Automated machine learning to predict the co-occurrence of isocitrate dehydrogenase mutations and O^6-methylguanine-DNA methyltransferase promoter methylation in patients with gliomas

Simin Zhang, PhD,[1,2] Huaiqiang Sun, PhD,[1] Xiaorui Su, PhD,[1,2] Xibiao Yang, MM,[2,3] Weina Wang, PhD,[1] Xinyue Wan, MM,[1] Qiaoyue Tan, MM,[1,4] Ni Chen, MD, PhD,[5] Qiang Yue, MD, PhD,[2,3]* and Qiyong Gong, MD, PhD[1]*

Abstract

Combining isocitrate dehydrogenase mutation (IDHmut) with O^6-methylguanine-DNA methyltransferase promoter methylation (MGMTmet) has been identified as a critical prognostic molecular marker for gliomas. The aim of this study was to determine the ability of glioma radiomics features from magnetic resonance imaging (MRI) to predict the co-occurrence of IDHmut and MGMTmet by applying the tree-based pipeline optimization tool (TPOT), an automated machine learning (autoML) approach. This was a retrospective study, in which 162 patients with gliomas were evaluated, including 58 patients with co-occurrence of IDHmut and MGMTmet and 104 patients with other status comprising: IDH wildtype and MGMT unmethylated (n = 67), IDH wildtype and MGMTmet (n = 36), and IDHmut and MGMT unmethylated (n = 1). Three-dimensional (3D) T1-weighted images, gadolinium-enhanced 3D T1-weighted images (Gd-3DT1WI), T2-weighted images, and fluid-attenuated inversion recovery (FLAIR) images acquired at 3.0 T were used. Radiomics features were extracted from FLAIR and Gd-3DT1WI images. The TPOT was employed to generate the best machine learning pipeline, which contains both feature selector and classifier, based on input feature sets. A 4-fold cross-validation was used to evaluate the performance of automatically generated models. For each iteration, the training set included 121 subjects, while the test set included 41 subjects. Student's t-test or a chi-square test was applied on different clinical characteristics between two groups. Sensitivity, specificity, accuracy, kappa score, and AUC were used to evaluate the performance of TPOT-generated models. Finally, we compared the above metrics of TPOT-generated models to identify the best-performing model. Patients' ages and grades between two groups were significantly different (p = 0.002 and p = 0.000, respectively). The 4-fold cross-validation showed that gradient boosting classifier trained on shape and textual features from the Laplacian-of-Gaussian-filtered Gd-3DT1 achieved the best performance (average sensitivity = 81.1%, average specificity = 94%, average accuracy = 89.4%, average kappa score = 0.76, average AUC = 0.951). Using autoML based on radiomics features from MRI, a high discriminatory accuracy was achieved for predicting co-occurrence of IDHmut and MGMTmet in gliomas.

Level of Evidence: 3

Technical Efficacy Stage: 3

J. MAGN. RESON. IMAGING 2021.

View this article online at wileyonlinelibrary.com. DOI: 10.1002/jmri.27498

Received Oct 2, 2020, Accepted for publication Dec 18, 2020.

*Correspondence Qiang Yue, Department of Radiology, West China Hospital of Sichuan University, #37 GuoXue Xiang, Chengdu, Sichuan 610041, China.
Email: scu_yq@163.com
Qiyong Gong, Huaxi MR Research Center (HMRRC), Department of Radiology, West China Hospital of Sichuan University, Chengdu, Sichuan 610041, China.
Email: qiyonggong@hmrrc.org.cn
Simin Zhang and Huaiqiang Sun contributed equally to this work.

From the [1]Huaxi MR Research Center (HMRRC), Functional and Molecular Imaging Key Laboratory of Sichuan Province, Department of Radiology, West China Hospital of Sichuan University, Chengdu, China; [2]Huaxi Glioma Center, West China Hospital of Sichuan University, Chengdu, China; [3]Department of Radiology, West China Hospital of Sichuan University, Chengdu, China; [4]Division of Radiation Physics, State Key Laboratory of Biotherapy and Cancer Center, West China Hospital of Sichuan University, Chengdu, China; and [5]Department of Pathology, West China Hospital of Sichuan University, Chengdu, China

Additional supporting information may be found in the online version of this article

INTRODUCTION

Gliomas are the most common malignant primary brain tumors in adults.[1] High-grade gliomas tend to have poor outcomes despite aggressive treatment, including surgery and adjuvant concurrent chemoradiation therapy with temozolomide.[2] Hence, the early identification of potential prognostic characteristics is of value. Overall survival and the response to therapy are highly dependent on both the World Health Organization (WHO) grade and molecular characteristics, particularly the isocitrate dehydrogenase mutation (IDHmut) and O^6-methylguanine-DNA methyltransferase promoter methylation (MGMTmet) status.[3–5] A previous study reported that glioblastomas with co-occurrence of IDHmut and MGMTmet had the longest survival, followed by patients with IDHmut and MGMT unmethylated (MGMTunmet) or IDH wildtype (IDHwt) and MGMTmet glioblastoma, while patients with IDHwt and MGMTunmet glioblastoma showed the shortest survival; and the study suggested that the combination of IDHmut and MGMTmet outperforms either IDHmut or MGMTmet alone to predict survival in glioblastoma.[6] Similarly, another study demonstrated that the co-occurrence of IDHmut and MGMTmet is associated with long-term survival in low-grade glioma.[7] Thus, combination of IDHmut and MGMTmet is a predictor with more prognostic power than IDHmut or MGMTmet alone. Currently, information on the two genotypes can be obtained only via invasive tissue sampling and immunohistochemistry and genomic sequence analysis. However, these methods may not have high accuracy on account of intratumoral heterogeneity and insufficiency of tumor samples.[8]

As a noninvasive method, radiomics from magnetic resonance imaging (MRI) may take full advantage of the deep information, which play an important role in the genetic-level classification of gliomas. This approach employs a set of qualitative and quantitative analyses of high-throughput image features to acquire predictive information derived from MRI.[9] With the potential ability to improve the predictive accuracy, radiomics was widely utilized to classify molecular subtypes based on various MRI images in gliomas.[10,11] However, most previous radiomics studies aimed to predict IDHmut or MGMTmet alone,[12,13] which rarely focused on the co-occurrence of IDHmut and MGMTmet.

Predicting the IDH and MGMT status is possible by constructing radiomics-based classifiers applying machine learning (ML) methods. However, during ML, choosing the most appropriate combination of feature selector and classifier and corresponding hyperparameters tends to be complicated for researchers without ML experience.[14] To overcome these limitations, a relatively new system called "automated machine learning (autoML)" has been developed to automate these challenging and time-consuming processes.[15] The tree-based pipeline optimization tool (TPOT) is an autoML system that adopts a higher-level architecture for the optimization process by utilizing genetic programming to search and generate optimal ML pipelines.[16] The performance of TPOT-generated ML pipelines has been observed to match or exceed that of the models tuned by ML experts while requiring minimal intervention.[17,18]

In the current study, we aimed to evaluate the ability of TPOT based on radiomics features from MRI to predict the co-occurrence of IDHmut and MGMTmet in gliomas.

MATERIALS AND METHODS

Patient Cohort

This retrospective study was approved by the local research ethics committee, and the requirement to obtain individual informed consent was waived because of the retrospective nature of this study. Four hundred and fifty-seven patients with brain tumor were selected from routine clinical scans performed between October 2016 and November 2020. MRI scanning was performed according to the current clinical protocols, and no additional scans or sequences were carried out for purely research purposes. All the included patients met the following criteria: (1) histopathologically confirmed primary Grade II, III, or IV gliomas with reference to the latest WHO classification; (2) available IDH and MGMT genotype; and (3) full preoperative MRI scanning, including high-resolution three-dimensional (3D) T1-weighted images (3D T1WI), T2-weighted images (T2WI), fluid-attenuated inversion recovery (FLAIR), and gadolinium-enhanced 3D T1-weighted images (Gd-3DT1WI). Initially, 171 patients met the inclusion criteria; however, nine patients were excluded because of poor image quality. Finally, 162 patients were enrolled for the subsequent training and testing of autoML models, Of these, 58 patients with co-occurrence of IDHmut and MGMTmet, and 104 patients with other status, including 67 patients with IDHwt and MGMTunmet, 36 patients with IDHwt and MGMTmet, and one patient with IDHmut and MGMTunmet. The process of patients' enrollment was presented in Figure S1.

Image Acquisition

Images were obtained with a 3.0 T MR scanner (Skyra, Siemens Healthineers) and 20-channel head coil. The protocol included the following sequences: 3D T1WI (repetition time [TR]/echo time [TE]/flip angle = 1540 ms/2.4 ms/8°, matrix = 256 × 256, field of view [FOV] = 230 × 230 mm, slice thickness = 1 mm); T2WI (TR/TE/flip angle = 4500/105/150°, matrix = 320 × 230, FOV = 176 × 220 mm, slice thickness = 5 mm); FLAIR (TR/TE/flip angle = 6000 ms/81 ms/90°, matrix = 256 × 198, FOV = 195 × 220 mm, thickness = 5 mm); Gd-3DT1WI (0.1 mmol/kg body weight of gadobenate dimeglumine [Multihance, Braccosine, Shanghai, China] was administered intravenously at a rate of 4.0 ml/s, followed by a 30-ml saline flush, TR/TE/flip angle = 1540 ms/2.4 ms/8°, matrix = 230 × 230 mm, slice thickness = 1 mm).

Glioma Segmentation

A region of interest (ROI) encompassing whole tumors was segmented using the open-source NiftyNet platform,[19] which is a deep learning based fully automatic segmentation approach that has been shown to be more reliable and faster than other semi-automatic methods.[20] The details of the segmentation process are described as follows: (1) the 3DT1 volume was resampled to 1 mm^3 isotropic resolution; (2) the resampled 3DT1 volume was automatically cropped using robustfov from the fsl package (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/InitialProcessing), and the skull was removed using the fsl brain extraction tool (BET) (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET); (3) the remaining Gd-3DT1, FLAIR, and T2 volumes were co-registered to the processed 3DT1 volume and multiplied with the brain mask generated by BET; (4) all four processed volumes were merged and fed into the segmentation model trained on the multimodal brain tumor segmentation (BraTS) dataset, then a mask encompassing both tumor core and edema area was generated.[21] The initial mask was then visually inspected and manually corrected by a radiologist (Xibiao Yang. with 10 years of experience in neuro-oncology) utilizing ITK-SNAP software (http://www.itksnap.org). A detailed illustration of steps on manual correction is provided in Table S1.

Radiomics Feature Extraction

For radiomics analysis, we applied an open-source python package to extract radiomics features (Pyradiomics Library, version 2.0.0).[22] By doing so, the radiomics features, including tumor shape and texture features, were automatically extracted from each ROI.[22,23] Laplacian-of-Gaussian (LoG) and wavelet filters were applied to the original images, and radiomics features were extracted from these images, as well. The kernel width (σ) of LoG is listed as follows: 1 mm (fine-scale filtration), 3 mm (medium-scale filtration), and 5 mm (coarse-scale filtration). In total, 1,102 radiomics features (14 shape features + 68 textural features × (1 original image + 3 LoG filtered images + 4 wavelet filtered images) × 2 imaging sequences) were calculated for each subject. Table S2 displays details regarding the extracted features.

Tree-Based Pipeline Optimization Tool

TPOT OVERVIEW. Here, we used TPOT (http://epistasislab.github.io/tpot/) as an autoML system to construct optimal ML pipelines for our image dataset.[24] In brief, TPOT employs genetic programming to select a set of featured preprocessing functions and ML classification or regression algorithms to maximize model performance for the dataset of interest.[23,25] The performance of the model is then evaluated utilizing a fitness function that selects more powerful features over weaker features.[26]

TPOT ANALYSIS. To explore the diagnostic value of Gd-3DT1 and FLAIR and the scale of the disease-related features displayed, the extracted radiomics features from above two sequences were categorized into 23 feature sets (the 23 feature sets are presented in Table S3).

To ensure the stability and reproductivity, an outer 4-fold cross validation was employed to evaluate the performance of automatic generated models. Briefly, the dataset is randomly divided into four equal subsets; during each iteration, the four subsets are alternated between training and testing cohorts. Note that each iteration of cross-validation procedure represents a new training and testing phase on a unique combination of the four subsets. Thereby, the TPOT procedure run four times to generalize the reliability of the results. The evaluation procedure using cross validation was presented in Figure S2.

The settings of TPOT are as follows: generation number, 50; size of population, 100; and inner 10-fold cross validation. For each generation, TPOT replicated the pipelines that currently performed best and then made random changes to them (e.g., adding or deleting an operation or adjusting the parameter settings). These random changes can positively or negatively affect pipeline performance; thus, TPOT continued to detect new pipelines. At the end of each generation, the pipeline with the worst performance was removed, and TPOT entered the next generation. Finally, the best-performing pipeline created during the optimization process was recommended using TPOT. The workflow in each iteration is shown in Figure 1.

Further, according to the feature importance scores, TPOT automatically produced features from the best model which are most useful in predicting co-occurrence of IDHmut and MGMTmet. The scores of each feature were added together across 4-fold cross validation, and the top 10 features were selected as important features that contributed to final prediction model.

Statistical Analysis

Student's *t*-test was performed to evaluate the significant differences in age between patients with co-occurrence of IDHmut and MGMTmet and patients with other status. The chi-square test was performed to determine significant differences in the gender and grades between groups. The level of confidence was kept at 95%, and results with *p*-values less than 0.05 were considered to be significant.

The predictive power of the automatic generated models was measured using the average of metrics including sensitivity, specificity, accuracy, kappa score, and AUC across 4-fold cross-validation.

RESULTS

Demographics

No significant difference was found in gender ($p = 0.162$) between the co-occurrence of IDHmut and MGMTmet and other status. However, the mean age of the co-occurrence of the IDHmut and MGMTmet groups was younger than that of the other status group (41.60 ± 11.40 years vs. 48.34 ± 14.19 years, $p = 0.002$). The constituent ratios of glioma grades are significantly different ($p = 0.000$). The constituent ratio of lower grade (II + III) in co-occurrence of IDHmut and MGMTmet groups was higher than in the other status group. In contrast, the ratio of high grade (IV) in other status was higher than in the co-occurrence of IDHmut and MGMTmet. The clinical and statistical results of the study are summarized in Table 1.

Classification Model Derived from TPOT

The TPOT generated the optimal ML models for every feature set in the training cohort. The performance of each

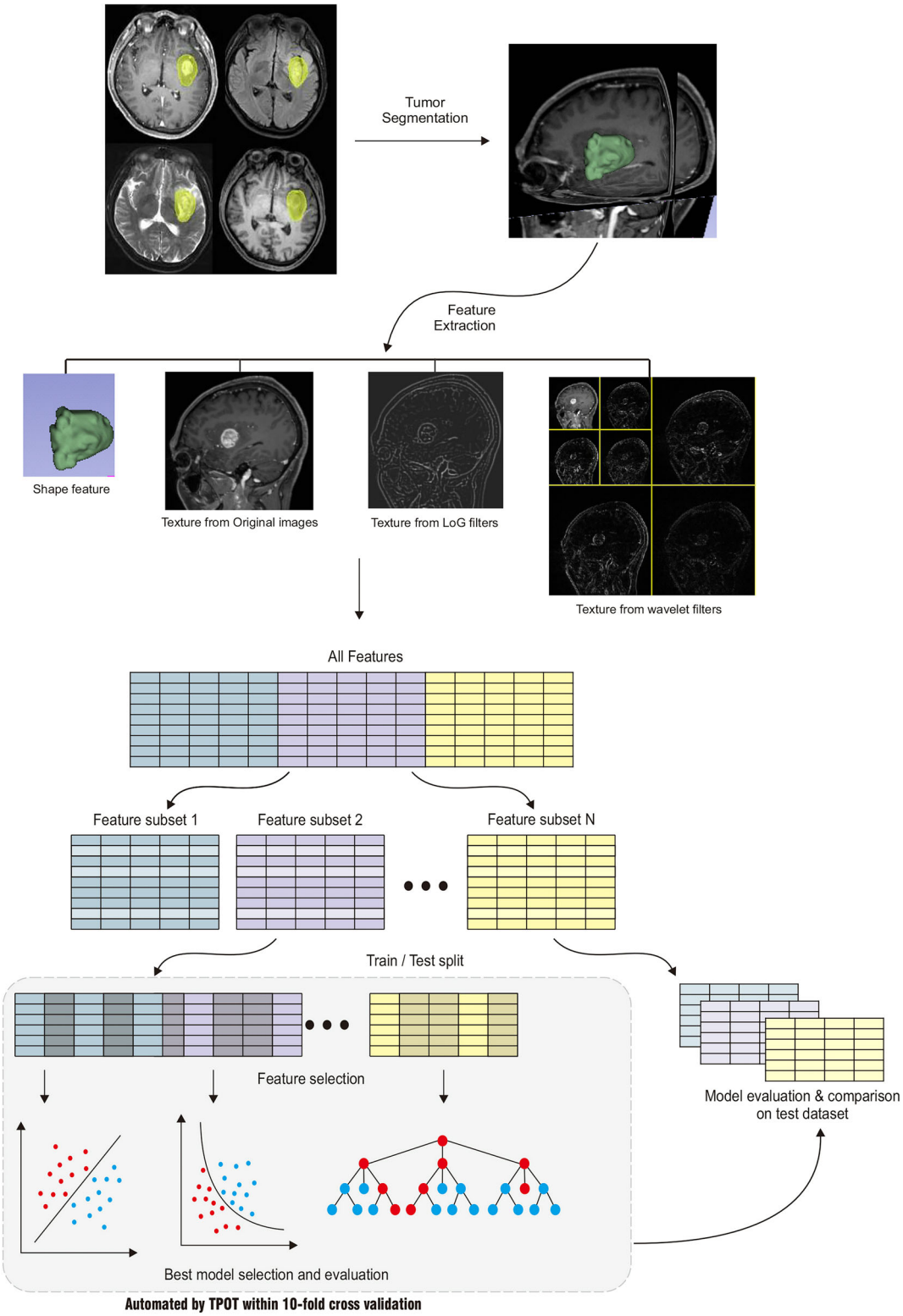

FIGURE 1: An overview of the current study. The subject's glioma was automatically segmented from Gd-3DT1, FLAIR, T2, and 3DT1 (the yellow region indicates the tumor mask on each sequence) using the open-source NiftyNet platform. We first extracted the radiomics features. Shape features were extracted from the segmented result (the green shape); texture features were extracted from original images, Laplacian-of-Gaussian (LoG) filter images, and wavelet filter images. Then, all the extracted features were categorized into 23 feature sets. Each feature set split into training set (dark color) and test set (light color) then entered to TPOT analysis. In the training phase, an inner 10-fold cross validation is used to tune the hyperparameters and select the optimal model. We then evaluated the optimal model on the independent test set. Finally, the best classification pipeline for our dataset was constructed

generated model was assessed in a separate test set. Table 2 lists all the generated models and their average performance of 4-fold cross validation. According to the model comparison, gradient boosting classifier trained on the shape and textual features from the LoG-filtered Gd-3DT1 identified by TPOT (Model 7) achieved the best performance (average sensitivity = 81.1%, average specificity = 94%, average accuracy = 89.4%, average kappa score = 0.76, average AUC = 0.951). Table S4a–d provides the performance of 23 models in each iteration and detailed parameters used in these models. Figure 2 shows the receiver operating characteristic curves illustrating the predictive performance of Model 7 in each iteration.

Important Features Contributing to Classification

The top 10 features with the strongest predictive ability are presented in Figure 3 and include gray-level co-occurrence matrix (GLCM) correlation from LoG-filtered Gd-3DT1, GLCM cluster prominence from LoG-filtered Gd-3DT1, GLCM Informational Measure of Correlation (IMC) 1 from LoG-filtered Gd-3DT1, GLCM Inverse Difference Moment Normalized (IDMN) from LoG-filtered Gd-3DT1, the shape surface area to volume ratio (SAVR), gray-level size zone matrix (GLSZM) zone entropy, shape-maximum two-dimensional (2D) diameter column, GLCM maximum probability from LoG-filtered Gd-3DT1, GLCM cluster shade from LoG-filtered Gd-3DT1, and shape-sphericity.

DISCUSSION

We evaluated the ability of autoML methods implemented in TPOT to predict the co-occurrence of IDHmut and MGMTmet in patients with gliomas. Over the 4-fold cross validation, the gradient-boosting classifier built on shape and texture derived from LoG-filtered Gd-3DT1 achieved the best performance.

The ability of TPOT is to identify the best pipeline that can be used to fit the statistical properties of the underlying dataset while controlling for overfitting and reliability.[27] Prior studies have compared the models derived by TPOT with a basic random forest (RF) and relevance vector regression (RVR) and shown that the models from TPOT had a significant improvement than RF or RVR.[24,27] Our findings also showed a relatively high accuracy of prediction and suggested that TPOT is feasible and promising for genotype status prediction.

Considering that Gd-3DT1 contains information on angiogenesis of the tumor and the disruption of the blood–brain barrier, while FLAIR contains information on the density of tumor cells,[28] these two sequences were involved in current study. As a crucial step in radiomics studies, tumor segmentation has a significant impact on subsequent feature extraction.[29] Manual segmentation is often required for

TABLE 1. Patient demographics and tumor genotypes

Demographics	IDHmut + MGMTmet *N* = 58	Other status, IDHwt + MGMTunmet = 67; IDHwt + MGMTmet = 36; IDHmut + MGMTunmet = 1 *N* = 104	*p*-value (IDHmut + MGMTmet vs. Other status)
Age (years)			
Mean, years (range)	41.60 (21–78)	48.34 (12–79)	0.002*
Sex			
Male, *n* (%)	29 (50%)	63 (60.1%)	0.193
Grade, *n* (%)			
Lower grade (II + III)	48 (82.8%)	30 (28.8%) (subgroup 1 = 18; subgroup 2 = 11; subgroup 3 = 1)	0.000*
High grade (IV)	10 (17.2%)	74 (71.2%) (subgroup 1 = 49; subgroup 2 = 25)	

Abbreviations: IDHmut, isocitrate dehydrogenase mutation; IDHwt, isocitrate dehydrogenase wildtype; MGMTmet, O^6-methylguanine-DNA methyltransferase promoter methylation; MGMTunmet, O^6-methylguanine-DNA methyltransferase promoter unmethylated.
*$p < 0.05$.

TABLE 2. Twenty-three optimal models identified by TPOT with different input feature sets and their average performance of the 4-fold cross validation

Index	Model	Sensitivity	Specificity	Accuracy	Kappa score	AUC
1	Shape	0.656 ± 0.048	0.865 ± 0.022	0.79 ± 0.015	0.533 ± 0.036	0.848 ± 0.032
2	Original Gd-3DT1	0.537 ± 0.102	0.885 ± 0.044	0.756 ± 0.052	0.446 ± 0.122	0.774 ± 0.023
3	LoG-filtered Gd-3DT1	0.766 ± 0.112	0.946 ± 0.042	0.881 ± 0.066	0.734 ± 0.152	0.932 ± 0.066
4	Wavelet-filtered Gd-3DT1	0.577 ± 0.078	0.827 ± 0.067	0.738 ± 0.048	0.414 ± 0.098	0.769 ± 0.023
5	Shape + Original + LoG-filtered + wavelet-filtered Gd-3DT1	0.636 ± 0.160	0.885 ± 0.083	0.796 ± 0.0499	0.538 ± 0.115	0.879 ± 0.051
6	Shape + Original Gd-3DT1	0.588 ± 0.103	0.875 ± 0.048	0.772 ± 0.011	0.481 ± 0.04	0.793 ± 0.037
7	Shape + LoG-filtered Gd-3DT1	*0.811 ± 0.215*	*0.94 ± 0.089*	*0.894 ± 0.062*	*0.76 ± 0.153*	*0.951 ± 0.039*
8	Shape + wavelet-filtered Gd-3DT1	0.67 ± 0.095	0.846 ± 0.083	0.784 ± 0.073	0.524 ± 0.155	0.84 ± 0.065
9	Shape + original + LoG-filtered Gd-3DT1	0.675 ± 0.065	0.901 ± 0.066	0.821 ± 0.063	0.597 ± 0.131	0.893 ± 0.054
10	Shape + Original + wavelet-filtered Gd-3DT1	0.605 ± 0.051	0.884 ± 0.031	0.784 ± 0.023	0.51 ± 0.051	0.823 ± 0.048
11	Original + wavelet-filtered Gd-3DT1	0.64 ± 0.053	0.875 ± 0.037	0.79 ± 0.0146	0.53 ± 0.032	0.808 ± 0.053
12	Original + LoG-filtered + wavelet-filtered Gd-3DT1	0.586 ± 0.06	0.913 ± 0.036	0.796 ± 0.038	0.529 ± 0.09	0.852 ± 0.07
13	LoG-filtered + wavelet-filtered Gd-3DT1	0.638 ± 0.119	0.846 ± 0.07	0.772 ± 0.064	0.492 ± 0.146	0.854 ± 0.058
14	Original FLAIR	0.569 ± 0.064	0.788 ± 0.067	0.71 ± 0.039	0.362 ± 0.068	0.707 ± 0.038
15	LoG-filtered FLAIR	0.436 ± 0.253	0.885 ± 0.083	0.722 ± 0.053	0.335 ± 0.184	0.72 ± 0.057
16	Wavelet-filtered FLAIR	0.511 ± 0.177	0.856 ± 0.0578	0.732 ± 0.058	0.38 ± 0.155	0.758 ± 0.034
17	Shape + Original + LoG-filtered + wavelet-filtered FLAIR	0.534 ± 0.103	0.779 ± 0.036	0.692 ± 0.049	0.317 ± 0.114	0.746 ± 0.032
18	Shape + Original FLAIR	0.383 ± 0.156	0.846 ± 0.137	0.679 ± 0.055	0.248 ± 0.116	0.729 ± 0.049
19	Shape + LoG-filtered FLAIR	0.451 ± 0.189	0.875 ± 0.048	0.722 ± 0.059	0.345 ± 0.168	0.761 ± 0.04
20	Original + LoG-filtered FLAIR	0.497 ± 0.144	0.798 ± 0.065	0.691 ± 0.03	0.303 ± 0.091	0.732 ± 0.044
21	Original + wavelet-filtered FLAIR	0.523 ± 0.148	0.798 ± 0.085	0.7 ± 0.053	0.328 ± 0.13	0.716 ± 0.048
22	LoG-filtered + wavelet-filtered FLAIR	0.603 ± 0.103	0.788 ± 0.038	0.722 ± 0.028	0.392 ± 0.071	0.755 ± 0.041
23	Wavelet-filtered Gd-3DT1 + LoG-filtered FLAIR	0.569 ± 0.064	0.798 ± 0.048	0.716 ± 0.048	0.373 ± 0.107	0.77 ± 0.036

Note: Values in italics indicate the best score in each metric.
Abbreviations: FLAIR, fluid-attenuated inversion recovery; Gd-3DT1, gadolinium-enhanced, high-resolution 3D T1-weighted images; LoG-filtered, Laplacian of Gaussian-filtered; TPOT, tree-based pipeline optimization tool.

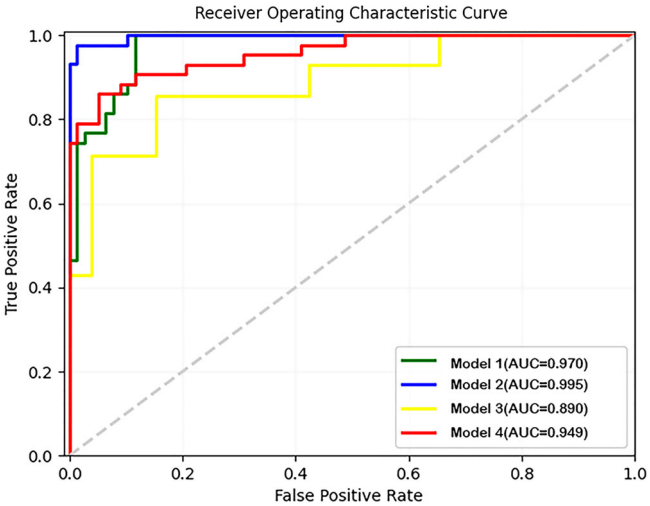

FIGURE 2: Receiver operating characteristic curve analysis for optima model generated by tree-based pipeline optimization tool. Separate curves were plotted for each outer cross-validation fold along with corresponding AUC value

tumors with irregular shapes or fuzzy borders. However, manual segmentation is a time-consuming and labor-intensive work. In this study, we integrated a fully automated, deep learning-based approach for glioma segmentation, making the whole method more practical in the clinical setting. From our results, the trained segmentation model performed satisfactorily in most cases; only five patients were mis-segmented the periventricular white matter hyperintensities resulted from small vessel disease as edema area of tumor. Of note, the mis-segmented regions were easy to recognize and remove manually, as shown in Table S1. Generally, this automatic approach is beneficial to the feasibility of the proposed workflow, which has great potential to be applied in routine clinical practice.

Notably, the predictive ability of texture is associated with the sequence of MRI and whether the LoG filter is applied. Features from the Gd-3DT1 sequence (Models 2–13) were found to be more predictive than those from FLAIR images (Models 14–23). This finding might be interpreted as a consequence of the relatively lower resolution of the FLAIR images, which might limit the radiomics feature extraction, thereby reducing the predictive ability. Additionally, in the feature sets from Gd-3DT1 sequence, when LoG filtering was employed, the predictive performance of the models outperformed the models without LoG filtering. The LoG filter is a tool that can reduce the MRI high-frequency signal noise and simultaneously reduce the effects of large signal variations that can be detected in each image slice—for example, the variation from an inhomogeneous magnetic field.[30] The extraction of LoG-filtered features was strongly determined by the kernel size. In brief, low σ values indicate fine textures, whereas high σ values indicate coarse textures. The current study demonstrated that the filtered volumes encode the characteristics of tissue heterogeneity, which may be related to the co-occurrence of IDHmut and MGMTmet.

In the top 10 features with contribution to prediction from the best-performing model, seven of them were derived from texture features, including GLCM correlation, IMC1, and IDMN. The GLCM correlation parameter is interpreted as the grayscale linear dependence between adjacent pixels.[22] A lower GLCM correlation in fine-filtered images in the

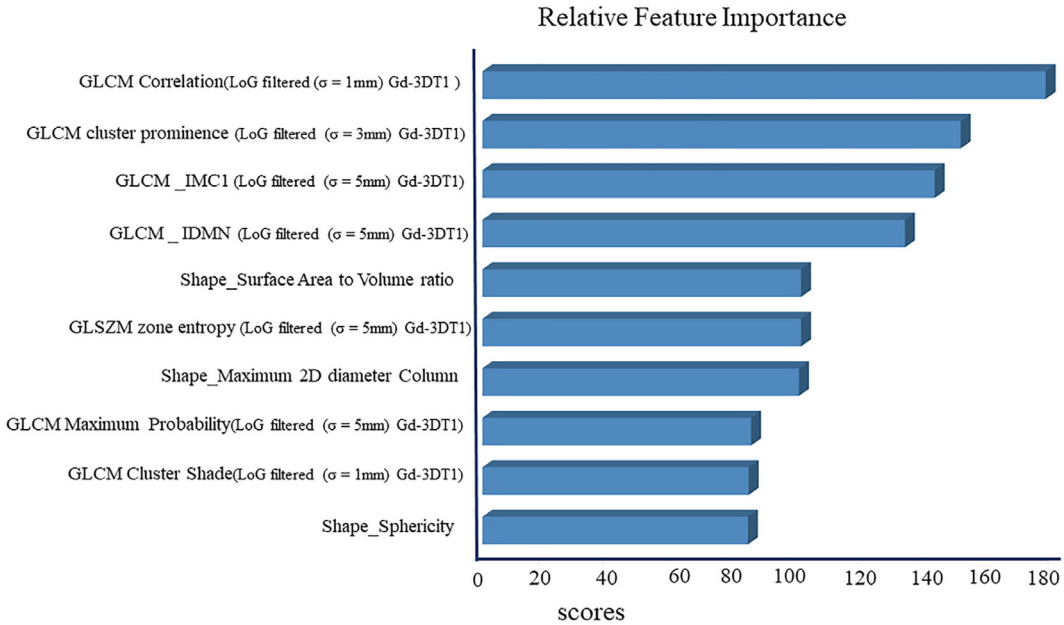

FIGURE 3: Top 10 features that contributed to the prediction from the optimal model. Abbreviations: GLCM_IMC1, gray-level co-occurrence matrix Informational Measure of Correlation 1; GLCM_ IDMN, gray-level co-occurrence matrix Inverse Difference Moment Normalized; GLSZM, gray-level size zone matrix; LoG, Laplacian of Gaussian.

group with co-occurrence group indicated a sign of less infiltration of peripheral tissues in this group. This finding agrees with a previous study finding that a correlation metric could differentiate the pretreatment glioma grades in MR images.[31] According to the definition, IMC1 quantifies the complexity of the texture and IDMN measures the local homogeneity within a ROI. Both features reflect the levels of tissue homogeneity in different ways: a lower IMC1 and a higher IDMN are related to the homogeneous nature of the abnormality.[32,33] Unsurprisingly, our research showed lower values of IMC1 and higher IDMN in the co-occurrence group. Additionally, GLCM cluster prominence, cluster shade, and maximum probability also showed high power in predicting the co-occurrence of IDHmut and MGMTmet. Cluster prominence and cluster shade are the measures of the skewness and uniformity of the GLCM, and a higher cluster shade indicates greater asymmetry about the mean, while the maximum probability of GLCM is the occurrence of the most significant pair of neighboring intensity values.[22] The association between the features and metastatic gastrointestinal stromal tumors has been reported.[34] As regard to the GLSZM zone entropy, this texture measures the uncertainty/randomness in the distribution of zone sizes and gray levels. A higher value indicates more heterogeneity in the texture patterns. Generally, the differences in these features between the two groups in the present study indicate a unique heterogeneous textural pattern in the IDHmut and MGMTmet groups.

Three shape features were also key features in the best model, the feature with the dominant predictive capability was the SAVR. This feature is a potential marker of the infiltrative capacity, while a large SAVR represents the complexity at the tumor margin and an increased interaction between the tumor and surrounding tissues.[35] Briefly, tumors with a low SAVR may be inherently less aggressive than those with a high SAVR, possibly explaining why patients with co-occurrence of IDHmut and MGMTmet have a longer survival than those with other status.[35] The maximum 2D diameter column is the largest pairwise Euclidean distance between the tumor surface mesh vertices, and the sphericity is defined as the measure of roundness of the shape region relative to a circle. These two features also contributed the most to the model for predicting IDH status in a previous study.[36]

Moreover, we observed that the patients with co-occurrence of IDHmut and MGMTmet were significantly younger than those with other status. It has been proved that IDHmut is associated with younger age and rarely occurs in patients above the age of 65.[37] Our results are consistent with this study and suggest that age might also be an important demographic factor in gliomas with co-occurrence of IDHmut and MGMTmet. Additionally, we found that the ratio of lower grade (II + III) glioma in co-occurrence of IDHmut and MGMTmet group was higher than in the other status group. This finding was in line with prior studies, which showed IDHmut is more commonly seen in Grade II and III gliomas.[3,38]

Limitations

First, the relatively small sample size is a major limitation, although radiomics can be performed with as few as 100 patients,[9] the inclusion of more patients with different scanner should provide more power and is a better choice in the future. Second, we had no independent external validation dataset, which would be the best strategy to deal with overfitting.[39] On the other hand, we utilized an inner 10-fold cross validation and an outer 4-fold cross validation to validate our models' predictive performance and minimize the potential bias. Third, we did not extract features from advanced MR sequences, such as diffusion tensor imaging, dynamic susceptibility contrast perfusion-weighted imaging, and proton MR spectroscopy, which may provide valuable information beyond conventional sequences; studies combining advanced MRI data may lead to an improvement in the model performances. Finally, we focused only on predicting the most meaningful group (IDHmut and MGMTmet), which can provide precise prognosis and help early choice of treatment in patients with glioma; in the future, as the amount of dataset enlarged, classifying a case into four groups (IDHmut and MGMTmet, IDHwt and MGMTunmet, IDHwt and MGMTmet, IDHmut and MGMTunmet) is worth exploring.

CONCLUSION

TPOT can be applied as a data-driven approach to identify optimal radiomics-based ML models to predict the co-occurrence of IDHmut and MGMTmet in gliomas with high accuracy. This method has the potential to serve as an alternative to invasive tissue sampling and may aid in the preoperative diagnosis and clinical decision-making for patients with gliomas.

ACKNOWLEDGMENTS

This research received funding from the Sichuan Provincial Foundation of Science and Technology (Grant Nos. 2019YFS0428, 2013SZ0047), the National Natural Science Foundation of China (Grant Nos. 81974278, 81621003, 81371528, 82027808) and the Foundation of the National Research Center of Geriatrics (Grant No. Z2018A07).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

References

1. Lapointe S, Perry A, Butowski NA. Primary brain tumours in adults. Lancet. 2018;392:432–46.

2. Stupp R, Brada M, van den Bent MJ, Tonn JC, Pentheroudakis G, ESMO Guidelines Working Group. High-grade glioma: ESMO clinical

practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2014;25 (Suppl 3):iii93–iii101.

3. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W, et al. IDH1 and IDH2 mutations in gliomas. N Engl J Med. 2009;360:765–73.
4. Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med. 2005;352:997–1003.
5. Weller M, Stupp R, Hegi ME, van den Bent M, Tonn JC, Sanson M, et al. Personalized care in neuro-oncology coming of age: why we need MGMT and 1p/19q testing for malignant glioma patients in clinical practice. Neuro Oncol. 2012;14 (Suppl 4):iv100–8.
6. Molenaar RJ, Verbaan D, Lamba S, Zanon C, Jeuken JWM, Boots-Sprenger SHE, et al. The combination of IDH1 mutations and MGMT methylation status predicts survival in glioblastoma better than either IDH1 or MGMT alone. Neuro Oncol. 2014;16:1263–73.
7. Leu S, von Felten S, Frank S, Vassella E, Vajtai I, Taylor E, et al. IDH/MGMT-driven molecular classification of low-grade glioma is a strong predictor for long-term survival. Neuro Oncol. 2013;15:469–79.
8. Sottoriva A, Spiteri I, Piccirillo SG, Touloumis A, Collins VP, Marioni JC, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc Natl Acad Sci U S A. 2013;110:4009–14.
9. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278:563–77.
10. Zhou H, Vallières M, Bai HX, Su C, Tang H, Oldridge D, et al. MRI features predict survival and molecular markers in diffuse lower-grade gliomas. Neuro Oncol. 2017;19:862–70.
11. Ren Y, Zhang X, Rui W, Pang H, Qiu T, Wang J, et al. Noninvasive prediction of IDH1 mutation and ATRX expression loss in low-grade gliomas using multiparametric MR radiomic features. J Magn Reson Imaging. 2019;49:808–17.
12. Wei J, Yang G, Hao X, Gu D, Tan Y, Wang X, et al. A multi-sequence and habitat-based MRI radiomics signature for preoperative prediction of MGMT promoter methylation in astrocytomas with prognostic implication. Eur Radiol. 2019;29:877–88.
13. Zhang B, Chang K, Ramkissoon S, Tanguturi S, Bi WL, Reardon DA, et al. Multimodal MRI features predict isocitrate dehydrogenase genotype in high-grade gliomas. Neuro Oncol. 2017;19:109–17.
14. Wang S, Summers RM. Machine learning and radiology. Med Image Anal. 2012;16:933–51.
15. Hutter F, Lücke J, Schmidt-Thieme L. Beyond manual tuning of hyperparameters. Künstliche Intell. 2015;29:329–37.
16. Olson RS, Bartley N, Urbanowicz RJ, Moore JH. Evaluation of a tree-based pipeline optimization tool for automating data science. *Proceedings of the genetic and evolutionary computation conference*. New York, NY, USA: Association for Computing Machinery; 2016. p. 485–92.
17. Olson RS, Urbanowicz RJ, Andrews PC, Lavender NA, Moore JH. Automating biomedical data science through tree-based pipeline optimization. *European conference on the applications of evolutionary computation*. Cham: Springer; 2016. p. 123–37.
18. Sohn A, Olson RS, Moore JH. Toward the automated analysis of complex diseases in genome-wide association studies using genetic programming. *Proceedings of the genetic and evolutionary computation conference*. New York, NY, USA: Association for Computing Machinery; 2017. p. 489–96.
19. Gibson E, Li W, Sudre C, Fidon L, Shakir DI, Wang G, et al. NiftyNet: a deep-learning platform for medical imaging. Comput Methods Programs Biomed. 2018;158:113–22.
20. Kavur AE, Gezer NS, Barış M, Şahin Y, Özkan S, Baydar B, et al. Comparison of semi-automatic and deep learning-based automatic methods for liver segmentation in living liver transplant donors. Diagn Interv Radiol. 2020;26:11–21.
21. Wang G, Li W, Ourselin S, Vercauteren T. Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. *International MICCAI brainlesion workshop*. Cham: Springer; 2017. p. 178–90.
22. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res. 2017;77:e104–7.
23. Koza JR. *Genetic programming: on the programming of computers by means of natural selection*. Vol 1. Cambridge, Massachusetts: MIT Press; 1992.
24. Olson RS, Moore JH. TPOT: a tree-based pipeline optimization tool for automating machine learning. In: Hutter F, Kotthoff L, Vanschoren J, editors. *Automated machine learning: methods, systems, challenges*. Cham: Springer International Publishing; 2019. p. 151–60.
25. Fortin FA, FMd R, Gardner MA, Parizeau M, Gagne C. DEAP: evolutionary algorithms made easy. J Mach Learn Res. 2012;13:2171–5.
26. Orlenko A, Kofink D, Lyytikäinen LP, Nikus K, Mishra P, Kuukasjärvi P, et al. Model selection for metabolomics: predicting diagnosis of coronary artery disease using automated machine learning. Bioinformatics. 2020;36:1772–8.
27. Dafflon J, Pinaya WHL, Turkheimer F, Cole JH, Leech R, Harris MA, et al. An automated machine learning approach to predict brain age from cortical anatomical measures. Hum Brain Mapp. 2020;41:3555–66.
28. Keek SA, Leijenaar RT, Jochems A, Woodruff HC. A review on radiomics and the future of theranostics for patient selection in precision medicine. Br J Radiol. 2018;91:20170926.
29. Larue RT, Defraene G, De Ruysscher D, Lambin P, van Elmpt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. Br J Radiol. 2017;90:20160665.
30. Neycenssac F. Contrast enhancement using the Laplacian-of-a-Gaussian filter. CVGIP-Graph Model IM. 1993;55:447–63.
31. Vamvakas A, Williams SC, Theodorou K, Kapsalaki E, Fountas K, Kappas C, et al. Imaging biomarker analysis of advanced multiparametric MRI for glioma grading. Phys Med. 2019;60:188–98.
32. Chen X, Wei X, Zhang Z, Yang R, Zhu Y, Jiang X. Differentiation of true-progression from pseudoprogression in glioblastoma treated with radiation therapy and concomitant temozolomide by GLCM texture analysis of conventional MRI. Clin Imaging. 2015;39:775–80.
33. Huang SY, Franc BL, Harnish RJ, Liu G, Mitra D, Copeland TP, et al. Exploration of PET and MRI radiomic features for decoding breast cancer phenotypes and prognosis. NPJ Breast Cancer. 2018;4:24.
34. Fu J, Fang MJ, Dong D, Li J, Sun YS, Tian J, et al. Heterogeneity of metastatic gastrointestinal stromal tumor on texture analysis: DWI texture as potential biomarker of overall survival. Eur J Radiol. 2020;125:108825.
35. Cordova JS, Gurbani SS, Holder CA, Olson JJ, Schreibmann E, Shi R, et al. Semi-automated volumetric and morphological assessment of glioblastoma resection with fluorescence-guided surgery. Mol Imaging Biol. 2016;18:454–62.
36. Zhou H, Chang K, Bai HX, Xiao B, Su C, Bi WL, et al. Machine learning reveals multimodal MRI patterns predictive of isocitrate dehydrogenase and 1p/19q status in diffuse low- and high-grade gliomas. J Neurooncol. 2019;142:299–307.
37. Wirsching HG, Galanis E, Weller M. Glioblastoma. Handb Clin Neurol. 2016;134:381–97.
38. Cohen AL, Holmen SL, Colman H. IDH1 and IDH2 mutations in Gliomas. Curr Neurol Neurosci Rep. 2013;13:345–5.
39. Koçak B, Durmaz E, Ateş E, Kılıçkesmez Ö. Radiomics with artificial intelligence: a practical guide for beginners. Diagn Interv Radiol. 2019;25:485–95.