Cancer Letters 451 (2019) 128–135

Contents lists available at ScienceDirect

Cancer Letters

journal homepage: www.elsevier.com/locate/canlet

Original Articles

Differentiation of glioblastoma from solitary brain metastases using radiomic machine-learning classifiers

Zenghui Qian[a,1], Yiming Li[b,1], Yongzhi Wang[c], Lianwang Li[c], Runting Li[c], Kai Wang[d], Shaowu Li[e], Ke Tang[f], Chuanbao Zhang[c], Xing Fan[b], Baoshi Chen[a,*], Wenbin Li[a,**]

[a] Department of Neuro-oncology, Neurosurgery Center, Beijing Tiantan Hospital, Capital Medical University, Beijing, China
[b] Beijing Neurosurgical Institute, Capital Medical University, Beijing, China
[c] Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing, China
[d] Department of Nuclear Medicine, Beijing Tiantan Hospital, Capital Medical University, Beijing, China
[e] Department of Neuroradiology, Beijing Neurosurgical Institute, Capital Medical University, Beijing, China
[f] Department of Neurosurgery, The 309th Hospital of Chinese People's Liberation Army, Beijing, China

ARTICLE INFO

Keywords:
Brain metastases
Glioblastoma
Radiomics
Machine learning

ABSTRACT

This study aimed to identify the optimal radiomic machine-learning classifier for differentiating glioblastoma (GBM) from solitary brain metastases (MET) preoperatively. Four hundred and twelve patients with solitary brain tumors (242 GBM and 170 solitary brain MET) were divided into training (n = 227) and test (n = 185) cohorts. Radiomic features extraction was performed with *PyRadiomics* software. In the training cohort, twelve feature selection methods and seven classification methods were evaluated to construct favorable radiomic machine-learning classifiers. The performance of the classifiers was evaluated using the mean area under the curve (AUC) and relative standard deviation in percentile (RSD). In the training cohort, thirteen classifiers had favorable predictive performances (AUC ≥ 0.95 and RSD ≤ 6). In the test cohort, receiver operating characteristic (ROC) curve analysis revealed that support vector machines (SVM) + least absolute shrinkage and selection operator (LASSO) (AUC, 0.90) classifiers had the highest prediction efficacy. Furthermore, the clinical performance of the best classifier was superior to neuroradiologists in accuracy, sensitivity, and specificity. In conclusion, employing radiomic machine-learning technology could help neuroradiologist in differentiating GBM from solitary brain MET preoperatively.

1. Introduction

Brain metastases (MET) and glioblastoma (GBM) are the most common malignant brain neoplasms in adults [1,2]. Therapeutic strategies for these two diseases are very different: the current standard of treatment for newly diagnosed GBM is maximal resection of the tumor, followed by radiotherapy and temozolomide [3], while stereotactic radiosurgery is considered an effective strategy in the treatment of MET with the advantage of excellent local control rates with minimal invasiveness [4]. Consequently, accurate preoperative differentiation of MET and GBM is critical for the individualized therapeutic decision making. Some studies have revealed that clinical history and multiplicity of lesions can facilitate differentiation between these two neoplasms [5]. Nonetheless, conventional solitary brain MET MRI features are similar to GBM, making clinical differentiation difficult.

Recently, due to the evolution of imaging hardware and technology, advanced imaging modalities play an important role in differentiating GBM from brain MET [6–8]. Advanced imaging modalities including

Abbreviations: GBM, glioblastoma; MET, metastases; AUC, area under the curve; RSD, relative standard deviation; ROC, receiver operating characteristic; SVM, support vector machines; LASSO, least absolute shrinkage and selection operator; CE, contrast-enhancement; TCGA, the Cancer Genome Atlas; T1, T1-weighted; T2, T2-weighted; ROI, region of interest; LoG, Laplacian of Gaussian; ICCs, intraclass correlation coefficients; T-test-score, TSCR; RELF, Relief; IFGN, information gain; GNRO, gain ratio; EUDT, Euclidean distance; FAOV, F-anova; WLCX, Wilcoxon rank sum; LG, logistic regression; MUIF, mutual information; SVM, support vector machine; LASSO, least absolute shrinkage and selection operator; RF, random forest; Ada, Adaboost Classifier; KNN, k-nearest neighbor; MLP, multi-layer perceptron; DT, decision tree; NB, naïve Bayes; ACC, accuracy; PPV, positive prediction value; NPV, negative predictive value

* Corresponding author. Beijing Tiantan Hospital, Department of Neurosurgery, 6 Tiantanxili, Beijing, 100050, China.
** Corresponding author. Beijing Tiantan Hospital, Capital Medical University, 6 Tiantanxili, Beijing, 100050, China.
E-mail addresses: tiantanchenbaoshi@163.com (B. Chen), liwenbin@ccmu.edu.cn (W. Li).
[1] Zenghui Qian and Yiming Li contributed equally to this work.

https://doi.org/10.1016/j.canlet.2019.02.054
Received 26 October 2018; Received in revised form 26 January 2019; Accepted 28 February 2019

perfusion MRI [9], diffusion weighted imaging [10] and magnetic resonance spectroscopy [11], provide ways to obtain information regarding the physiological and metabolic characterization of brain tumors. Although these advanced imaging modalities are attracting attention, no single finding has been substantive enough to guide clinical practice. Three major limitations remain in the previous studies. First, most studies to date have a relatively small sample size; thus, the predictive models have been typically designed and validated on the same cohort, consequently decreasing their generalizability. Second, initial attempts were limited to a handful of predefined imaging features, as opposed to extracting high-dimensional, mineable, and quantitative imaging features via high-throughput data-mining algorithms. Finally, the power of modern machine-learning technology has not yet been well utilized. Thus, there are limited data available regarding radiomic features as a means to distinguish GBM from MET.

In this study, we extracted a large panel of radiomic features from contrast-enhancement (CE) imaging data for 412 patients with solitary brain tumors. Machine-learning based feature selection and classification evaluation were performed in a training cohort and tested in an independent test cohort. Our hypothesis is that combining machine-learning technologies and high-dimensional radiomic features will facilitate the differentiation of GBM from solitary brain MET.

2. Materials and methods

2.1. Patient enrollment

This study was approved by the ethics committee of Beijing Tiantan Hospital. A total of 142 patients with histologically confirmed GBM diagnoses from the Cancer Genome Atlas (TCGA) (http://cancergenome.nih.gov) and 85 patients with solitary brain MET from the Beijing Tiantan Hospital (from September 2014 to May 2016) were collected as a training cohort, and an independent test cohort was built by the data from 185 patients (100 with GBM and 85 patients with solitary brain MET), which were retrospectively collected from Beijing Tiantan Hospital (from June 2016 to January 2018). The inclusion criteria were listed as follows: (I) pathologically confirmed GBM or solitary brain MET according to the World Health Organization (WHO) criteria [12]; (II) preoperative multi-parametric MRI scans including T1-weighted (T1), T2-weighted (T2) and CE data were available; (III) patients had no history of preoperative treatment; (IV) all the GBM cases were newly-diagnosed and without history of prior lower-grade brain cancer; and (V) clinical characteristics were available. The detailed characterization of each clinical case was listed in Supplementary Table 1.

2.2. MRI data acquisition and region-of-interest segmentation

MRI data for 142 patients with GBM from TCGA were downloaded from the Cancer Imaging Archive (http://www.cancerimagingarchive.net). MRI scans in the CGGA were generally performed using a Siemens Trio 3-T scanner (Siemens Healthcare, Erlangen, Germany), and typically included axial T1 data (repetition time, 450 ms; echo time, 15 ms; slice thickness, 5 mm), CE scans using 0.1 mmol/kg of DTPA-Gd injections (Beijing Beilu Pharmaceutical Co., Beijing China) (repetition time, 450 ms; echo time, 15 ms; slice thickness, 5 mm), and T2 images (repetition time, 5800 ms; echo time, 110 ms; slice thickness, 5 mm), with a 24 cm field of view and a matrix size of 384 × 300 pixels. In order to make the model simple and easy to use, we only segmented the tumor on CE images. The region of interest (ROI) was manually drawn according to the tumor boundary on CE images by two neuroradiologists (one with 5 years of experience, another with 10 years of experience) independently, with an open-source software MRIcro (http://www.mccauslandcenter.sc.edu/mricro/). These two neuroradiologists were blinded to the clinical data. An overview of the radiomics workflow is shown in Fig. 1.

2.3. Radiomic feature extraction

The extraction of the radiomic features was conducted by using *PyRadiomics* [13], a flexible open-source platform capable of extracting a large panel of engineered features from medical images; this radiomic quantification platform enables the standardization of both feature definitions and image processing. The parameter settings of image processing and feature extraction were provided in the supplementary material in order to facilitate the application of our findings. To avoid data heterogeneity bias, all MRI data were subjected to imaging normalization (the intensity of the image was scaled to 0–100) and resampled to the same resolution (3 mm × 3 mm × 3 mm) before feature extraction. For each ROI, six built-in filters (wavelet, Laplacian of Gaussian (LoG), square, square root, logarithm, and exponential) were applied and five feature classes (first order statistics, shape descriptors, and texture classes (gray level cooccurrence matrix, gray level run length matrix, and gray level size zone matrix)) were calculated, which resulted in a total of 1,303 radiomic features (13 shape features, 18 first-order intensity statistics features, 68 texture features, 86 square features, 86 square root features, 86 logarithm features, 86 exponential features, 172 LoG features, and 688 wavelet features). Both the two sets of ROIs generated by the two neuroradiologists would be subject to the feature extraction process, and two sets of radiomic features were obtained. Intraclass correlation coefficients (ICCs) analysis was performed on the two sets of radiomic features in order to assess the stability of each feature. Stable radiomic features were defined by ICCs > 0.8. Subsequently, one set of the stable radiomic features (obtained from the neuroradiologist with 10 years of experience) was selected and normalized to z distribution ([value – mean value]/standard deviation) for further analysis. Unsupervised K-means clustering analysis was performed on the stable radiomic features to determine if GBM and MET are differentiable by imaging subtype.

2.4. Feature selection methods

In this study, twelve feature selection methods were used in the analysis, which were divided into three categories [15,16]: filter methods (T-test-score (TSCR), Relief (RELF), information gain (IFGN), gain ratio (GNRO), Euclidean distance (EUDT), F-anova (FAOV), Wilcoxon rank sum (WLCX), logistic regression (LG), and mutual information (MUIF)), wrapper methods (support vector machine (SVM)), and embedded methods (10-fold cross-validation least absolute shrinkage and selection operator (LASSO) and random forest (RF)). Feature selection methods were conducted using R software (version 3.4.0, R Foundation for Statistical Computing, Vienna, Austria) and Python (version 3.6.4). Seven of the filter methods were univariate methods (TSCR, RELF, IFGN, GNRO, EUDT, FAOV, and WLCX) and two were multivariate methods (LG and MUIF). For each of the seven univariate methods, we listed the first 100 high-scoring features based on a relevance score. Then, we combined logistic regression model and ROC curve analysis to identify features that could achieve the maximal area under the curve (AUC) by continuing to add features in the list from top to bottom.

2.5. Feature classification methods

We applied seven supervised machine-learning algorithms; these classification methods were Adaboost Classifier (Ada), k-nearest neighbor (KNN), multi-layer perceptron (MLP), decision tree (DT), naïve Bayes (NB), RF, and SVM. Feature classification methods were all implemented using the Python (version 3.6.4) machine-learning library known as scikit-learn (version 19.0) [14].

2.6. Predictive performance of machine-learning classifiers

The 12 subsets of selected features were evaluated by using each of

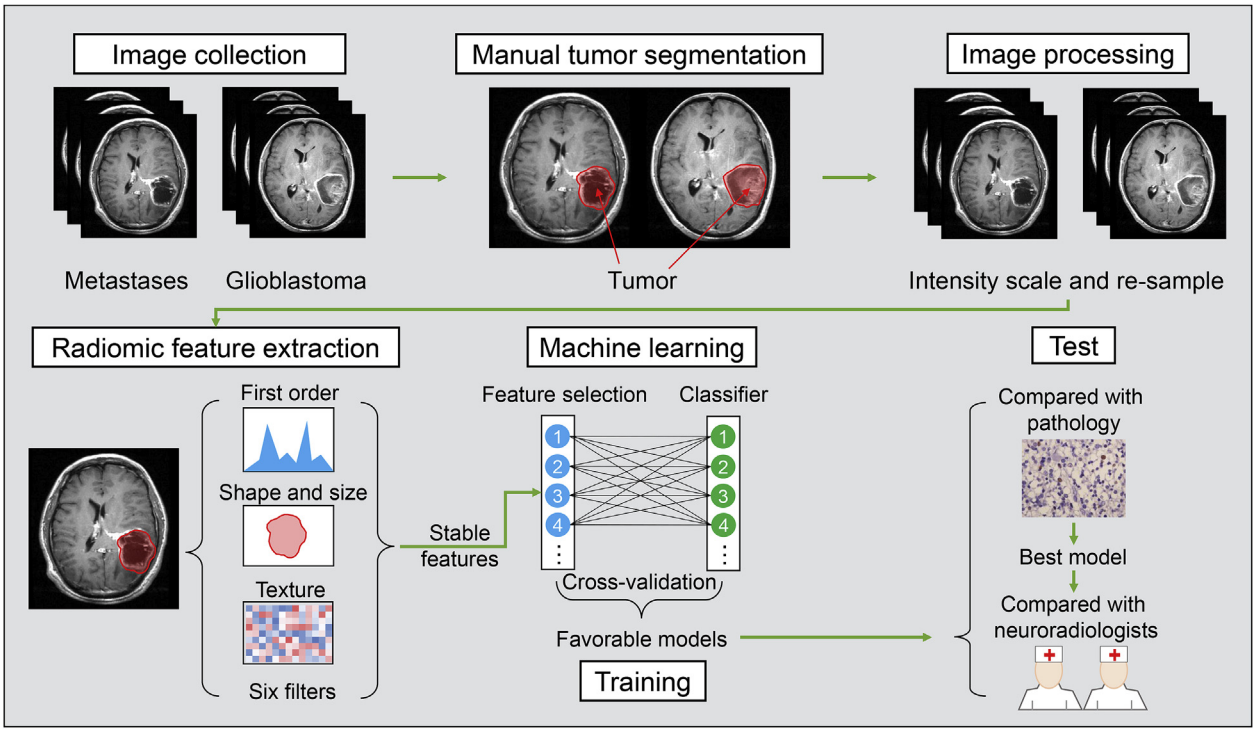

Fig. 1. The workflow of radiomic analysis in the current study. After feature extraction, stable features were selected for further analysis. Multiple feature selection scheme and classifiers were combined and favorable models were selected with the help of cross-validation in the training cohort. In an independent test cohort, the best model was identified by compared with pathology, and then the performance of the best model was compared with two neuroradiologists.

the 7 classification methods, which generated 84 (12*7 = 84) machine-learning classifiers. We adopted a nomenclature for these classifiers that combined the names of the feature selection method and the classification method. For example, "EUDT_LASSO" was a model based on LASSO classification methods with radiomic features selected by the EUDT feature selection method.

We applied 5-fold cross-validation as the criteria for each of the 84 machine-learning classifiers on the training cohort using the StratifiedKFold iterator in scikit-learn. The area under the curve (AUC) and relative standard deviation in percentile (RSD) were used to evaluate the predictive performance and stability of the classifiers, respectively. RSD was defined as the ratio between the standard deviation and mean of the 5-fold cross-validated AUC values: RSD = (sdAUC/meanAUC) * 100. Then, the trained models were evaluated in an independent test cohort using the AUC. Classifiers with cross-validated values for AUC ≥ 0.95 and RSD ≤ 6 in the training cohort, and the model with the highest AUC value in the test cohort was considered as the final model. Confusion matrix derived metrics including accuracy (ACC), sensitivity, specificity, positive prediction value (PPV) and negative predictive value (NPV) were calculated to further define the final model.

Subsequently, we employed the abovementioned two neuroradiologists, who were blinded to the clinical data, to manually differentiate GBM from MET according to T1, T2 and CE images. The performance of the two neuroradiologists was evaluated with ROC curve analysis, and was compared with the performance of the final machine-leaning model. Chi-square test was performed the compare the proportion of predicted GBM/MET between the neuroradiologist and the final machine-leaning model.

2.7. Statistical analysis

Differences in clinical and MRI characteristics between GBM and MET were assessed by the Student's *t*-test or chi-square test, as appropriate. All statistical tests were two-tailed and conducted with a statistical significance level set at 0.05. The statistical analysis and figure plots were performed using R software (version 3.0.1; http://www.R-project.org) and SPSS software (SPSS Inc., Chicago, IL).

3. Results

3.1. Clinical and magnetic resonance imaging characteristics

The clinical and MRI characteristics of the patients in the training and test cohorts are shown in Table 1. There were no significant differences in age or sex between patients with GBM and MET in the two cohorts. The anatomical characteristics revealed via MRI indicated that more patients with infratentorial structures were diagnosed with MET ($p < 0.001$). However, no differences were found between GBM and MET in the characteristics of necrosis appearance or edema appearance for both cohorts.

3.2. Qualitative radiomic feature analysis

Fig. 2 depicts the robustness of radiomic features using ICC analysis. Shape features (ICC = 0.983 ± 0.015), first-order intensity statistics features (ICC = 0.985 ± 0.019), texture features (ICC = 0.964 ± 0.046), square features (ICC = 0.979 ± 0.041), square root features (ICC = 0.951 ± 0.058), logarithm features (ICC = 0.938 ± 0.131), exponential features (ICC = 0.985 ± 0.042), and LoG features (ICC = 0.989 ± 0.015) had high ICCs, and were thus more robust, whereas wavelet features had low ICCs (ICC = 0.670 ± 0.256) and were less robust. In total, 898 of the 1,303 (68.9%) extracted radiomic features showed high stability (ICC > 0.8) including 13 shape features, 18 first-order intensity statistics features, 68 texture features, 82 square features, 82 square root features, 82 logarithm features, 85 exponential features, 172 LoG features, and 296 wavelet features. As depicted in Supplementary Fig. 1, we observed two imaging subtypes by conducting unsupervised clustering on these stable features. However, the association between the imaging subtypes and histology subtypes was

Table 1
Clinical and MRI characteristics of patients with MET and GBM.

	Training cohort			Test cohort		
	MET (n = 85)	GBM (n = 142)	*P* value	MET (n = 85)	GBM (n = 100)	*P* value
Age (years)	55.9	57.7	0.278[b]	56.3	53.4	0.107[b]
Sex						
Female	36	59	0.940[a]	33	42	0.661[a]
Male	49	82		52	58	
NA		1				
Localization						
Supratentorial	69	142	< **0.001**[a]	65	97	< **0.001**[a]
Infratentorial	16	0		20	3	
Appearance of necrosis						
Yes	73	133	0.050[a]	78	95	0.242[a]
No	12	9		7	5	
Appearance of edema						
Yes	74	130	0.278[a]	74	83	0.443[a]
No	11	12		11	17	

Abbreviations: MET, metastases; GBM, glioblastoma; NA, not available.
[a] Chi-square test.
[b] Student's t-test.

not obvious.

3.3. Feature selection

In the training cohort, 12 subsets of selected features were screened based on the 12 feature selection methods and were described as follows (the number of selected features was displayed following each method's name): EUDT_38, FAOV_41, GNRO_17, IFGN_31, LASSO_24, LG_4, MUIF_35, RELF_17, RF_43, SVM_3, TSCR_50, and WLCX_46. Then, based on these 12 subsets of selected features and 7 feature classification methods, we examined 84 machine-learning classifiers. The heatmap of the mean AUC and RSD values of the 84 machine-learning classifiers are shown in Fig. 3. The mean AUC value was 0.90 (range, 0.74 to 0.96) and the mean RSD value was 7.99 (range, 4.19 to 12.66).

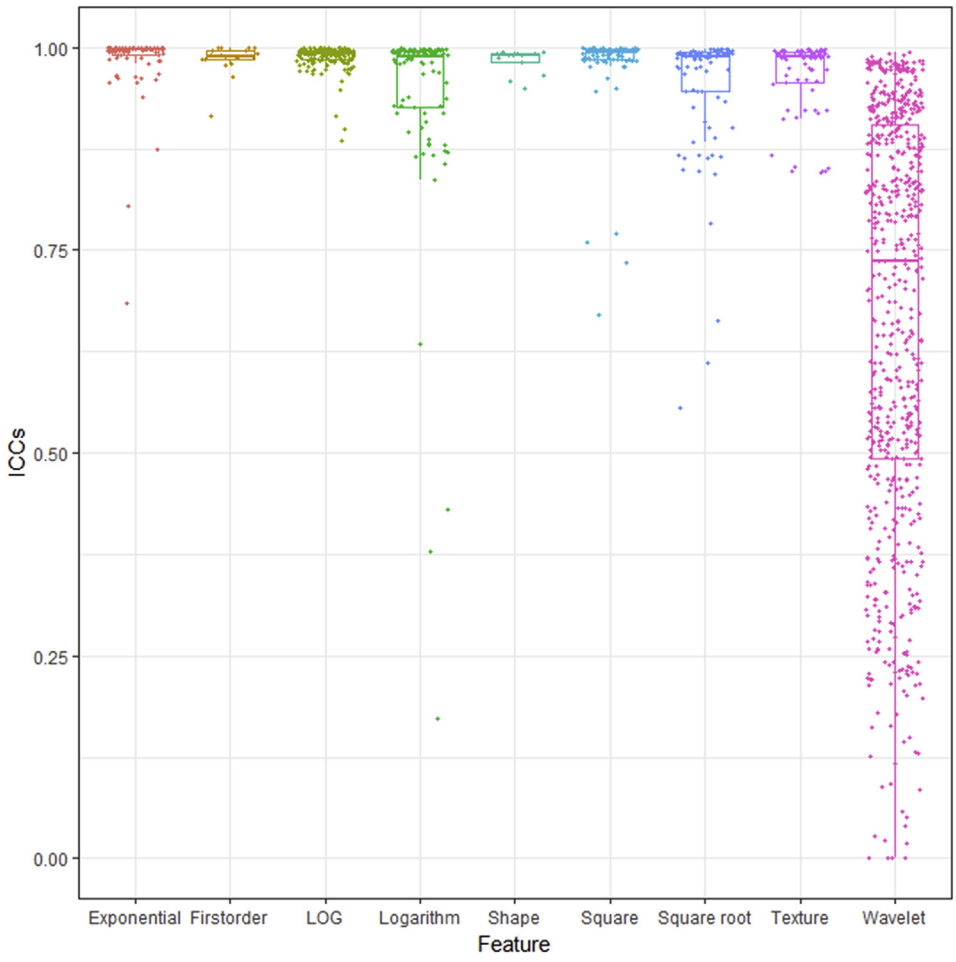

Fig. 2. Boxplot of intraclass correlation coefficients (ICCs) of features extracted from nine feature classes.

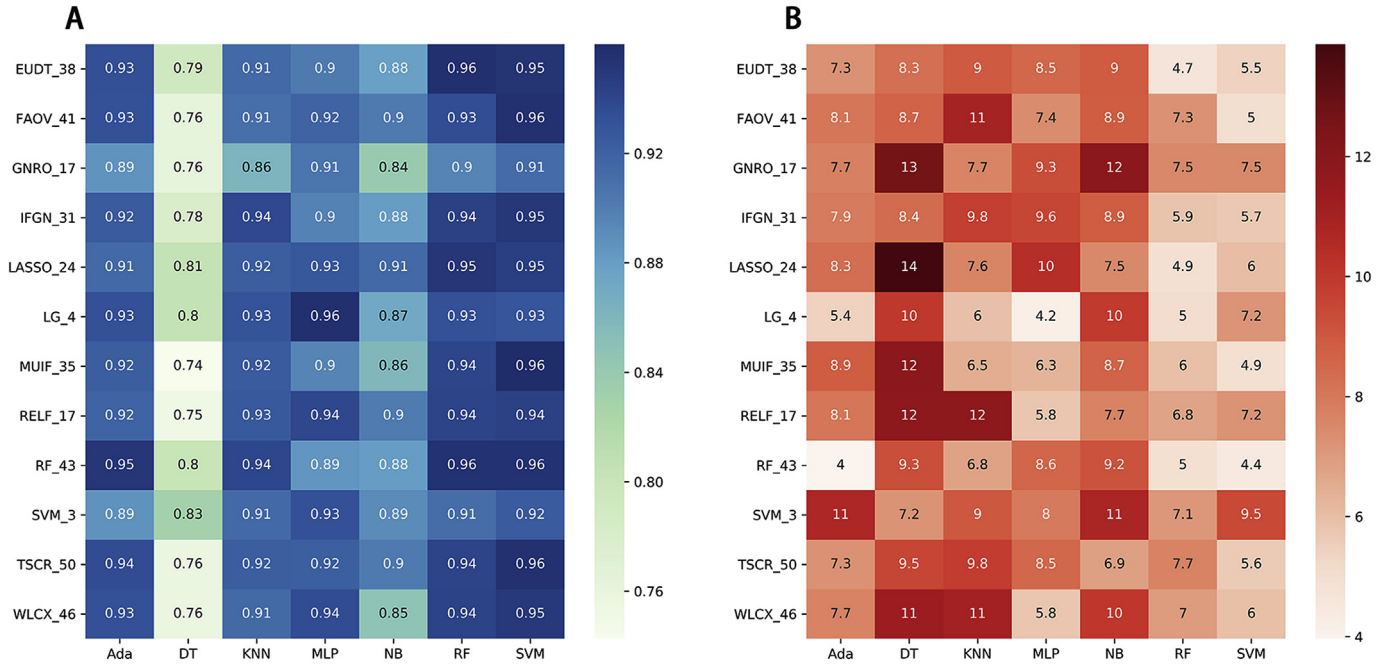

Fig. 3. Heatmaps illustrating the predictive performance (AUC) of different combinations of feature selection methods (rows) and classification algorithms (columns) in the training cohort. (**A**) Cross-validated AUC values of 84 classifiers. (**B**) RSD values of 84 classifiers. AUC, area under the curve; RSD, relative standard deviation in percentile.

SVM classification outperformed other classifications; the mean AUC value of the 12 classifiers using the SVM classification reached 0.945 in the training cohort (Fig. 3A). Regarding prediction stability, the 12 SVM classifiers had a mean RSD value of 6.21, which was the lowest compared to other classifiers (Fig. 3B).

According to the criteria for the cross-validated values (AUC ≥ 0.95 and RSD ≤ 6), thirteen classifiers (RF_Ada, LG_MLP, EUDT_RF, LASSO_RF, RF_RF, EUDT_SVM, FAOV_SVM, IFGN_SVM, LASSO_SVM, MUIF_SVM, RF_SVM, TSCR_SVM, and WLCX_SVM) were selected. The established classifiers were then applied to the test cohort. Among the selected 13 classifiers, the classifier of LASSO_SVM has the highest AUC value (0.90, Fig. 4A), which was chosen as the optimal classifier. The 24 features selected by LASSO and their corresponding descriptions were listed in Supplementary Table 2, including 3 shape features, 4 texture features, 1 square feature, 1 square root feature, 1 logarithm feature, 2 exponential features, 4 LoG features, and 8 wavelet features (Supplementary Table 2). Analysis of the confusion matrix related classification metrics of LASSO_SVM revealed that the accuracy was 82.7%, and the sensitivity, specificity, PPV, and NPV were 79.8%, 87.3%, 90.0%, and 72.9%, respectively (Fig. 4B). The illustration of the 5-fold cross-validated ROC curve of the LASSO_SVM classifier in the training cohort and ROC curve of the LASSO_SVM classifier in the test

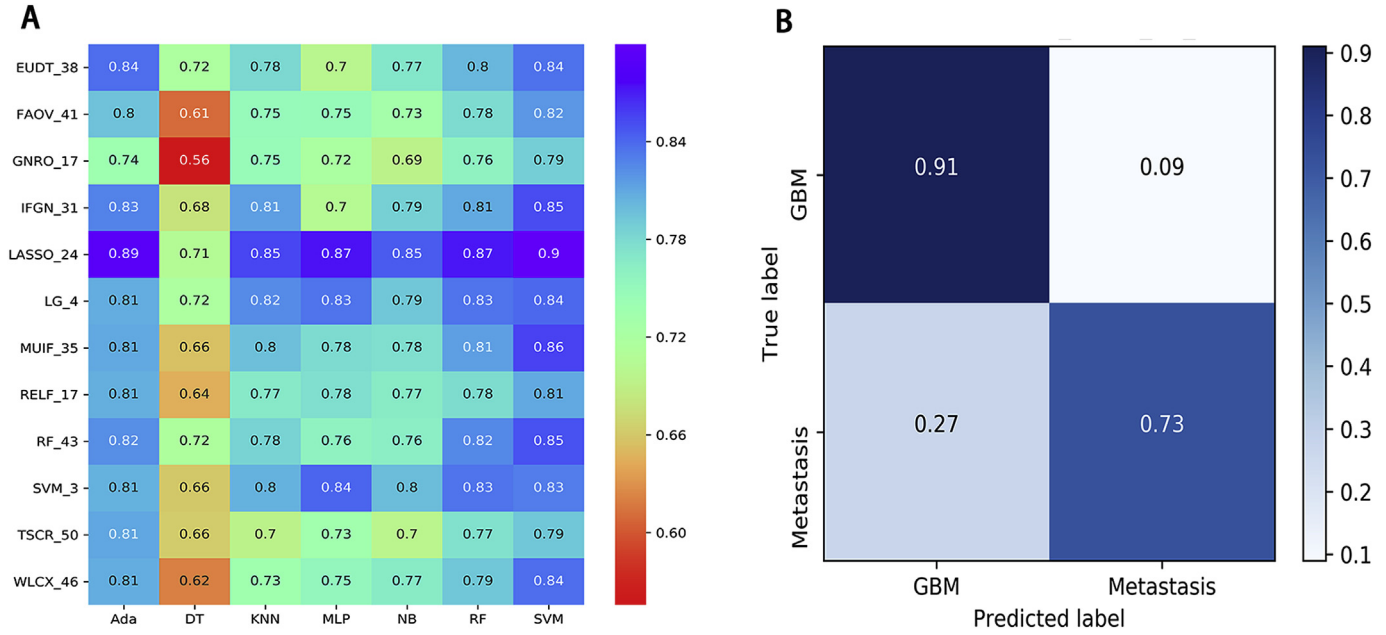

Fig. 4. Predictive performance of the constructed classifier in the test cohort. (**A**) AUC values of 84 classifiers. (**B**) Confusion matrix with the LASSO_SVM classifier. AUC, area under the curve; LASSO, least absolute shrinkage and selection operator; SVM, support vector machines.

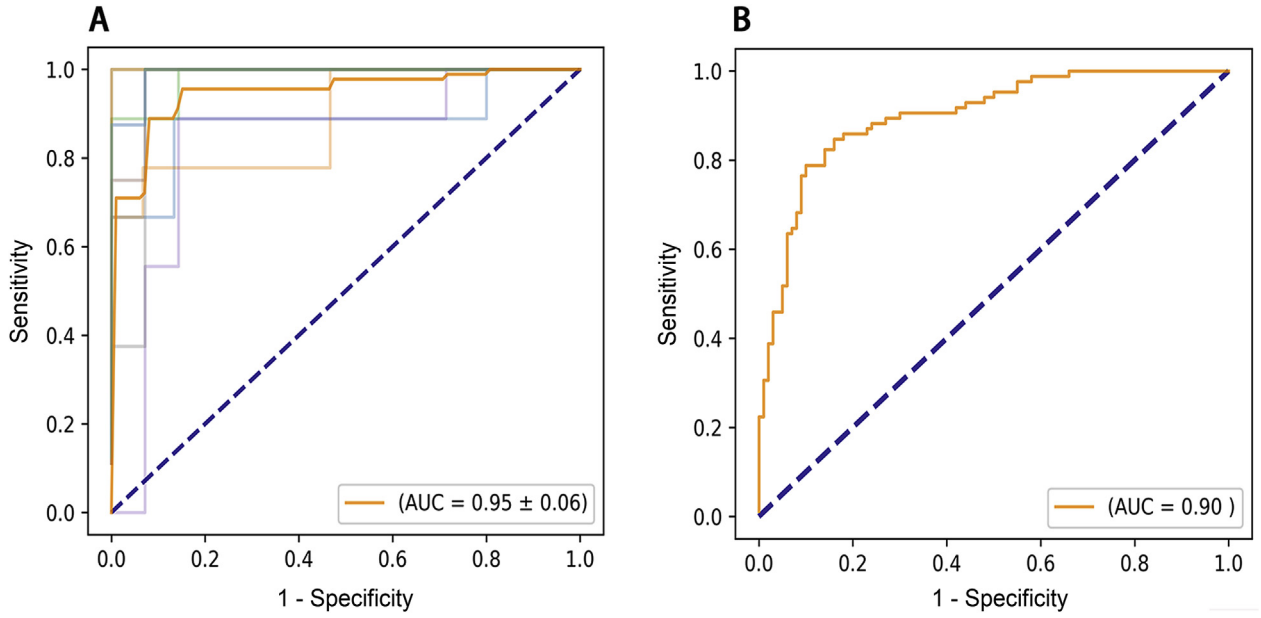

Fig. 5. Receiver operating characteristic (ROC) curve of the optimal classifier. (**A**) The 5-fold cross-validated ROC curve of the optimal LASSO_SVM classifier in the training cohort. (**B**) The ROC curve of the LASSO_SVM classifier in the test cohort. LASSO, least absolute shrinkage and selection operator; SVM, support vector machines.

Table 2
Comparison of predictive performance between radiomic machine-learning classifier and neuroradiologists in the test cohort.

Neuroradiologists	Sen, *P*	Spe, *P*	NPV	PPV	Accuracy, *P*
5 year	0.66, 0.022[a]	0.60, < 0.001[a]	0.67	0.59	0.63, < 0.001[a]
10 year	0.66, 0.025[a]	0.65, 0.003[a]	0.75	0.54	0.65, < 0.001[a]
LASSO_SVM	0.80, ——	0.87, ——	0.90	0.73	0.83, ——

Abbreviations: Sen, sensitivity; Spe, specificity; PPV, positive prediction value; NPV, negative prediction value; LASSO, least absolute shrinkage and selection operator; SVM, support vector machine.
[a] Chi-square test.

cohort were shown in Fig. 5, panels A and B. Also, the clinical value of the selected radiomic machine-learning classifier was further evaluated by comparing its performance with the assessment of the test cohort data by two neuroradiologists. As shown in Table 2, the classifier of LASSO_SVM was superior to neuroradiologists in accuracy, sensitivity, and specificity.

4. Discussion

In this study, by integrating a large panel of radiomic features and machine-learning techniques, we established an accurate and reproducible classifier to distinguish patients with MET from patients with GBM. Importantly, this efficient classifier was obtained using cross-validation and was subsequently tested on an entirely new test cohort, bolstering its generalizability. In addition, we found that this radiomic machine-learning classifier was superior to neuroradiologists in diagnostic accuracy, sensitivity, and specificity.

Traditionally, several conventional observable MRI features such as the maximum diameter of peritumoral edema compared to the diameter of the contrast enhancing tumor, peritumoral T2 prolongation, and tumor morphology have been used to differentiate between GBM and MET [17,18]. However, these morphological parameters are non-specific and subject to individual interpretation, resulting in a high level of inter-observer variability. The inclusion of advanced imaging features, which could characterize tumor cellularity, microvascular permeability, and vascular proliferation, can improve diagnostic accuracy compared to the use of conventional MRI features. For instance, a previous study indicated that diffusion tensor imaging is a very promising imaging tool for differentiating GBM from MET [6]. Similarly, some recent studies also showed that measuring the relative cerebral blood volume in a tumor by using the dynamic susceptibility contrast technique could discriminate GBM from MET, significantly enhancing diagnostic abilities [19,20]. However, other studies showed inconsistent results [21–23]. The discrepancies might be explained by different of parameter analyses, section selection, or ROI positioning [18]. Moreover, these advanced imaging modalities are not readily available tools in most radiology departments and require more examination time, which limit their use in clinical practice.

Recently, radiomic analysis has been considered as a useful tool to extract a large number of quantitative features from medical images [24,25]. It can provide more metabolic and biologic information in addition to conventional MRI metrics [26]. In recent years, this technique has emerged as a powerful way to facilitate treatment personalization in clinical practice, particularly in relation to tumor detection, subtype classification, and prognostic estimation. Considerable progress has been made in a variety of pathologies such as Alzheimer disease [27], multiple sclerosis [28], lung cancer [29], breast cancer [30], and also gliomas [31]. To date, there have been a few reports on the differentiation of GBM and MET with radiomics analysis. For instance, a previous radiomics analysis study indicated that volumetric textural features from CE MRI data may satisfactorily discriminate between these two types of tumors [32]. However, this study was limited by small number of available radiomic features and small number of patients, which decreased the accuracy and creditability of the results.

To expand beyond the previous studies, we extracted a total of 1,303 radiomic features from preoperative MRI data by using *PyRadiomics*. To further reduce the variations in the manual segmentations between radiologists, the ICCs were calculated for each feature and only stable features were selected for further analysis [33]. We believe that these methods of image postprocessing and feature extraction in our study may supersede some of the inherent challenges of radiomics analysis.

Like any high-throughput data-mining field, the curse of dimensionality presents a challenge for radiomics analysis [34]. As depicted in Supplementary Fig. 1, the imaging subtypes derived from

unsupervised clustering could not capture the unique characteristics of different histology subtypes precisely, while feature selection methods combined with feature classification methods have mitigated the challenge of dimensionality in large datasets [16,35]. In our study, we evaluated twelve feature selection methods and seven classification methods to identify optimal radiomic machine-learning classifiers for differentiating GBM and MET. Those feature selection and classification methods were chosen in our analysis because of their simplicity, efficiency, and popularity in the field of radiomics analysis. As a result, we identified a machine-learning classifier by combining the feature selection method LASSO with the feature classification method SVM; this classifier has the highest prognostic performance with high stability in this study.

LASSO is a compression estimation method in variable selection that has been widely used for biomarker selections in high-dimensional data [36]. It constructs a more refined model by creating a penalty function, causing it to compress some coefficients while setting other coefficients to zero. In this way, dimension reduction (feature screening) is achieved and over-fitting is prevented in the process of training the model. SVM is a robust, powerful, and effective machine-learning classifier that has been predominately used in the fields of neuroimaging and molecular biology [37,38]. With the combination of these properties, the LASSO regression model and the SVM classifier collaborated excellently in the radiomics analysis. Additionally, the identified radiomics features selected by LASSO algorithm were derived from different filters and feature classes, suggesting that different categories of features could provide complementary information in distinguishing the GBM and MET. Although the biological behavior behind these radiomics features is still unclear, we suppose that these features could capture the fine radiomics characteristics of microstructure and the local environment of the tumor.

Our study has a few limitations. First, the imaging data used were acquired from multiple MRI systems with varying protocols, which may explain the model performance discrepancy in the two independent cohorts. Future studies using better quality images and consistent protocols will further improve the predictive efficiency of radiomics. Second, only CE sequences were used in our radiomics analysis since the boundaries of MET and GBM are better defined in CE sequences compared to T2 sequences. Our model can be improved by integrating multi-model imaging data (such as diffusion tensor and perfusion imaging) in the future. Finally, although our study has a large sample size containing independent training and test cohorts, prospective studies are warranted.

In conclusion, the presented radiomic machine-learning classifier is a non-invasive tool for differentiating GBM from MET preoperatively with favorable predictive accuracy and stability. We believe that combining with machine-learning methods, radiomics analysis will improve oncology precision and clinical practice.

Conflicts of interests

The authors declare no conflict of interest.

Acknowledgments

This work was supported by funds from National Natural Science Foundation of China (No. 81601452), Key laboratory of functional and clinical translational medicine, Fujian province university (JNYLC1808), National Natural Science Foundation of China (No. 81773208), Beijing Nova Program (Z16110004916082).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.canlet.2019.02.054.

References

[1] H. Ohgaki, P. Kleihues, Epidemiology and etiology of gliomas, Acta Neuropathol. 109 (2005) 93–108.
[2] C.S. Platta, D. Khuntia, M.P. Mehta, J.H. Suh, Current treatment strategies for brain metastasis and complications from therapeutic techniques: a review of current literature, Am. J. Clin. Oncol. 33 (2010) 398–407.
[3] M. Weller, M. van den Bent, K. Hopkins, J.C. Tonn, R. Stupp, A. Falini, E. Cohen-Jonathan-Moyal, D. Frappaz, R. Henriksson, C. Balana, O. Chinot, Z. Ram, G. Reifenberger, R. Soffietti, W. Wick G, European Association for Neuro-Oncology Task Force on Malignant, EANO guideline for the diagnosis and treatment of anaplastic gliomas and glioblastoma, Lancet Oncol. 15 (2014) e395–403.
[4] C. Gaudy-Marqueste, R. Carron, C. Delsanti, A. Loundou, S. Monestier, E. Archier, M.A. Richard, J. Regis, J.J. Grob, On demand Gamma-Knife strategy can be safely combined with BRAF inhibitors for the treatment of melanoma brain metastases, Ann. Oncol. 25 (2014) 2086–2091.
[5] R.N. Al-Okaili, J. Krejza, J.H. Woo, R.L. Wolf, D.M. O'Rourke, K.D. Judy, H. Poptani, E.R. Melhem, Intraaxial brain masses: MR imaging-based diagnostic strategy–initial experience, Radiology 243 (2007) 539–550.
[6] S. Wang, S.J. Kim, H. Poptani, J.H. Woo, S. Mohan, R. Jin, M.R. Voluck, D.M. O'Rourke, R.L. Wolf, E.R. Melhem, S. Kim, Diagnostic utility of diffusion tensor imaging in differentiating glioblastomas from brain metastases, AJNR. American journal of neuroradiology 35 (2014) 928–934.
[7] N. Mouthuy, G. Cosnard, J. Abarca-Quinones, N. Michoux, Multiparametric magnetic resonance imaging to differentiate high-grade gliomas and brain metastases, Journal of neuroradiology. Journal de neuroradiologie 39 (2012) 301–307.
[8] D.O. Kamson, S. Mittal, A. Buth, O. Muzik, W.J. Kupsky, N.L. Robinette, G.R. Barger, C. Juhasz, Differentiation of glioblastomas from metastatic brain tumors by tryptophan uptake and kinetic analysis: a positron emission tomographic study with magnetic resonance imaging comparison, Mol. Imag. 12 (2013) 327–337.
[9] O. Halshtok Neiman, S. Sadetzki, A. Chetrit, S. Raskin, G. Yaniv, C. Hoffmann, Perfusion-weighted imaging of peritumoral edema can aid in the differential diagnosis of glioblastoma mulltiforme versus brain metastasis, Isr. Med. Assoc. J.: Isr. Med. Assoc. J. 15 (2013) 103–105.
[10] G. Yang, T.L. Jones, T.R. Barrick, F.A. Howe, Discrimination between glioblastoma multiforme and solitary metastasis using morphological features derived from the p:q tensor decomposition of diffusion tensor imaging, NMR Biomed. 27 (2014) 1103–1111.
[11] I. Tsougos, P. Svolos, E. Kousi, K. Fountas, K. Theodorou, I. Fezoulidis, E. Kapsalaki, Differentiation of glioblastoma multiforme from metastatic brain tumor using proton magnetic resonance spectroscopy, diffusion and perfusion metrics at 3 T, Cancer Imaging: Off. Publ. Intern. Cancer Imaging Soc. 12 (2012) 423–436.
[12] D.N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W.K. Cavenee, H. Ohgaki, O.D. Wiestler, P. Kleihues, D.W. Ellison, The 2016 World Health organization classification of tumors of the central nervous system: a summary, Acta Neuropathol. 131 (2016) 803–820.
[13] J.J.M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R.G.H. Beets-Tan, J.C. Fillion-Robin, S. Pieper, H. Aerts, Computational radiomics system to decode the radiographic phenotype, Cancer Res. 77 (2017) e104–e107.
[14] A. Abraham, F. Pedregosa, M. Eickenberg, P. Gervais, A. Mueller, J. Kossaifi, A. Gramfort, B. Thirion, G. Varoquaux, Machine learning for neuroimaging with scikit-learn, Front. Neuroinf. 8 (2014) 14.
[15] Z.M. Hira, D.F. Gillies, A review of feature selection and feature extraction methods applied on microarray data, Adv. Bioinform. 2015 (2015) 198363.
[16] C. Parmar, P. Grossmann, J. Bussink, P. Lambin, H.J. Aerts, Machine learning methods for quantitative radiomic biomarkers, Sci. Rep. 5 (2015) 13087.
[17] M.H. Maurer, M. Synowitz, H. Badakshi, L.N. Lohkamp, J. Wustefeld, M.L. Schafer, E. Wiener, Glioblastoma multiforme versus solitary supratentorial brain metastasis: differentiation based on morphology and magnetic resonance signal characteristics, Röfo: Fortschritte dem Geb. Rontgenstrahlen Nukl. 185 (2013) 235–240.
[18] X.Z. Chen, X.M. Yin, L. Ai, Q. Chen, S.W. Li, J.P. Dai, Differentiation between brain glioblastoma multiforme and solitary metastasis: qualitative and quantitative analysis based on routine MR imaging, AJNR. American journal of neuroradiology 33 (2012) 1907–1912.
[19] S. Cha, J.M. Lupo, M.H. Chen, K.R. Lamborn, M.W. McDermott, M.S. Berger, S.J. Nelson, W.P. Dillon, Differentiation of glioblastoma multiforme and single brain metastasis by peak height and percentage of signal intensity recovery derived from dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging, AJNR. American journal of neuroradiology 28 (2007) 1078–1084.
[20] J.H. Ma, H.S. Kim, N.J. Rim, S.H. Kim, K.G. Cho, Differentiation among glioblastoma multiforme, solitary metastatic tumor, and lymphoma using whole-tumor histogram analysis of the normalized cerebral blood volume in enhancing and perienhancing lesions, AJNR. American journal of neuroradiology 31 (2010) 1699–1706.
[21] K. Tsuchiya, A. Fujikawa, M. Nakajima, K. Honya, Differentiation between solitary brain metastasis and high-grade glioma by diffusion tensor imaging, Br. J. Radiol. 78 (2005) 533–537.
[22] E.J. Lee, K. terBrugge, D. Mikulis, D.S. Choi, J.M. Bae, S.K. Lee, S.Y. Moon, Diagnostic value of peritumoral minimum apparent diffusion coefficient for differentiation of glioblastoma multiforme from solitary metastatic lesions, AJR. American journal of roentgenology 196 (2011) 71–76.
[23] A. Server, T.E. Orheim, B.A. Graff, R. Josefsen, T. Kumar, P.H. Nakstad, Diagnostic examination performance by using microvascular leakage, cerebral blood volume, and blood flow derived from 3-T dynamic susceptibility-weighted contrast-

enhanced perfusion MR imaging in the differentiation of glioblastoma multiforme and brain metastasis, Neuroradiology 53 (2011) 319–330.
[24] M.D. Kuo, N. Jamshidi, Behind the numbers: decoding molecular phenotypes with radiogenomics–guiding principles and technical considerations, Radiology 270 (2014) 320–325.
[25] H.J. Aerts, The potential of radiomic-based phenotyping in precision medicine: a review, JAMA oncology 2 (2016) 1636–1642.
[26] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures, They Are Data, Radiology 278 (2016) 563–577.
[27] J. Zhang, C. Yu, G. Jiang, W. Liu, L. Tong, 3D texture analysis on MRI images of Alzheimer's disease, Brain imaging and behavior 6 (2012) 61–69.
[28] Y. Zhang, H. Zhu, J.R. Mitchell, F. Costello, L.M. Metz, T2 MRI texture analysis is a sensitive measure of tissue injury and recovery resulting from acute inflammatory lesions in multiple sclerosis, Neuroimage 47 (2009) 107–111.
[29] S.H. Yoon, C.M. Park, S.J. Park, J.H. Yoon, S. Hahn, J.M. Goo, Tumor heterogeneity in lung cancer: assessment with dynamic contrast-enhanced MR imaging, Radiology 280 (2016) 940–948.
[30] J.H. Kim, E.S. Ko, Y. Lim, K.S. Lee, B.K. Han, E.Y. Ko, S.Y. Hahn, S.J. Nam, Breast cancer heterogeneity: MR imaging texture analysis and survival outcomes, Radiology (2016) 160261.
[31] O. Gevaert, L.A. Mitchell, A.S. Achrol, J. Xu, S. Echegaray, G.K. Steinberg, S.H. Cheshier, S. Napel, G. Zaharchuk, S.K. Plevritis, Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features, Radiology 273 (2014) 168–174.
[32] P. Georgiadis, D. Cavouras, I. Kalatzis, D. Glotsos, E. Athanasiadis, S. Kostopoulos, K. Sifaki, M. Malamas, G. Nikiforidis, E. Solomou, Enhancing the discrimination accuracy between metastases, gliomas and meningiomas on brain MRI by volumetric textural features and ensemble pattern recognition methods, Magn. Reson. Imag. 27 (2009) 120–130.
[33] H.J. Aerts, E.R. Velazquez, R.T. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers, M.M. Rietbergen, C.R. Leemans, A. Dekker, J. Quackenbush, R.J. Gillies, P. Lambin, Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, Nat. Commun. 5 (2014) 4006.
[34] C.N. Andreassen, L.M. Schack, L.V. Laursen, J. Alsner, Radiogenomics - current status, challenges and future directions, Cancer Lett. 382 (2016) 127–136.
[35] B. Zhang, X. He, F. Ouyang, D. Gu, Y. Dong, L. Zhang, X. Mo, W. Huang, J. Tian, S. Zhang, Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma, Cancer Lett. 403 (2017) 21–27.
[36] J. Gui, H. Li, Penalized Cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data, Bioinformatics 21 (2005) 3001–3008.
[37] H. Han, X. Jiang, Overcome support vector machine diagnosis overfitting, Canc. Inf. 13 (2014) 145–158.
[38] Y. Zhu, Y. Tan, Y. Hua, M. Wang, G. Zhang, J. Zhang, Feature selection and performance evaluation of support vector machine (SVM)-based classifier for differentiating benign and malignant pulmonary nodules by computed tomography, J. Digit. Imaging 23 (2010) 51–65.