Computers in Biology and Medicine 132 (2021) 104320

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

Radiomics-based machine learning model for efficiently classifying transcriptome subtypes in glioblastoma patients from MRI

Nguyen Quoc Khanh Le [a,b,c,*], Truong Nguyen Khanh Hung [d,e], Duyen Thi Do [f], Luu Ho Thanh Lam [d,g], Luong Huu Dang [h], Tuan-Tu Huynh [i,j]

[a] Professional Master Program in Artificial Intelligence in Medicine, College of Medicine, Taipei Medical University, Taipei, 106, Taiwan
[b] Research Center for Artificial Intelligence in Medicine, Taipei Medical University, Taipei, 106, Taiwan
[c] Translational Imaging Research Center, Taipei Medical University Hospital, Taipei, 110, Taiwan
[d] International Master/Ph.D. Program in Medicine, College of Medicine, Taipei Medical University, Taipei, 110, Taiwan
[e] Orthopedic and Trauma Department, Cho Ray Hospital, Ho Chi Minh City, 70000, Viet Nam
[f] Graduate Institute of Biomedical Informatics, Taipei Medical University, Taipei, 106, Taiwan
[g] Children's Hospital 2, Ho Chi Minh City, 70000, Viet Nam
[h] Department of Otolaryngology, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, 70000, Viet Nam
[i] Department of Electrical Engineering, Yuan Ze University, No. 135, Yuandong Road, Zhongli, 320, Taoyuan, Taiwan
[j] Department of Electrical Electronic and Mechanical Engineering, Lac Hong University, No. 10, Huynh Van Nghe Road, Bien Hoa, Dong Nai, 76120, Viet Nam

ARTICLE INFO

Keywords:
Radiogenomics
Glioblastoma
Neuroimaging
Transcriptome subtypes
Radiomics biomarker
XGBoost
Artificial intelligence
Magnetic resonance imaging

ABSTRACT

Background: In the field of glioma, transcriptome subtypes have been considered as an important diagnostic and prognostic biomarker that may help improve the treatment efficacy. However, existing identification methods of transcriptome subtypes are limited due to the relatively long detection period, the unattainability of tumor specimens via biopsy or surgery, and the fleeting nature of intralesional heterogeneity. In search of a superior model over previous ones, this study evaluated the efficiency of eXtreme Gradient Boosting (XGBoost)-based radiomics model to classify transcriptome subtypes in glioblastoma patients.

Methods: This retrospective study retrieved patients from TCGA-GBM and IvyGAP cohorts with pathologically diagnosed glioblastoma, and separated them into different transcriptome subtypes groups. GBM patients were then segmented into three different regions of MRI: enhancement of the tumor core (ET), non-enhancing portion of the tumor core (NET), and peritumoral edema (ED). We subsequently used handcrafted radiomics features (n = 704) from multimodality MRI and two-level feature selection techniques (Spearman correlation and F-score tests) in order to find the features that could be relevant.

Results: After the feature selection approach, we identified 13 radiomics features that were the most meaningful ones that can be used to reach the optimal results. With these features, our XGBoost model reached the predictive accuracies of 70.9%, 73.3%, 88.4%, and 88.4% for classical, mesenchymal, neural, and proneural subtypes, respectively. Our model performance has been improved in comparison with the other models as well as previous works on the same dataset.

Conclusion: The use of XGBoost and two-level feature selection analysis (Spearman correlation and F-score) could be expected as a potential combination for classifying transcriptome subtypes with high performance and might raise public attention for further research on radiomics-based GBM models.

1. Introduction

Glioblastoma (GBM) is a term indicating the most aggressive and exceptionally invasive brain tumors and is categorized based on its cell of origin [1]. GBM has an incidence of two to three per 100,000 people worldwide per year, accounting for 52% of all primary brain tumors. Although GBM has a relatively low incidence, its poor prognosis (aggressive and infiltrative growth pattern) renders the curative treatment impossible [2,3]. Despite plenty of efforts worldwide, the treatment of GBM is still considered the most difficult work in clinical

* Corresponding author. Professional Master Program in Artificial Intelligence in Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan.
E-mail address: khanhlee@tmu.edu.tw (N.Q.K. Le).

https://doi.org/10.1016/j.compbiomed.2021.104320
Received 10 December 2020; Received in revised form 5 March 2021; Accepted 5 March 2021
Available online 9 March 2021

oncology. Histopathological analysis of the tumor tissue is the reference standard of a definitive diagnosis of GBM. More recently, tumor genomic characterization (tissue available from the surgical resection) has advanced GBM's clinical evaluation to provide additional predictors of the response and outcome to treatment.

Radiomics is a new approach that has been commonly used to identify the association between clinical symptoms and underlying genetic characteristics [4]. Potentially, tumor phenotypes can be determined by collecting large quantities of features from high-throughput medical radiographic images by breakthrough data characterization [5]. Different subtypes of gliomas have been classified in the previously published works using the radiomics model. For instance, glioma grading has been qualified by using magnetic resonance imaging (MRI) transformed gray-scale invariant textures [6], five significant features [7], wavelet-based features [8], or even deep learning features [9]. For the molecular subtype classification, isocitrate dehydrogenase (IDH) mutation status, 1p/19q codeletion status, TERT promoter status, and O6-methylguanine-DNA methyltransferase (MGMT) methylation status have been served as key biomarkers. For that reason, Lu et al. conducted a radiomics-based machine learning model for three-level classification of gliomas [10]. These molecular subtypes have been defined by mutation status of IDH gene and codeletion status of 1p/19q. For each mutation status itself, radiomics model has been used in predicting the status of IDH [11–14], 1p/19q deletion [12,15], TERT promoter [16, 17], p53 [18], and MGMT [19,20].

In addition to the WHO grade or molecular subtypes, the classification of transcriptome subtypes also plays an essential role in the assessment and treatment of GBM. According to Ref. [21], GBM could be classified into four transcriptome subtypes: classical, mesenchymal, neural, and proneural, respectively. These subtypes include different biological biomarkers such as collective loss in chromosome 10 and chromosome 7 amplification in classical subtypes, the largest occurrence of focal hemizygous deletions in the region at 17q11.2, including NF1 genes in mesenchymal subtypes, PDGFRA aberrations and IDH1 mutations in proneural subtypes, and GABRA1, SYT1, NEFL, and SLC12A5 in neurine subtypes. Recent research [22] has also shown that subcategory genetic abnormalities have the ability to act as predictive markers and therapeutic targets. With the relatively high importance of these transcriptome subtypes in GBM, they are considered a critical factor in temozolomide resistance and poor progression-free survival (PFS). Thus, non-invasive imaging biomarkers for identifying the transcriptome subtypes could pave the way to a better future for GBM treatment with accurate treatment guidance and prognosis.

The question of how to classify the transcriptome subtypes of GBM using MRI features is a challenging task nowadays. Some radiomics studies have been performed to deal with this problem. For example, Saima et al. [23] classified transcriptome subtypes of GBMs with an average accuracy of 71%. Macyszyn et al. [24] achieved a multi-classification accuracy of 76% when creating a multiparametric MRI-based model with 105 patients and 60 diverse features. Lee et al. [25] solved this multi-classification problem by separating it into four binary classifications. In an attempt to find efficient radiomics features for this problem, Yang et al. [26] confirmed that textual features might play an important role in classifying GBM transcriptome subtypes. For more detail on a specific transcriptome subtype, Kourosh et al. [27] retrieved 46 GBM patients for identifying the mesenchymal molecular subtype only. Nicholas et al. [28] created a radiogenomics-based pilot study for classification between mesenchymal and classical subtypes.

Previous works have proposed some promising machine learning models and radiomics signatures for classifying transcriptome subtypes in GBMs. However, the performance is still not satisfactory and it is challenging to look at the other models that could help to enhance the classification performance. Aiming at addressing this problem, this study suggested an efficient radiomics signature by using eXtreme Gradient Boosting (XGBoost) and two-level (Spearman correlation and F-score) feature selection.

2. Materials and methods

Fig. 1 demonstrates our proposed radiomics framework and systematic is described in the following subsections.

2.1. Patients cohort

Our patient cohort has been obtained from The Cancer Imaging Archive (TCIA) [29], a public resource for a large collection of cancer medical images that can be downloaded. Since we aimed to classify the transcriptome subtypes of GBM, we selected the patient data from TCGA-GBM project [30]. It is a project including 262 participants (collected from eight institutions in U.S.A. and Italy) with their pre-operative multimodal magnetic resonance imaging (MRI) images. The selected MRI modalities were T1-weighted pre-contrast (T1), T1-weighted post-contrast (T1-Gd), T2, and T2-FLAIR. The subject counts of the transcriptome subtypes are 20, 34, 11, and 21 for classical, mesenchymal, neural, and proneural, respectively [31]. Meanwhile, we used these 86 patients with transcriptome subtypes information to be included in our next analyses.

Most radiomics guidelines suggest an external validation set to believe its reproducibility and repeatability [32]. Therefore, we also retrieved another GBM set from TCIA to validate our predictive performance as well as our radiomics features. The selected set was Ivy Glioblastoma Atlas (IvyGAP) [33] that contained genomic alterations and gene expression patterns for GBM patients. There were 34 GBM patients that contained transcriptome subtype information with the number of 15, 12, 12, and 10 for classical, mesenchymal, neural, and proneural, respectively. Notice here is that few patients were marked as belonging to two or even greater than two transcriptome subtypes together, and we used all of this information. Therefore, we used these 34 GBM patients as our external validation cohort.

2.2. MRI segmentation and radiomics features

The medical images have been segmented consistently based on the criteria of BraTS challenge [34]. On the basis of this criterion, GBM MRIs may be broken down into three distinct regions: (1) the enhancement of the tumor core (ET) is distinguished by areas with hypertensity in T1-Gd relative to both T1 and normal/healthy white matter (WM) in T1-Gd; (2) the non-enhancing portion of the tumor core (NET) reveals hypo-intense presence in T1-Gd compared to both T1 and normal/healthy WM in T1-Gd; and (3) finally, hyper-intense signal in T2-FLAIR volumes is used to characterize peritumoral edema (ED).

BraTS challenge also released the benchmark segmentations [35] that contained the standard ground truth and radiomics features of TCGA-GBM project. In detail, before segmentations, a preprocessing step was performed on mMRI volumes including re-orienting to the LPS (left-posterior-superior), co-registered to the same T1 anatomic template, resampling to 1 mm3 voxel resolution, and skull-stripping (using FMRIB Software Library (FSL) [36]). Next, GLISTRboost [37], an advanced version of GLioma Image SegmenTation and Registration (GLISTR) [38] was used to produce the segmentation labels. An important notice shown in the original article is that they did not use any non-parametric, non-uniform intensity normalization algorithm to standardize the MRIs before segmenting. The reason is that they observed that the application of such an algorithm obliterated the T2-FLAIR signal.

An example of how to segment a GBM patient into three different regions has been shown in Fig. 2. Next, the original study also developed Cancer Imaging Phenomics Toolkit (CaPTk) [39] for extracting the radiomics features from images as well as segmentation. All the radiomics features using in this study were consistent with 174 standardized features that have been introduced by the Image Biomarker Standardisation Initiative (ISBI) [40]. The radiomics features included the intensity information, image derivative, geodesic information, texture

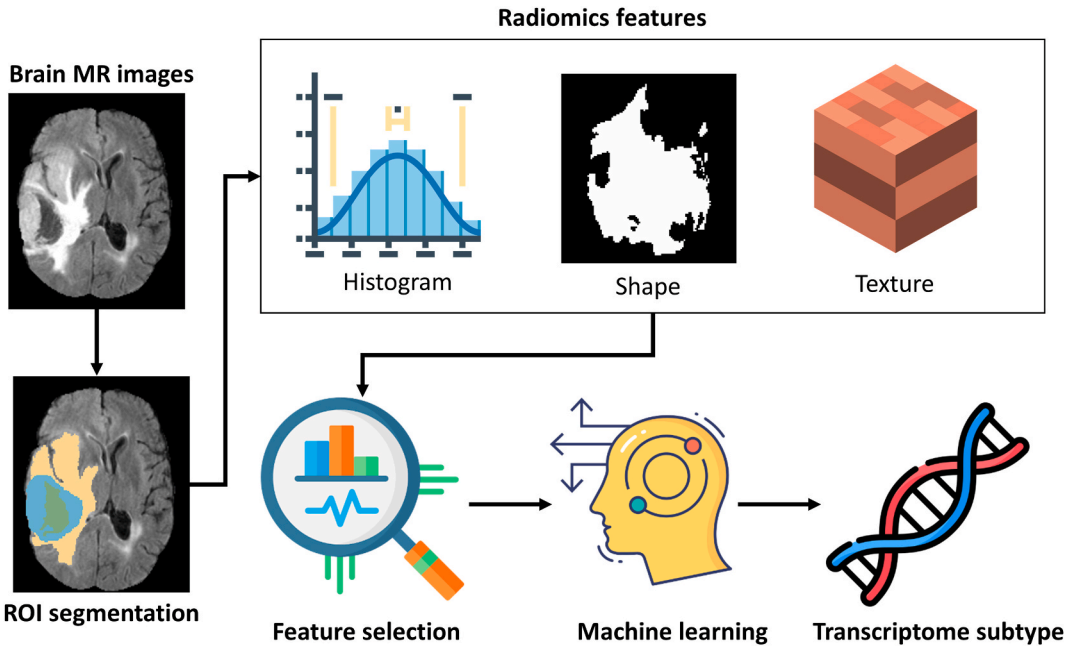

Fig. 1. A flowchart of this study. It comprises of four sub-processes: data cohort collection, radiomics feature extraction, feature selection, and machine learning implementation.

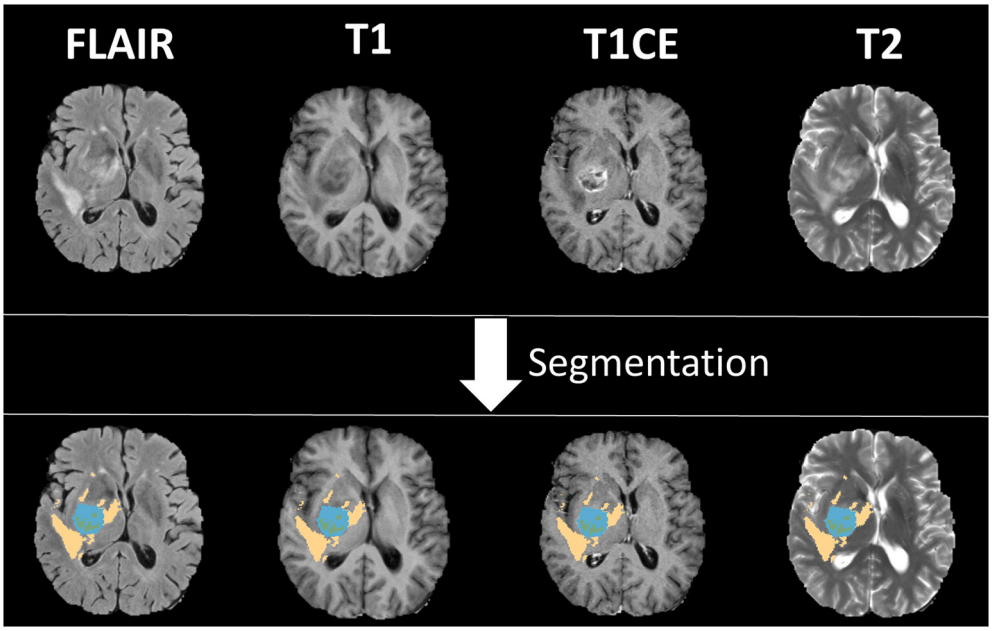

Fig. 2. An example of segmenting GBM patients on multimodal MRI images (patient ID: TCGA-06-5413, neural transcriptome subtype).

features, and GLISTR posterior probability maps. Bin discretization of 64 have been chosen when extracting radiomics features. We then re-used 704 radiomics features that had been extracted and fed into our machine learning model to evaluate the predictive performance.

2.3. *Two-level radiomics feature selection*

The most important problem in radiomics-based machine learning model is the dimension of data. Since a big number of radiomics are used as feature sets in machine learning models, it will expand the computational complexity of the model as well as the overfitting problem. Therefore, an essential task that needs to be performed is to reduce the number of features. There are many common methods for solving this task (such as correlation-based, information theory-based, etc.), and this study would like to propose a two-level feature selection technique on it. First, we performed a Spearman correlation test to evaluate the significant features for transcriptome subtype classification. A feature was significant if it had a high correlation coefficient (>0.8) among any of other features. Consequently, we would like to test the possibility of F-score analysis in extracting the best features among these aforementioned ones. The idea behind F-score is to simply check the differences in the performance results between two sets of values [41] as follows:

$$F(i) = \frac{\left(\bar{x}_i^{(+)} - \bar{x}_i\right)^2 + \left(\bar{x}_i^{(-)} - \bar{x}_i\right)^2}{\frac{1}{n_+ - 1}\sum_{k=1}^{n+}\left(x_{k,i}^{(+)} - \bar{x}_i^{(+)}\right)^2 + \frac{1}{n_- - 1}\sum_{k=1}^{n-}\left(x_{k,i}^{(-)} - \bar{x}_i^{(-)}\right)^2}$$

where n_+ is the number of positive instances and n_- is the number of negative instances. In addition, $\bar{x}_i$, $\bar{x}_i^{(+)}$, and $\bar{x}_i^{(-)}$ are the averages of the ith feature of the entire, positive, and negative data sets, respectively; $x_{k,i}^{(+)}$ is the ith feature of the kth positive instance; and $x_{k,i}^{(-)}$ is the ith feature of the kth negative instance. F-score analysis has been applied in the previous study to search for the optimal radiomics features in GBM binary classification [42]. Here we extended the F-score analysis into a multi-classification in which we measured the accuracy of a test from four sets of real numbers (corresponding to four classes of GBMs). We firstly calculated F-score values for all significant radiomic features, and then ranked them descending to see the top-rank important features. After that, we fed the important features one-by-one increasingly to our model to test the cut-off value for feature selection. The point that has achieved the best value will be chosen as our optimum cut-off point for the selection of the F-score features. To display the performance results at different number of features, we used recursive feature elimination (RFE) technique to present. It is an efficient approach for displaying the training performance after increasing the number of features one by one.

2.4. Machine learning implementation

Different machine-learning models were implemented in this research to see which algorithms perform well for these forms of radiomics. These included k-nearest neighbors (kNN), Naïve Bayes, Random Forest, support vector machine (SVM), and XGBoost. Our machine learning models were implemented using Python programming language and scikit-learn library [43]. Each machine learning algorithm needed a process namely hyperparameter optimization to enable the best results.

2.5. Statistical analysis and measurement metrics

To deal with a multiple classification problem, we treated it as multiple binary classifications and then calculated the individual prediction metrics. Because of the limited data, we used leave-one-out cross-validation (LOOCV) as evaluation method to validate the complete output. In this method, each sample is used as a test collection, while the other samples are used for training and the accuracy stated is the mean of the accuracy of all tests. After creating the model, we also had an external validation set to evaluate the predictive performance of our model on unseen data. We adopted different performance metrics in the prediction model such as accuracy, receiver operating characteristic (ROC) curve, and Area under the curve (AUC) to stratify the training data to improve machine learning-based GBM subtype classification. Among them, ROC curve and AUC have been used to overcome the imbalance dataset problem by showing the overall performance at different threshold points. These measurement metrics are common in machine learning and they have been successfully used in many biomedical works with high evidence [42,44].

3. Results

3.1. Patient's clinical characteristics

Table 1 shows the patient's characteristics of our training and validation cohort. Our training data contained 20, 34, 11, and 21 patients for classical, mesenchymal, neural, and proneural transcriptome groups, respectively. Most of our patients had IDH1-wildtype in their genomic information, thus this study could be treated as a transcriptome subtype classification among IDH1-wildtype GBM patients. According to the data characteristics, patients with proneural subtype are younger than the other subtypes (average age of 54.6 compared to 61.4, 59.7, and 62.8 from the others). There are also many differences in gender among four transcriptome subtypes, with a higher number of males in mesenchymal and neural subtypes. On the other hand, there are a higher number of female in classical transcriptome subtype. Furthermore, there are not many differences in MGMT methylation status among different transcriptome subtypes and methylation classes. Data statistics also showed a consistent level between the training data and validation data, which means that we could use this IvyGAP dataset as a quality external dataset to validate the performance results.

Table 1
Patients' characteristics of our training and validation cohort.

	Training (n = 86)	Validation (n = 34)
Age(mean ± SD, years)	59.22 ± 12.7	59.6 ± 10.3
Gender		
Male	57	17
Female	29	17
Transcriptome subtype		
Classical	20	15
Mesenchymal	34	12
Neural	11	12
Proneural	21	10
IDH1 status		
Wildtype	68	31
Mutant	18	3
MGMT status		
Methylated	22	12
Unmethylated	27	20

(MGMT: O6 methylguanine DNA methyltransferase).

3.2. Generating the efficient MRI radiomics signatures

Two-level feature selection has been applied to find the radiomics signatures of GBM transcriptome subtype classification. As a detail, we first conducted the Spearman correlation test to statistically look at the significant features that might affect the predictive performance. After this step, we found that there were 470 significant radiomics features (with correlation coefficient > 0.8). Next, we applied F-score analysis in these features and determined some top-rank features with high F-score values such as VOLUME_NET_OVER_ED (f-score = 0.18037), TEXTURE_GLCM_ET_FLAIR_Entropy (F-score = 0.15062), and TEXTURE_GLSZM_ET_FLAIR_SZHGE (F-score = 0.1432). All the detailed information on radiomics features and their F-score values are shown in Supplementary Table S1. After that, our RFE curve (Fig. 3) shows that we can use the first 13 features (with optimum cut-off F-score of 0.1085) as input into the model to get the highest results. The aforementioned features might be our innovative signature to help to classify GBM transcriptome subtypes with high performance and the use of features as little as possible.

3.3. Model performance

The first experiments were conducted to look at the optimal parameters of individual machine learning classifiers. The same setting was applied to each algorithm; and then the ranges of all searched hyperparameters as well as the optimal values were given in Table 2. For the experimental results, our radiomics-based classifiers reached an average accuracy of 69.8%, 75%, 77.3%, 71.5%, and 80.2% for kNN, Naïve Bayes, Random Forest, SVM, and XGBoost respectively (as shown in Table 3). The average sensitivity and specificity of XGBoost (51.9% and 87.5%, respectively) were also superior to the others. Wilcoxon tests have been also conducted to see the significant improvements of XGBoost compared to the other classifiers. We are able to observe that the XGBoost had improvements in most measurement metrics compared to the other classifiers. Among the individual classifiers of four subtypes,

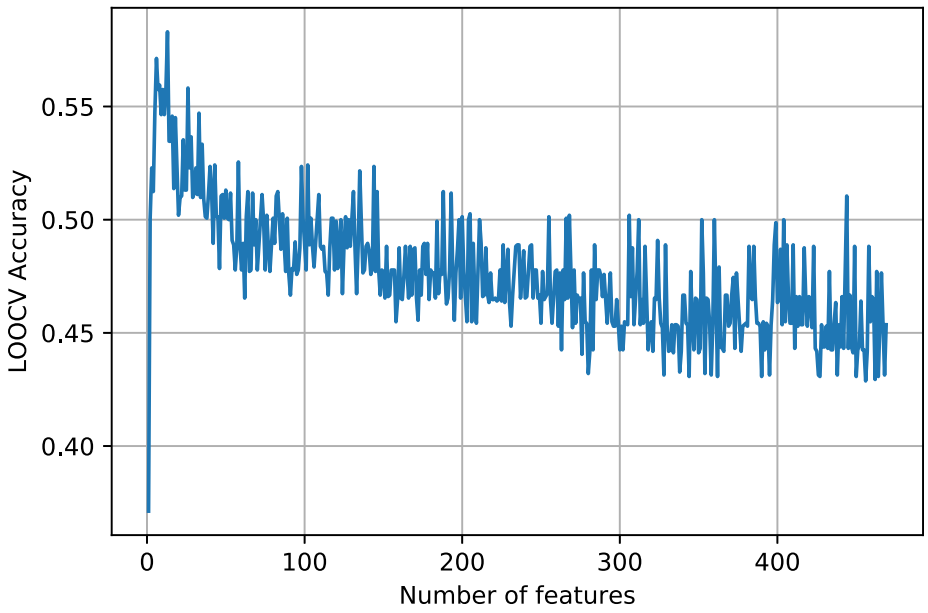

Fig. 3. RFE curve for classifying GBM transcriptome subtypes using different number of features. The best performance was reached with the use of the first 13 features after Spearman correlation test and F-score analysis.

Table 2
Hyperparameters optimization for each machine learning algorithm.

Machine learning	Ranges of hyperparameters	Optimal value
kNN	n_neighbors = [1,2,3,..,10]	1
	weights = [uniform, distance]	uniform
	metric = [euclidean, manhattan, minkowski]	minkowski
Random Forest	max_depth = [80, 90, 100, 110]	110
	max_features = [2, 3]	3
	min_samples_leaf = [3, 4, 5]	4
	min_samples_split = [8, 10, 12]	8
	n_estimators = [100, 200, 300, 1000]	100
SVM	C = [0.001, 0.01, 0.1, 1, 10]	1
	gammas = [0.001, 0.01, 0.1, 1]	0.001
	kernels = [rbf, linear]	rbf
XGBoost	min_child_weight = [1, 5, 10]	1
	gamma = [0.5, 1, 1.5, 2, 5]	1
	subsample = [0.6, 0.8, 1]	0.8
	colsample_bytree = [0.6, 0.8, 1]	1
	max_depth = [3, 4, 5]	4

Table 3
Leave-One-Out cross-validation among different subtypes and machine learning algorithms.

Machine learning	Subtype	Sensitivity	Specificity	Accuracy
kNN	Classical	50.0	80.3	73.3
	Mesenchymal	67.6	42.3	52.3
	Neural	9.1	94.7	83.7
	Proneural	0.0	92.3	69.8
Naïve Bayes	Classical	45.0	81.8	73.3
	Mesenchymal	61.8	63.5	62.8
	Neural	36.4	88.0	81.4
	Proneural	42.9	95.4	82.6
Random Forest	Classical	25.0	90.9	75.6
	Mesenchymal	85.3	50.0	64.0
	Neural	0.0	93.3	81.4
	Proneural	61.9	96.9	88.4
SVM	Classical	15.0	97.0	77.9
	Mesenchymal	82.4	28.8	50.0
	Neural	0.0	98.7	86.0
	Proneural	28.6	86.2	72.1
XGBoost	Classical	25.0	84.8	70.9
	Mesenchymal	70.6*	75.0*	73.3*
	Neural	45.5*	94.7	88.4*
	Proneural	66.7*	95.4	88.4

All machine-learning algorithms have used the optimal hyperparameters. * shows the significant improvements after Wilcoxon tests.

the model held good potential on identifying mesenchymal (sensitivity of 70.6%) and proneural (sensitivity of 66.7%) subtypes. The worst performance came from the other two subtypes; however, it was also an acceptable level for a challenging multi-classification problem and the specificities were high enough to show the efficiency of the models.

In order to see the performance at different threshold levels, it is important to present the ROC curves and AUCs. We provided ROC curves of our XGBoost models on different transcriptome subtypes one by one (Fig. 4). According to the information shown in this figure, we again observed that our model worked well with the use of signature radiomics features (AUCs of four transcriptome subtypes reached 0.711, 0.763, 0.745, and 0.854, respectively. Therefore, it strongly suggests that we could incorporate these 13 features and XGBoost algorithm to classify GBM transcriptome subtypes with high performance.

3.4. Validation results

To ensure the efficiency of our final model, we used a validation dataset to evaluate the predictive performance of this model and radiomics signature. In detail, we used the same 13 radiomics features from 39 IvyGAP patients [33] and inserted them into our model. Fig. 5 then shows the comparative performance between training and validation dataset in terms of sensitivity, specificity, and accuracy. We observed that there were consistent between these two sets and this ensured that our model was reliable and did not contain much overfitting. It also means that our 13-signature features might be significant in classifying transcriptome subtypes of GBM patients.

4. Discussions

Medical imaging features are deemed as the cornerstone of the algorithm for managing and evaluating the further response of cancer. Many disease characteristics, especially in the field of oncology, have been uncovered thanks to the extraction of hundreds of quantitative radiomics features including CT, PET, and MR scans. Recently, with the increasing opening of public medical datasets [29], it is possible to

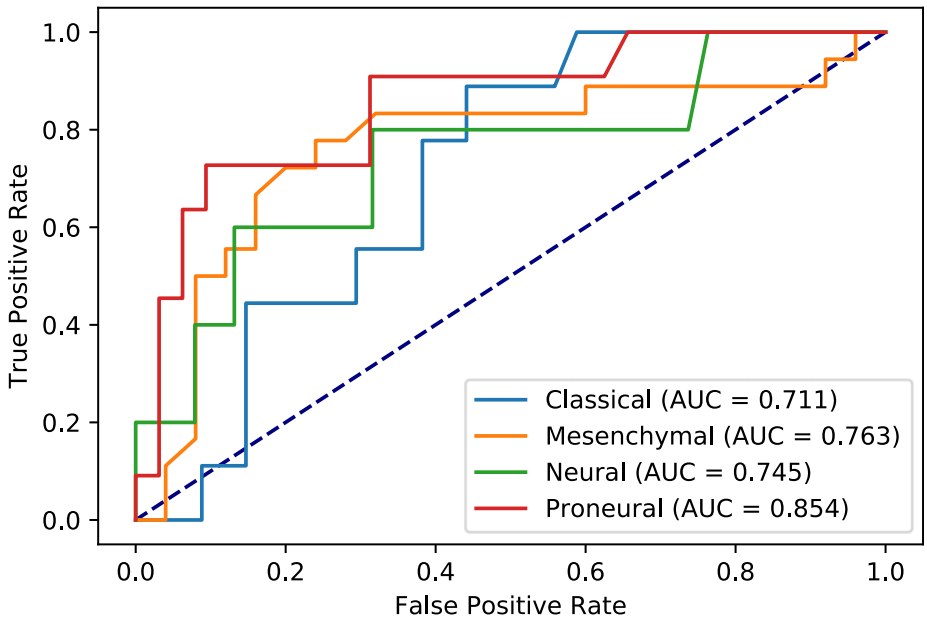

Fig. 4. ROC curve analysis of classifying GBM transcriptome subtypes using XGBoost on 13 top-rank radiomics features.

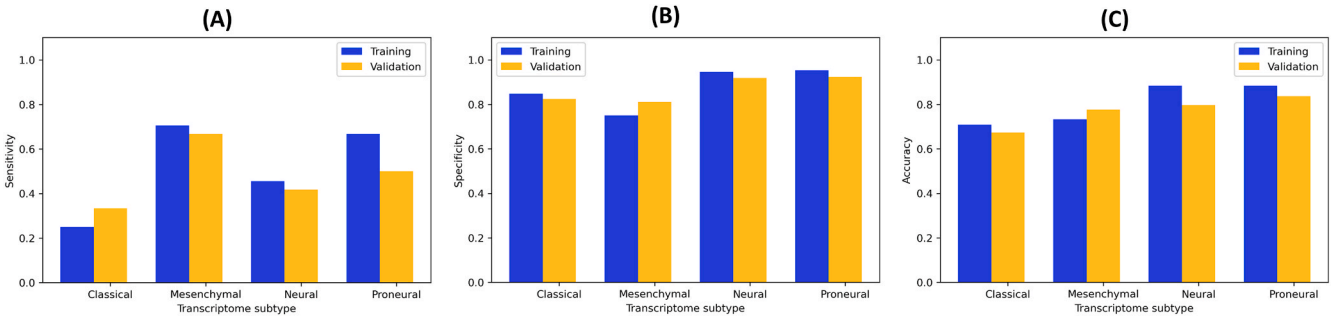

Fig. 5. Comparative performance results among training and validation data.

greatly enhance the predictive performance of radiomics-based machine learning problems in general and GBM transcriptome subtypes in particular. In this study, the capacity of multimodal fusion and two-level feature selection techniques to identify GBM subtypes was demonstrated through investigations of the TCGA-GBM [30] dataset. The final model was also validated on an external dataset (IvyGAP) [33] with valuable results.

Traditionally, a lot of feature selection techniques have been applied in radiomics-based machine learning models to classify GBM transcriptome subtype such as sequential forward feature selection [23], individual predictive test [24], spatial point pattern analysis [25], or Mann–Whitney test [27]. This study, therefore, provided evidence on applying another efficient technique for this classification purpose. Our 13-signature features might differentiate from the previous methods and they could help to generate a promising performance in multi-classification. Moreover, via using a combination of Spearman and F-score analysis, our radiomics features are reliable and could be interpretable easily.

Our study also analyzed the efficiencies of different machine learning models in learning radiomics features. The results of this study (Table 3) showed that XGBoost is superior to other methods to classify transcriptome subtypes for patients with gliomas. This finding is consistent with the previous study where XGBoost is also the optimal model for glioblastoma studies such as [42,45]. Therefore, this study again emphasizes the significance of ensemble learning (especially XGBoost) in radiomics. Further radiomics studies could consider it as the first choice in learning features efficiently.

Different radiomics-based GBM studies used different patient cohorts, thus it is challenging to have a comparison among different studies fairly and accurately. However, to have a relative view on the effectiveness of our method, we also provided the comparison results among our model and previous works on the same transcriptome subtype classification. There were some published works on the classification of GBM transcriptome subtypes with promising performance such as [23,26]. Their performance results were retrieved to support the comparison purpose. For instance, compared to the work of [23] on the same TCIA dataset, our model has improved at about 5% on average accuracy. Focusing on detailed performance results among the individual subtypes, our model achieved better performance than [26] in mesenchymal, neural, and proneural subtypes. The comparative performance observed that at the same and fair level of comparison, our model's performance had a little bit better than the others had.

In addition, our optimal set of 13 radiomics features might attract much attention to GBM research. This radiomics signature combing with XGBoost classifier could hold potential in classifying transcriptome subtypes with high performance. Another finding is that in our biomarker set containing 13 features, most of them were from textual analysis. It has been consistent with previous works on radiomics-based GBM models such as prediction of MGMT methylation status [46,47], IDH1 mutation, and 1p/19q-codeletion status [48,49]. In addition, it is observed that the transformation of the wavelet (i.e. GLSZM) appeared mainly in the essential set of features. Therefore, with the joint of either

wavelet transform features or NET-hit features, it becomes far more possible for a computational model to classify transcriptome non-invasively and pre-operatively in GBM patients.

Despite the positive results of this research, it is also important to look at the limitations of this study. First, to improve the quality of this research, it is necessary to increase the sample sizes for assessing the generalization of our model. However, by repeating the training many times via LOOCV, this study so that somehow addressed this limitation. In addition, we have released the detailed information of the selected radiomics features; it would be useful for other centers to validate the results in further studies. Second, the updated TCGA molecular subtype of GBM has removed the neural subtype, which is more likely to be the adjacent brain tissues other than GBM tumor mass. Thus, further transcriptome subtype-based studies could exclude the neural subtype in TCGA-GBM project to avoid insufficient data. Finally, as shown in previous radiomics study [50], repeatability of radiomic features is affected due to test–retest and image registration. Thus, yet importantly, image normalization or standardization could be performed to the heterogeneous data before applying radiomics analyses.

5. Conclusion

Through investigations on the dataset of 86 GBM patients, this study investigates the role of two-level feature selection and radiomics-based XGBoost model in classifying transcriptome subtypes in GBM patients. Multi-classification has been resolved by treating it as individual binary classifications. Thereafter, our feature selection analysis with 13 features can help to boost the predictive performance and stability of our model. The final predictive model that used 13 features as input achieved accuracies of 70.9%, 73.3%, 88.4%, and 88.4% for four subtypes, respectively. Our AI-based radiomics model also shows a significant performance on an external validation dataset. This study shows that combining Spearman correlation, F-score and XGBoost functionality is a promising approach for the classification of GBM transcriptome subtypes. This finding might be replicated to enhance the predictive performance of further radiogenomics studies.

Funding

This work was supported by the Research Grant for Newly Hired Faculty, Taipei Medical University [grant number: TMU108-AE1-B26] and Higher Education Sprout Project, Ministry of Education, Taiwan [grant number: DP2-109-21121-01-A-06].

Declaration of competing interest

The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2021.104320.

References

[1] F. Hanif, et al., Glioblastoma multiforme: a review of its epidemiology and pathogenesis through clinical presentation and treatment, Asian Pac. J. Cancer Prev. APJCP 18 (1) (2017) 3–9.
[2] J.P. Thakkar, et al., Epidemiologic and molecular prognostic review of glioblastoma, Cancer Epidemiol. Biomark. Prev. 23 (10) (2014) 1985.
[3] D.R. Johnson, H.E. Leeper, J.H. Uhm, Glioblastoma survival in the United States improved after Food and Drug Administration approval of bevacizumab: a population-based analysis, Cancer 119 (19) (2013) 3489–3495.
[4] J. Song, et al., A review of original articles published in the emerging field of radiomics, Eur. J. Radiol. 127 (2020) 108991.
[5] P. Lambin, et al., Radiomics: the bridge between medical imaging and personalized medicine, Nat. Rev. Clin. Oncol. 14 (12) (2017) 749–762.
[6] K. Li-Chun Hsieh, C.-Y. Chen, C.-M. Lo, Quantitative glioma grading using transformed gray-scale invariant textures of MRI, Comput. Biol. Med. 83 (2017) 102–108.
[7] H.-h. Cho, et al., Classification of the glioma grading using radiomics analysis, PeerJ 6 (2018), e5982.
[8] R. Kumar, et al., CGHF: a computational decision support system for glioma classification using hybrid radiomics- and stationary wavelet-based features, IEEE Access 8 (2020) 79440–79458.
[9] T. Xiao, et al., Glioma Grading Prediction by Exploring Radiomics and Deep Learning Features, in Proceedings of the Third International Symposium on Image Computing and Digital Medicine, Association for Computing Machinery, Xi'an, China, 2019, p. 208–213.
[10] C.-F. Lu, et al., Machine learning–based radiomics for molecular subtyping of gliomas, Clin. Canc. Res. 24 (18) (2018) 4429.
[11] X. Liu, et al., IDH mutation-specific radiomic signature in lower-grade gliomas, Aging 11 (2) (2019) 673–696.
[12] Z.-C. Li, et al., Multiregional radiomics profiling from multiparametric MRI: identifying an imaging predictor of IDH1 mutation status in glioblastoma 7 (12) (2018) 5999–6009.
[13] P. Lohmann, et al., Predicting IDH genotype in gliomas using FET PET radiomics, Sci. Rep. 8 (1) (2018) 13328.
[14] S. Wu, et al., Radiomics-based machine learning methods for isocitrate dehydrogenase genotype prediction of diffuse gliomas, J. Canc. Res. Clin. Oncol. 145 (3) (2019) 543–550.
[15] B. Shofty, et al., MRI radiomics analysis of molecular alterations in low-grade gliomas, Int. J. Comput. Assisted Radiol. Surg. 13 (4) (2018) 563–571.
[16] R. Fukuma, et al., Prediction of IDH and TERT promoter mutations in low-grade glioma from magnetic resonance images using a convolutional neural network, Sci. Rep. 9 (1) (2019) 20311.
[17] C. Jiang, et al., Conventional magnetic resonance imaging–based radiomic signature predicts telomerase reverse transcriptase promoter mutation status in grade II and III gliomas, Neuroradiology (2020).
[18] Y. Li, et al., MRI features predict p53 status in lower-grade gliomas via a machine-learning approach, Neuroimage: Clin. 17 (2018) 306–311.
[19] T. Sasaki, et al., Radiomics and MGMT promoter methylation for prognostication of newly diagnosed glioblastoma, Sci. Rep. 9 (1) (2019) 14435.
[20] C. Jiang, et al., Fusion radiomics features from conventional MRI predict MGMT promoter methylation status in lower grade gliomas, Eur. J. Radiol. 121 (2019) 108714.
[21] R.G.W. Verhaak, et al., Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1, Canc. Cell 17 (1) (2010) 98–110.
[22] A.K. Park, et al., Subtype-specific signaling pathways and genomic aberrations associated with prognosis of glioblastoma, Neuro Oncol. 21 (1) (2018) 59–70.
[23] S. Rathore, et al., Multivariate Analysis of Preoperative Magnetic Resonance Imaging Reveals Transcriptomic Classification of de novo Glioblastoma Patients, Front. Comput. Neurosci. 13 (2019) 81.
[24] L. Macyszyn, et al., Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques, Neuro Oncol. 18 (3) (2015) 417–425.
[25] J. Lee, et al., Spatial habitat features derived from multiparametric magnetic resonance imaging data are associated with molecular subtype and 12-month survival status in glioblastoma multiforme, PLOS ONE 10 (9) (2015), e0136557.
[26] D. Yang, et al., Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma, Med. Phys. 42 (11) (2015) 6725–6735.
[27] K.M. Naeini, et al., Identifying the mesenchymal molecular subtype of glioblastoma using quantitative volumetric analysis of anatomic magnetic resonance images, Neuro Oncol. 15 (5) (2013) 626–634.
[28] M.C. Nicholas, et al., Radiogenomics of glioblastoma: a pilot multi-institutional study to investigate a relationship between tumor shape features and tumor molecular subtype, Proc. SPIE (2016).
[29] K. Clark, et al., The cancer imaging archive (TCIA): maintaining and operating a public information repository, J. Digit. Imag. 26 (6) (2013) 1045–1057.
[30] L. Scarpace, et al., Radiology data from the cancer genome atlas glioblastoma multiforme [TCGA-GBM] collection, Canc. Imag. Arch. 11 (4) (2016) 1.
[31] Cameron W. Brennan, et al., The somatic genomic landscape of glioblastoma, Cell 155 (2) (2013) 462–477.
[32] P. Lohmann, et al., Radiomics in neuro-oncology: basics, workflow, and applications, Methods (2020).
[33] R.B. Puchalski, et al., An anatomic transcriptional atlas of human glioblastoma, Science 360 (6389) (2018) 660.
[34] B.H. Menze, et al., The multimodal brain tumor image segmentation benchmark (BRATS), IEEE Trans. Med. Imag. 34 (10) (2015) 1993–2024.
[35] S. Bakas, et al., Advancing the Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features, Sci. Data 4 (1) (2017) 170117.
[36] M. Jenkinson, et al., FSL, NeuroImage 62 (2) (2012) 782–790.
[37] S. Bakas, et al., GLISTRboost: combining multimodal MRI segmentation, registration, and biophysical tumor growth modeling with gradient boosting machines for glioma segmentation. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Springer International Publishing, Cham, 2016.
[38] A. Gooya, et al., GLISTR: glioma image segmentation and registration, IEEE Trans. Med. Imag. 31 (10) (2012) 1941–1954.
[39] D. Christos, et al., Cancer imaging phenomics toolkit: quantitative imaging analytics for precision diagnostics and predictive modeling of clinical outcome, J. Med. Imag. 5 (1) (2018) 1–21.

[40] A. Zwanenburg, et al., The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping, Radiology 295 (2) (2020) 328–338.
[41] Y.-W. Chen, C.-J. Lin, Combining SVMs with various feature selection strategies. Feature Extraction, Springer, 2006, pp. 315–324.
[42] N.Q.K. Le, et al., XGBoost Improves Classification of MGMT Promoter Methylation Status in IDH1 Wildtype Glioblastoma, J. Pers. Med. 10(3) (2020) 128.
[43] F. Pedregosa, et al., Scikit-learn: machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[44] D.T. Do, N.Q.K. Le, Using extreme gradient boosting to identify origin of replication in Saccharomyces cerevisiae via hybrid features, Genomics 112 (3) (2020) 2445–2451.
[45] M. Nakagawa, et al., Machine learning based on multi-parametric magnetic resonance imaging to differentiate glioblastoma multiforme from primary cerebral nervous system lymphoma, Eur. J. Radiol. 108 (2018) 147–154.
[46] P. Korfiatis, et al., MRI texture features as biomarkers to predict MGMT methylation status in glioblastomas, Med. Phys. 43 (6Part1) (2016) 2835–2844.
[47] Levner, I., et al. Predicting MGMT methylation status of glioblastomas from MRI texture. in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009*. 2009. Berlin, Heidelberg: Springer Berlin Heidelberg.
[48] K.L.-C. Hsieh, C.-Y. Chen, C.-M. Lo, Radiomic model for predicting mutations in the isocitrate dehydrogenase gene in glioblastomas, Oncotarget 8 (28) (2017).
[49] P.P. Batchala, et al., Neuroimaging-based classification algorithm for predicting 1p/19q-codeletion status in IDH-mutant lower grade gliomas, Am. J. Neuroradiol. 40 (3) (2019) 426–432.
[50] I. Shiri, et al., Repeatability of radiomic features in magnetic resonance imaging of glioblastoma: test–retest and image registration analyses, Med. Phys. 47 (9) (2020) 4265–4280.