualization of different feature subcategories are also shown in Figure 4, illustrating that the DL signatures possess better discriminative ability than the CR features.

TABLE 3 Average validation performance of different feature subcategories.

Subcategory	Accuracy	AUC	RSD _{AUC}
CR	0.754 \pm 0.072	0.799 \pm 0.095	11.9%
DL	0.835 \pm 0.056	0.900 \pm 0.058	6.4%
DL+CR	0.841 \pm 0.051	0.900 \pm 0.065	7.2%

Numbers in bold font represent the best performance among different models.

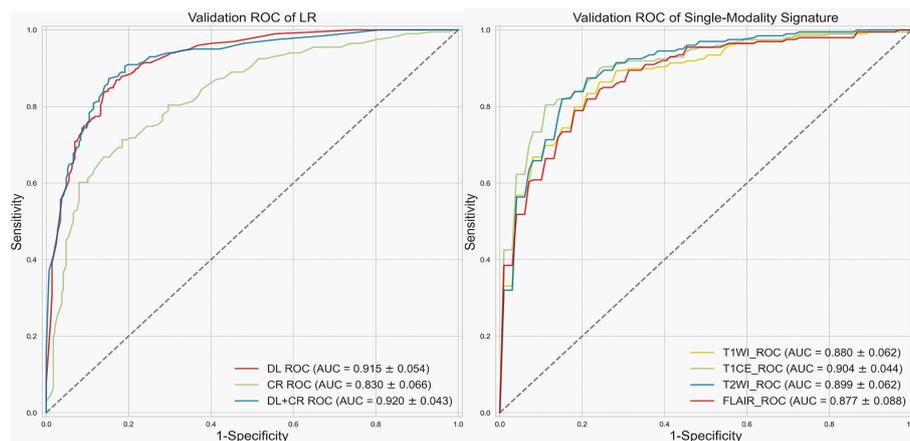


FIGURE 3

Receiver operating characteristic (ROC) curve. Left, validation ROC curve of LR by using different feature subcategories as predictors. Right, validation ROC curve for single-modality signature.

4 Discussion

According to the newest WHO classification of CNS tumors, adult diffuse gliomas were classified not only by pathological characteristics but also by genotyping, which was the important prognostic factor affecting patients' survival. In addition to *IDH* genotyping, genetic phenotypes and molecular characteristics such as chromosome *1p/19q* co-deletion, O-6-methylguanine-DNA methyltransferase (*MGMT*) methylation, and phosphatase and tensin homologue deleted on chromosome 10 (*PTEN*) genotyping also have important effects on the prognosis and treatment of gliomas (14). Tan et al. (35) used a radiomics nomograph to predict *IDH* genotyping (AUC: 0.900), confirming that most WHO LGG patients presented with *IDH* mutation and had better prognosis. Kanazawa et al. (36) found that ADC kurtosis (AUC: 0.728) and T2 kurtosis (AUC: 0.866) had the highest correlation with *1p/19q* codeletion through the radiomics texture analysis. In *MGMT* prediction, Li et al. (37) compared two different feature selection methods, among which the all-relevant features have the potential of offering better prediction power than the univariately-predictive and non-redundant features (AUC: 0.880). Based on multicenter and multimodal MRIs, Li et al. (38) illustrated that the radiomics features derived from T2WI were more correlated with *PTEN* genotyping (AUC: 0.787). Results from these studies suggest that radiomics analysis is indeed a powerful method for predicting glioma genotyping before surgery. Multimodal MRIs could further improve the prediction performance of glioma genotyping.

Multimodal MRI technology can obtain a variety of tumor information including tumor morphology, blood perfusion, and metabolism that can help further evaluate tumor prognosis and therapeutic effects. Several studies have explored the molecular-biomarkers-based classification of glioma subtypes using PET, DWI, DCE, and DSC-PWI, which are non-standard imaging modalities used to gather additional tumor information. Song et al. (39) demonstrated that both PET and DSC-PWI might be non-invasive predictors for *IDH* genotyping, in which PET

combined with CBV could improve the differentiation of *IDH*-mutant astrocytoma and *IDH*-wildtype glioblastoma (AUC: 0.903). Kim et al. (40) found that DWI and PWI can improve the diagnostic performance of *IDH* genotyping (AUC: 0.747) to further guide the LGG glioma subtyping, with DWI-ADC features playing a significant role. Furthermore, Yan et al. (17) showed that multimodal-MRIs-based radiomics may be useful for noninvasive detection of molecular groups and guiding glioma subtyping. The image fusion model (multivariate logistic regression) incorporating radiomic features from T1CE and DWI-ADC achieved an AUC of 0.884 and 0.669 for predicting *IDH* and *TERT* status, respectively. Pei et al. (41) have also investigated the integration of multimodal MRIs to improve the accuracy of glioma subclassification. Adding DSC-PWI to conventional MRIs can improve glioma subtype prediction in patients with diffuse gliomas (AUC: 0.864, 0.787, and 0.816 in *IDH* wildtype, *IDH* mutant and *1p/19q*-noncodeleted, and *IDH* mutant and *1p/19q*-codeleted, respectively). In summary, the utilization of non-standard imaging techniques and machine-learning in glioma subtyping has exhibited encouraging outcomes. As the research in this area advances, it is essential to emphasize the reproducibility and generalizability of these methods to facilitate their potential incorporation into routine clinical practice. Furthermore, additional comprehensive studies are necessary to explore the effects of integrating imaging features with multimodal imaging data, with the aim of improving the accuracy and comprehensiveness of glioma subtyping.

IDH mutant tumors can produce oncometabolite 2-hydroxyglutarate (2HG) which can be non-invasively detected by *in vivo* MR spectroscopy (MRS). Studies have shown that elevated levels of 2HG can be detected in *IDH* mutant tumors using MRS, allowing for non-invasive assessment and confirmation of *IDH* mutation status (42, 43). Furthermore, the quantification of 2HG levels through MRS has demonstrated prognostic significance. Higher levels of 2HG in *IDH* mutant tumors have been associated with better treatment response and improved overall survival outcomes (44). Currently, radiomics methods and 2HG MRS are

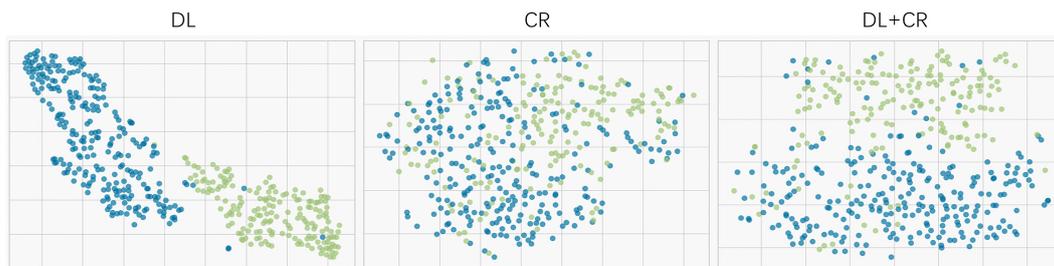


FIGURE 4

t-SNE visualization for different feature subcategories (the DL, CR and DL+CR). Every dot represents a patient. Blue represents the patients with *IDH* wildtype, whereas green represents the patients with *IDH* mutation.

both non-invasive techniques used in the detection of *IDH* mutant tumors. The choice between these approaches depends on the specific clinical context, resource availability, and the required information for patient management.

In general, the radiomic method is superior to the 2HG MRS analysis manifesting in heterogeneity assessment, availability, generalizability, and predictive modeling for glioma *IDH* genotyping (45–47). *IDH* mutant gliomas often exhibit significant intratumoral heterogeneity, with diverse regions of aggressiveness, therapy resistance, and molecular characteristics. Radiomics methods can capture this heterogeneity by analyzing multiple regions within the tumor, allowing for a more comprehensive quantitative evaluation of the *IDH* mutant gliomas. However, 2HG MRS typically provides a global assessment of 2HG concentration throughout the entire tumor, potentially missing significant spatial variances. The majority of clinical MRI scanners are already equipped with imaging protocols essential for radiomics analysis. However, performing 2HG MRS often necessitates specific acquisition sequences and dedicated post-processing techniques, which may not be universally accessible in all clinical settings. Currently, radiomics methods are highly adaptable and enable the integration of multimodal imaging data, which can enhance diagnostic and predictive capabilities for *IDH* genotyping. In contrast, 2HG MRS is specific to MRI and solely focuses on measuring 2HG concentration. Radiomics methods often employ machine-learning modeling to extract valuable information from imaging data. By training models on large datasets, radiomics-based approaches not only have the ability to generate predictive models for *IDH* genotyping but also to guide comprehensive prognostic analysis for patients. This capability extends beyond the direct measurement of 2HG, providing a more comprehensive assessment of the tumor and its behavior.

In conventional radiomics analysis, some studies have analyzed the prediction potential of different machine-learning models for specific clinical tasks. Parmer et al. (48) compared 12 machine-learning models in terms of their prediction performance in patients with lung cancer, with random forest achieving the best result (AUC: 0.660 ± 0.030 , RSD_{AUC}: 4.5%). In another study about head and neck cancer, Parmer et al. (49) evaluated 11 machine-learning models in terms of their prediction performance of overall survival, with naive Bayes managing the highest prognostic performance (AUC: 0.670 ± 0.076 , RSD_{AUC}: 11.3%). Fang et al.

(33) compared three machine-learning models in predicting *TERT* genotyping of LGGs, including random forest (AUC: 0.827 ± 0.043 , RSD_{AUC}: 5.2%), adaboost (AUC: 0.820 ± 0.040 , RSD_{AUC}: 4.9%) and linear SVM (AUC: 0.840 ± 0.090 , RSD_{AUC}: 10.7%). In our study, we compared the predictive effectiveness of eight machine-learning classifiers by using only CR features as model predictors, with LDA exhibiting the best prediction performance and stability (AUC: 0.833 ± 0.062 , RSD_{AUC}: 7.4%). Overall, the prediction performance of classifiers using only CR features is generally low, and the predictive ability of the same classifier varies greatly across different clinical tasks. Conventional radiomics analysis always depends on a fixed feature extraction pipeline, which could limit the predictive potential of machine-learning models.

To improve the predictive potential of machine-learning models, extracting features with abundant representational information is the key. CR features extracted by conventional radiomics are predefined and limited in number, resulting in the limited acquisition of tumor heterogeneities. However, a deep-learning model with end-to-end prediction capability can automatically extract features from each layer or transform and represent features layer by layer to obtain a variety of complex features that are closely related to tumor heterogeneities. In medical imaging analysis, some studies have shown that DL models have a powerful ability to improve the understanding of tumor characteristics. When Li et al. (23) were constructing the hierarchical *IDH* and *1p/19q* prediction models, they confirmed that DL features had better stability and reproducibility than CR features, with better generalization ability in *IDH* genotyping prediction (AUC: 0.85–0.89). Chen et al. (21) confirmed that DL features tended to outperform CR features, and the prediction performance of the DL model after adding CR features was improved (AUC: 0.910) in *PTEN* genotyping. Compared to CR features, DL features are extracted more flexibly and can provide more comprehensive information for the construction of machine-learning models. However, when considering the mixture of DL and CR features as predictors, few studies have analyzed the prediction effectiveness of different machine-learning models, and the optimal machine-learning model has not yet been determined for *IDH* genotyping prediction.

Therefore, in our study, we dug deeper into the information contained in MR imaging data and extracted two groups of features (CR and DL) for *IDH* genotyping. We analyzed and compared eight common machine-learning models aiming to find the optimal

classifier. Features were extracted from different MRI modalities and tumor subregions, respectively. MRIs involved in this study were collected from our medical center and the public database TCGA, which may help to build a robust and broadly applicable model for *IDH* genotyping. Moreover, the region of interest in most studies was delineated by radiology specialists. However, we utilized a fully-automated segmentation approach to delineate glioma areas which were the focused zones of feature calculation. Compared to manual segmentation, the fully-automated segmentation methods can reduce the variation of delineation between different observers, produce more reproducible and stable features, and remain low-cost and time-saving options.

In recent studies, deep-learning models used for glioma genotyping were mainly from the CNN family. Chang et al. (50) developed a 2D-ResNet model to non-invasively predict *IDH* genotyping using conventional MR imaging data (AUC: 0.950). Li et al. (23) creatively built a novel 2.5D-ResNet18 model by designing an image sequence as independent input (AUC: 0.890). Some studies also extracted DL features directly from 3D glioma volumes. Chen et al. (21) integrated a 3D-ResNet model and CR features to collaboratively predict *PTEN* genotyping (AUC: 0.910). In our study, we constructed a 2.5D-CNN model to extract initial features from the image sequence and then used LSTM to learn the initial features in an ensemble way. Unlike the 2D- or 3D-CNN model, the 2.5D-CNN model can ensure access to rich tumor information and prevent model underfitting in a small cohort.

In our study for CR features selection, we used LASSO to reduce feature dimensionality and eliminate the risk of model overfitting caused by excessive features. In total, we selected the seven most valuable and stable CR features from ten feature subsets (Supplementary Table S3). Each feature had distinctive prediction potential in *IDH* genotyping ($P < 0.05$). We found that the final selected CR features were primarily texture information, reflecting the characteristics of slow or periodic changes throughout the tumor. It is difficult to observe these tumor texture features with the naked human eye in imaging data. Quantitative analysis of the texture features will be an effective way to help clinicians understand and treat tumors.

Metrics, such as accuracy, specificity, and sensitivity, need to set a threshold in advance to determine whether the predictive sample is positive or negative. However, AUC, an indicator used to evaluate the classifier's ability to distinguish samples between positive and negative, will not be affected by threshold adjustment. Therefore, we mainly refer to the AUC value to compare the prediction performance of classifiers (51). When we used a mixture of DL signatures and CR features as predictors, r-SVM had the highest accuracy but its stability was the worst (AUC: 0.911 ± 0.073 , accuracy: 0.856 ± 0.056 , RSD_{AUC} : 8.0%), while LR had the best AUC and stability (AUC: 0.920 ± 0.043 , accuracy: 0.843 ± 0.044 , RSD_{AUC} : 4.7%). If only the predefined CR features are used as the LR predictors, the learning ability of LR will be greatly limited and prediction performance will not be satisfactory (AUC: 0.830 ± 0.066 , accuracy: 0.771 ± 0.051 , RSD_{AUC} : 8.0%). Compared with CR features, DL signatures exhibit superior prediction values (AUC: 0.915 ± 0.054 , accuracy: 0.835 ± 0.061 , RSD_{AUC} : 5.9%). After adding CR features, the prediction performance of the DL-based model

does not exhibit a significant improvement (AUC: 0.920 ± 0.043 , accuracy: 0.843 ± 0.044 , RSD_{AUC} : 4.7%). Additionally, we confirmed that multimodal signatures improved the prediction performance of *IDH* genotyping and outperformed any single-modality signature in the comparison of prediction potentials. Among the single-modality signature, T1CE and T2WI had the best prediction results (T1CE, AUC: 0.904 ± 0.044 , RSD_{AUC} : 4.9%; T2WI, AUC: 0.899 ± 0.062 , RSD_{AUC} : 6.9%), considering as the important modalities in *IDH* genotyping. Furthermore, the t-SNE visualization suggested that the DL signatures possessed better discriminative ability compared to the CR features. Overall, the results from our analyses suggested that LR should be a preferable machine-learning classifier in *IDH* genotyping and DL signatures exhibit superior prediction values and discriminative capability.

The application of deep-learning for glioma genotyping is an inevitable trend in the future, but it is still in the initial stage and has certain limitations. Firstly, meeting the large sample size requirements of deep-learning is still a problem. A larger sample size and independent validation data are still required to assess the generalization of our model. Secondly, the study of multiple genotyping predictions will help us further understand the tumor characteristics and efficiently guide patients' treatment. Thirdly, the biological mechanisms and clinical interpretations of how DL features relate to *IDH* genotyping still remain unclear. Although we illustrated that DL features have promising prediction value in *IDH* genotyping, further research and understanding are required.

5 Conclusion

Our findings highlight the clinical utility of deep-learning-based radiomics analysis for *IDH* genotyping. Through a nested 10-fold cross-validation process, we developed an efficient LR model with robust performance by combining CR features and DL signatures as model predictors. Through subgroup analysis, it is observed that DL signatures consistently outperform CR features in terms of prediction performance and discriminative capability. The addition of CR features does not significantly enhance the prediction performance of DL-signature-based model, indicating that DL signatures alone exhibit favorable prediction capability, with potential as a standalone approach for accurate predictions. Through t-SNE cluster analysis, DL signatures also display markedly superior clustering and discriminative capability in comparison to CR features. Overall, the future direction of radiomics analysis may revolve around the utilization of custom deep-learning features, emphasizing the importance of incorporating deep-learning techniques to extract robust and informative features from medical imaging data.

Data availability statement

Data of the descriptive analysis, group comparison, and figure description are provided in the supplement. Additional data on the individual case level can be requested from the corresponding

author. Requests to access these datasets should be directed to liuyun@njmu.edu.cn.

Ethics statement

The studies involving humans were approved by The First Affiliated Hospital of Nanjing Medical University (2021-SR-098). 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 in accordance with the national legislation and the institutional requirements.

Author contributions

HZ, JW and WF conceived and designed the study. YoY, JZ, XF, ZW, JN, and FY collected the molecular pathology and imaging data. HZ and JW performed image pre-processing and tumor segmentation. GZ, CW, YH and XZ analyzed the data and performed the statistical analysis. HZ, YuY, and YL wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Industry Prospecting and Common Key Technology Key Projects of Jiangsu Province Science and Technology Department (grant no.: BE2020721), the

Industrial and Information Industry Transformation and Upgrading Special Fund of Jiangsu Province in 2021 (grant no.: [2021]92), the Key Project of Smart Jiangsu in 2020 (grant no.: [2021]1), and Jiangsu Province Engineering Research Center of Big Data Application in Chronic Disease and Intelligent Health Service (grant no.: [2020]1460).

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.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1143688/full#supplementary-material>

References

- Molinari AM, Taylor JW, Wiencke JK, Wrensch MR. Genetic and molecular epidemiology of adult diffuse glioma. *Nat Rev Neurol* (2019) 15:405–17. doi: 10.1038/s41582-019-0220-2
- Wen PY, Weller M, Lee EQ, Alexander BM, Barnholtz-Sloan JS, Barthel FP, et al. Glioblastoma in adults: a Society for Neuro-Oncology (SNO) and European Society of Neuro-Oncology (EANO) consensus review on current management and future directions. *Neuro-Oncology* (2020) 22:1073–113. doi: 10.1093/neuonc/noaa106
- Stupp R, Taillibert S, Kanner A, Read W, Steinberg DM, Lhermitte B, et al. Effect of tumor-treating fields plus maintenance temozolomide vs maintenance temozolomide alone on survival in patients with glioblastoma: A randomized clinical trial. *JAMA* (2017) 318:2306. doi: 10.1001/jama.2017.18718
- Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol* (2007) 114:97–109. doi: 10.1007/s00401-007-0243-4
- Reardon DA, Wen PY. Unravelling tumour heterogeneity—implications for therapy. *Nat Rev Clin Oncol* (2015) 12:69–70. doi: 10.1038/nrclinonc.2014.223
- 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
- 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
- Gilbert MR, Yuan Y, Wu J, Mendoza T, Vera E, Omuro A, et al. A phase II study of dose-dense temozolomide and lapatinib for recurrent low-grade and anaplastic supratentorial, infratentorial, and spinal cord ependymoma. *Neuro-Oncology* (2021) 23:468–77. doi: 10.1093/neuonc/noaa240
- Cairncross G, Wang M, Shaw E, Jenkins R, Brachman D, Buckner J, et al. Phase III trial of chemoradiotherapy for anaplastic oligodendroglioma: long-term results of RTOG 9402. *J Clin Oncol* (2013) 31:337–43. doi: 10.1200/JCO.2012.43.2674
- Van Den Bent MJ, Brandes AA, Taphoorn MJB, Kros JM, Kouwenhoven MCM, Delattre J-Y, et al. Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: long-term follow-up of EORTC brain tumor group study 26951. *JCO* (2013) 31:344–50. doi: 10.1200/JCO.2012.43.2229
- Kaminska B, Czapski B, Guzik R, Król SK, Gielniewski B. Consequences of IDH1/2 mutations in gliomas and an assessment of inhibitors targeting mutated IDH proteins. *Molecules* (2019) 24:968. doi: 10.3390/molecules24050968
- Platten M, Bunse L, Wick A, Bunse T, Le Cornet L, Harting I, et al. A vaccine targeting mutant IDH1 in newly diagnosed glioma. *Nature* (2021) 592:463–8. doi: 10.1038/s41586-021-03363-z
- Tan Y, Mu W, Wang X, Yang G, Gillies RJ, Zhang H. Whole-tumor radiomics analysis of DKI and DTI may improve the prediction of genotypes for astrocytomas: A preliminary study. *Eur J Radiol* (2020) 124:108785. doi: 10.1016/j.ejrad.2019.108785
- Haubold J, Hosch R, Parmar V, Glas M, Guberina N, Catalano OA, et al. Fully automated MR based virtual biopsy of cerebral gliomas. *Cancers* (2021) 13:6186. doi: 10.3390/cancers13246186
- Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur J Cancer* (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036
- Wang J, Zheng X, Zhang J, Xue H, Wang L, Jing R, et al. An MRI-based radiomics signature as a pretreatment noninvasive predictor of overall survival and

- chemotherapeutic benefits in lower-grade gliomas. *Eur Radiol* (2021) 31:1785–94. doi: 10.1007/s00330-020-07581-3
17. Yan J, Zhang B, Zhang S, Cheng J, Liu X, Wang W, et al. Quantitative MRI-based radiomics for noninvasively predicting molecular subtypes and survival in glioma patients. *NPJ Precis Onc* (2021) 5:72. doi: 10.1038/s41698-021-00205-z
18. Calabrese E, Rudie JD, Rauschecker AM, Villanueva-Meyer JE, Clarke JL, Solomon DA, et al. Combining radiomics and deep convolutional neural network features from preoperative MRI for predicting clinically relevant genetic biomarkers in glioblastoma. *Neuro-Oncol Adv* (2022) 4:vdac060. doi: 10.1093/oaajnl/vdac060
19. Wu S, Zhang X, Rui W, Sheng Y, Yu Y. A nomogram strategy for identifying the subclassification of IDH mutation and ATRX expression loss in lower-grade gliomas. *Eur Radiol* (2022) 32:3187–98. doi: 10.1007/s00330-021-08444-1
20. Li Z, Wang Y, Yu J, Guo Y, Cao W. Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. *Sci Rep* (2017) 7:5467. doi: 10.1038/s41598-017-05848-2
21. Chen H, Lin F, Zhang J, Lv X, Zhou J, Li Z-C, et al. Deep learning radiomics to predict PTEN mutation status from magnetic resonance imaging in patients with glioma. *Front Oncol* (2021) 11:734433. doi: 10.3389/fonc.2021.734433
22. Zeng H, Xing Z, Gao F, Wu Z, Huang W, Su Y, et al. A multimodal domain adaptive segmentation framework for IDH genotype prediction. *Int J CARS* (2022) 17:1923–31. doi: 10.1007/s11548-022-02700-5
23. Li Y, Wei D, Liu X, Fan X, Wang K, Li S, et al. Molecular subtyping of diffuse gliomas using magnetic resonance imaging: comparison and correlation between radiomics and deep learning. *Eur Radiol* (2022) 32:747–58. doi: 10.1007/s00330-021-08237-6
24. Lu C-F, Hsu F-T, Hsieh KL-C, Kao Y-CJ, Cheng S-J, Hsu JB-K, et al. Machine learning-based radiomics for molecular subtyping of gliomas. *Clin Cancer Res* (2018) 24:4429–36. doi: 10.1158/1078-0432.CCR-17-3445
25. Park CJ, Choi YS, Park YW, Ahn SS, Kang S-G, Chang J-H, et al. Diffusion tensor imaging radiomics in lower-grade glioma: improving subtyping of isocitrate dehydrogenase mutation status. *Neuroradiology* (2020) 62:319–26. doi: 10.1007/s00234-019-02312-y
26. Rohlfing T, Zahr NM, Sullivan EV, Pfefferbaum A. The SRI24 multichannel atlas of normal adult human brain structure. *Hum Brain Mapp* (2010) 31:798–819. doi: 10.1002/hbm.20906
27. Isensee F, Jäger PF, Kohl SAA, Petersen J, Maier-Hein KH. Automated design of deep learning methods for biomedical image segmentation. *Nat Methods* (2021) 18:203–11. doi: 10.1038/s41592-020-01008-z
28. Qu R, Xiao Z. An attentive multi-modal CNN for brain tumor radiogenomic classification. *Information* (2022) 13:124. doi: 10.3390/info13030124
29. Turan M, Durmus F. UC-NfNet: Deep learning-enabled assessment of ulcerative colitis from colonoscopy images. *Med Image Anal* (2022) 82:102587. doi: 10.1016/j.media.2022.102587
30. Huang Z, Xu W, Yu K. Bidirectional LSTM-CRF models for sequence tagging. (2015). doi: 10.48550/arXiv.1508.01991
31. He K, Zhang X, Ren S, Sun J. (2016). Deep residual learning for image recognition, in: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Las Vegas, NV, USA: IEEE), pp. 770–8. doi: 10.1109/CVPR.2016.90
32. Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative. *Radiology* (2020) 295:328–38. doi: 10.1148/radiol.2020191145
33. Fang S, Fan Z, Sun Z, Li Y, Liu X, Liang Y, et al. Radiomics features predict telomerase reverse transcriptase promoter mutations in world health organization grade II gliomas via a machine-learning approach. *Front Oncol* (2021) 10:606741. doi: 10.3389/fonc.2020.606741
34. Muthukrishnan R, Rohini R. (2016). LASSO: A feature selection technique in predictive modeling for machine learning, in: *2016 IEEE International Conference on Advances in Computer Applications (ICACA)* (Coimbatore, India: IEEE), pp. 18–20. doi: 10.1109/ICACA.2016.7887916
35. Tan Y, Zhang S, Wei J, Dong D, Wang X, Yang G, et al. A radiomics nomogram may improve the prediction of IDH genotype for astrocytoma before surgery. *Eur Radiol* (2019) 29:3325–37. doi: 10.1007/s00330-019-06056-4
36. Kanazawa T, Minami Y, Takahashi H, Fujiwara H, Toda M, Jinzaki M, et al. Magnetic resonance imaging texture analyses in lower-grade gliomas with a commercially available software: correlation of apparent diffusion coefficient and T2 skewness with 1p/19q codeletion. *Neurosurg Rev* (2020) 43:1211–9. doi: 10.1007/s10143-019-01157-6
37. Li Z-C, Bai H, Sun Q, Li Q, Liu L, Zou Y, et al. Multiregional radiomics features from multiparametric MRI for prediction of MGMT methylation status in glioblastoma multiforme: A multicentre study. *Eur Radiol* (2018) 28:3640–50. doi: 10.1007/s00330-017-5302-1
38. Li Y, Liang Y, Sun Z, Xu K, Fan X, Li S, et al. Radiogenomic analysis of PTEN mutation in glioblastoma using preoperative multi-parametric magnetic resonance imaging. *Neuroradiology* (2019) 61:1229–37. doi: 10.1007/s00234-019-02244-7
39. Song S, Wang L, Yang H, Shan Y, Cheng Y, Xu L, et al. Static 18F-FET PET and DSC-PWI based on hybrid PET/MR for the prediction of gliomas defined by IDH and 1p/19q status. *Eur Radiol* (2021) 31:4087–96. doi: 10.1007/s00330-020-07470-9
40. Kim M, Jung SY, Park JE, Jo Y, Park SY, Nam SJ, et al. Diffusion- and perfusion-weighted MRI radiomics model may predict isocitrate dehydrogenase (IDH) mutation and tumor aggressiveness in diffuse lower grade glioma. *Eur Radiol* (2020) 30:2142–51. doi: 10.1007/s00330-019-06548-3
41. Pei D, Guan F, Hong X, Liu Z, Wang W, Qiu Y, et al. Radiomic features from dynamic susceptibility contrast perfusion-weighted imaging improve the three-class prediction of molecular subtypes in patients with adult diffuse gliomas. *Eur Radiol* (2023) 33:3455–66. doi: 10.1007/s00330-023-09459-6
42. Leather T, Jenkinson M, Das K, Poptani H. Magnetic resonance spectroscopy for detection of 2-hydroxyglutarate as a biomarker for IDH mutation in gliomas. *Metabolites* (2017) 7:29. doi: 10.3390/metabo7020029
43. Hangel G, Schmitz-Abecassis B, Sollmann N, Pinto J, Arzanforoosh F, Barkhof F, et al. Advanced MR techniques for preoperative glioma characterization: part 2. *J Magnetic Resonance Imaging* (2023) 57:1676–95. doi: 10.1002/jmri.28663
44. Andronesi OC, Loebel F, Bogner W, Marjańska M, Vander Heiden MG, Iafrate AJ, et al. Treatment response assessment in IDH-mutant glioma patients by noninvasive 3D functional spectroscopic mapping of 2-hydroxyglutarate. *Clin Cancer Res* (2016) 22:1632–41. doi: 10.1158/1078-0432.CCR-15-0656
45. Bumes E, Fellner C, Fellner FA, Fleischanderl K, Häckl M, Lenz S, et al. Validation study for non-invasive prediction of IDH mutation status in patients with glioma using *in vivo* 1H-magnetic resonance spectroscopy and machine learning. *Cancers* (2022) 14:2762. doi: 10.3390/cancers14112762
46. Suh CH, Kim HS, Jung SC, Choi CG, Kim SJ. 2-Hydroxyglutarate MR spectroscopy for prediction of isocitrate dehydrogenase mutant glioma: a systemic review and meta-analysis using individual patient data. *Neuro-Oncology* (2018) 20:1573–83. doi: 10.1093/neuonc/noy113
47. Berrington A, Voets NL, Larkin SJ, De Pennington N, McCullagh J, Stacey R, et al. A comparison of 2-hydroxyglutarate detection at 3 and 7 T with long-TE semi-LASER. *NMR Biomed* (2018) 31:e3886. doi: 10.1002/nbm.3886
48. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep* (2015) 5:13087. doi: 10.1038/srep13087
49. Parmar C, Grossmann P, Rietveld D, Rietbergen MM, Lambin P, Aerts HJWL. Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer. *Front Oncol* (2015) 5:272. doi: 10.3389/fonc.2015.00272
50. Chang K, Bai HX, Zhou H, Su C, Bi WL, Agboda E, et al. Residual convolutional neural network for the determination of IDH status in low- and high-grade gliomas from MR imaging. *Clin Cancer Res* (2018) 24:1073–81. doi: 10.1158/1078-0432.CCR-17-2236
51. Wu S, Meng J, Yu Q, Li P, Fu S. Radiomics-based machine learning methods for isocitrate dehydrogenase genotype prediction of diffuse gliomas. *J Cancer Res Clin Oncol* (2019) 145:543–50. doi: 10.1007/s00432-018-2787-1