European Journal of Radiology 116 (2019) 128–134

Contents lists available at ScienceDirect

European Journal of Radiology

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

Research article

A deep learning radiomics model for preoperative grading in meningioma

Yongbei Zhu[a,b,c,1], Chuntao Man[a,1], Lixin Gong[d,1], Di Dong[b,e,1], Xinyi Yu[f], Shuo Wang[b,e], Mengjie Fang[b,e], Siwen Wang[b,e], Xiangming Fang[f,*], Xuzhu Chen[g,***], Jie Tian[b,c,d,h,**]

[a] School of Automation, Harbin University of Science and Technology, Heilongjiang, Harbin, 150080, China
[b] CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China
[c] Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University, Beijing, 100191, China
[d] Sino-Dutch Biomedical and Information Engineering School, Northeastern University, Shenyang, 110169, China
[e] University of Chinese Academy of Sciences, Beijing, 100080, China
[f] Imaging Center, Wuxi People's Hospital, Nanjing Medical University, Wuxi, 214000, China
[g] Department of Radiology, Beijing Tiantan Hospital, Capital Medical University, Beijing, 100050, China
[h] Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Science and Technology, Xidian University, Xi'an, 710126, China

ARTICLE INFO

Keywords:
Radiomics
Deep learning
Meningioma
Tumor grading
Magnetic resonance imaging

ABSTRACT

Objectives: To noninvasively differentiate meningioma grades by deep learning radiomics (DLR) model based on routine post-contrast MRI.

Methods: We enrolled 181 patients with histopathologic diagnosis of meningioma who received post-contrast MRI preoperative examinations from 2 hospitals (99 in the primary cohort and 82 in the validation cohort). All the tumors were segmented based on post-contrast axial T1 weighted images (T1WI), from which 2048 deep learning features were extracted by the convolutional neural network. The random forest algorithm was used to select features with importance values over 0.001, upon which a deep learning signature was built by a linear discriminant analysis classifier. The performance of our DLR model was assessed by discrimination and calibration in the independent validation cohort. For comparison, a radiomic model based on hand-crafted features and a fusion model were built.

Results: The DLR signature comprised 39 deep learning features and showed good discrimination performance in both the primary and validation cohorts. The area under curve (AUC), sensitivity, and specificity for predicting meningioma grades were 0.811(95% CI, 0.635–0.986), 0.769, and 0.898 respectively in the validation cohort. DLR performance was superior over the hand-crafted features. Calibration curves of DLR model showed good agreements between the prediction probability and the observed outcome of high-grade meningioma.

Conclusions: Using routine MRI data, we developed a DLR model with good performance for noninvasively individualized prediction of meningioma grades, which achieved a quantization capability superior over the hand-crafted features. This model has potential to guide and facilitate the clinical decision-making of whether to observe or to treat patients by providing prognostic information.

1. Introduction

Meningiomas are the most common extra-axial neoplasms in the supratentorial compartment. Extra-axial tumors are the tumors of extracerebral location and they are usually benign [1]. Pathologically, it is classified into WHO grade I, II, and III according to the updated 2016 WHO classification system [2]. Grade I (low-grade) meningiomas account for approximately 90% cases, while Grade II and III (high-grade)

Abbreviations: AUC, area under curve; CI, confidence interval; DLR, deep learning radiomics; GLCM, gray level co-occurrence matrix; GLRLM, gray level run length matrix; GLSZM, gray level size zone matrix; HCR, hand-crafted radiomics; LDA, linear discriminant analysis; MRI, Magnetic resonance imaging; ROI, Region of interest; SBS, sequential backward selection; SVM, support vector machine; T1WI, T1 weighted images

* Corresponding author at: Imaging Center, Wuxi People's Hospital, Nanjing Medical University, No. 299, Qingyang road, Liangxi District, Wuxi, Jiangsu, 214023, China.
** Corresponding author at: Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China.
*** Corresponding author at: Department of Radiology, Beijing Tiantan Hospital, Capital Medical University, No. 119 Nansihuan Xilu, Fengtai District, Beijing, 100160, China.
E-mail addresses: drfxm@163.com (X. Fang), radiology888@aliyun.com (X. Chen), tian@ieee.org (J. Tian).
[1] Yongbei Zhu, Chuntao Man, Lixin Gong, and Di Dong contributed equally to this work.

https://doi.org/10.1016/j.ejrad.2019.04.022
Received 14 January 2019; Received in revised form 9 April 2019; Accepted 29 April 2019

meningiomas take up only a few exhibiting aggressive growth behavior [3–5]. Patients with low- and high-grade meningiomas are significantly different in prognosis [6]. The survival rate is lower in high-grade meningiomas than that in the low-grade tumors [7]. Therefore, accurate preoperative grading is critical for predicting disease prognosis and may facilitate clinical decision-making.

Clinical characteristics and radiomic features have been used in Meningioma grading [8–14]. Imaging features, such as the absence or incompletement of peritumoral CSF signal around the tumor, especially the absence between the tumor and the adjacent brain tissues, may indicate the higher pathological grades. A research based on preoperative high-resolution gadolinium-enhanced T1WI 3D MRI showed that low intensity large area emphasis and high intensity large area emphasis were sensitive to low and high intensity large areas respectively [9]. Another study demonstrated that run length non-uniformity, grey-level non-uniformity and almost all of the selected histogram features had higher values in high-grade meningiomas [10]. The clinical characteristics, for example Ki67, were also proved to be relevant to grading [11].

Radiomic features [12,13] that were of great value in quantitating and analyzing radiography imaging in clinical oncology have been recognized as effective tools in differentiating tumor grades. Conventional radiomic features are based on hand-crafted radiomic (HCR) features, which can quantify tumor shape, intensity and texture information based on imaging [12]. According to the hypothesis of radiomics, heterogeneity in intra-tumor imaging could reflect potential genetic heterogeneity [13–15]. However, some intrinsic characteristics of intra-tumor imaging heterogeneity are difficult to be manually defined by low-order HCR features. In such a case, it is necessary to explore high-order features, such as DLR features, which may improve the prognosis performance of traditional radiomics model.

Recently, deep learning methods, especially convolutional neural networks, have rapidly become a promising technique for analyzing medical imaging [16–18]. Deep learning methods are composed of multi-types of self-learning units named ‘layers’ [19]. By hierarchical convolution operations of the image, valuable features of tumor that show promising prognostic value are extracted [20]. Compared with HCR features, more important information of tumor that may help the diagnoses was contained in DLR features [21].

Furthermore, clinical and HCR feature quantification based on preoperative high-resolution gadolinium-enhanced T1WI 3D MRI and ADC maps has been studied in Meningioma grading [9,10], while studies using deep learning techniques on routine MRI seldomly appeared. Thus, our study is to explore meningioma grading by DLR methods based on routine MRI.

In this study, we aimed to develop and validate a DLR model for meningioma grading based on routine post-contrast T1WI before operation.

2. Materials and methods

2.1. Patients

This retrospective study was approved by the institute review board of two hospitals. The requirement of informed consent was waived. The basic inclusion criteria for the primary and validation cohorts were as follows: (i) pathological diagnosis of intracranial meningiomas; (ii) previously untreated solitary enhancing extra-axial tumor; (iii) routine brain MR examination before surgical resection. The exclusion criteria were: (i) insufficient MRI qualities to obtain measurements (e.g., owing to motion artifacts); (ii) patients younger than 18; (iii) patients who had any previous relevant treatment history (including radiotherapy or surgery).

A total of 181 patients were identified for analysis, which were divided into two cohorts: a primarycohort (n = 99) and a validation cohort (n = 82). The primary cohort was consisted of 34 male and 65 female (mean age, 50.00 years; age range, 18–76 years), and the validation cohort comprised 23 male and 59 female (mean age, 57.37 years; age range, 28–80 years). The pathological grading was assured according to the updated 2016 WHO classification system [2].

2.2. MRI data acquisition and retrieval procedure

All MRI examinations were performed within two weeks before surgery. Enhanced and unenhanced sequences were included. The pre-contrast sequence included sagittal T1WI, axial T1WI, and T2WI. Once the pre-contrast imaging was completed, 0.2 ml/kg Gd-DTPA (Magnevist, Bayer Health Care Pharmaceuticals, Wayne, NJ) was administered manually via the antecubital vein by a registered nurse. Post-contrast images, including the coronal, sagittal, and axial images, were acquired immediately after the administration of contrast media. 181 patients were scanned on 3.0 T scanners (Discovery MR750 W, GE, Milwaukee, MI, USA, 29 cases; Cenesis Signa, GE, Milwaukee, MI, USA, 23 cases; TrioTim, Siemens Healthineers, Forchheim, Germany, 114 cases; Verio, Siemens Healthineers, Forchheim, Germany, 15 cases). T1-weighted sequence (TR/TE, 1800–2100 ms/2-20 ms) was performed with the same FOV (240 × 240 mm), matrix (512 × 512) and 8-channel head-matrix coil. The section thickness and gap were 5 mm and 6 mm, respectively, regardless of the scanner used.

2.3. Region of interest acquirement for DLR model

Region of interest (ROI) for meningioma refers to a tumor lesion area for quantitative analysis. For obtaining ROI, manual segmentation was performed on the meningioma images. The enhanced axial tumor images were imported into ITK-SNAP (http://www.itksnap.org) with the tumor lesion areas manually delineated by two experienced radiologists (5 years of experience and 10 years of experience) independently. The maximum area slice of the tumor lesion was selected and delineated and the adjacent bone invasion was excluded as much as possible, and then the software yielded a segmentation file in Nifti format. Each radiologist segmented the tumor images of one hospital. They were blind to each other and also the clinical and pathological information.

The inputs of DLR model should be rectangle images containing the corresponding tumor area, so the ROI was designed especially for satisfying the requirements of inputs. First, according to the tumor ROI, we located the tumor area of delineated slice using a rectangle bounding box and the bounding box should cover the primary tumor area in meningioma. The bias of manual segmentation made little difference to the rectangle bounding box. Afterward, we cropped three consecutive slices from the tumor images in DICOM format according to the bounding box and extracted the ROI of tumor for each patient. Finally, the tumor image was standardized by z-score normalization and scaled to a voxel size of 224 × 224 × 3. The extraction process of tumor ROI is shown in Fig. 1A.

2.4. Feature extraction for DLR model

DLR features generally referred to the outputs of convolutional neural networks’ hidden layers. In order to avoid over-fitting of the model, we employed the transfer learning strategy and used a pre-trained network to extract features [22]. For feature extraction, we used the Xception [23] tool box applied in Keras, which was a deep learning image classification model and pre-trained on ImageNet dataset (n = 1.3 million). Then we fed the tumor ROI into the deep learning network. The output of the deep learning network was 2048 features. The DLR feature extraction process is shown in Fig. 1B.

2.5. Dimension reduction of DLR features and building of DLR model

To avoid over-fitting, feature dimension reduction was necessary.

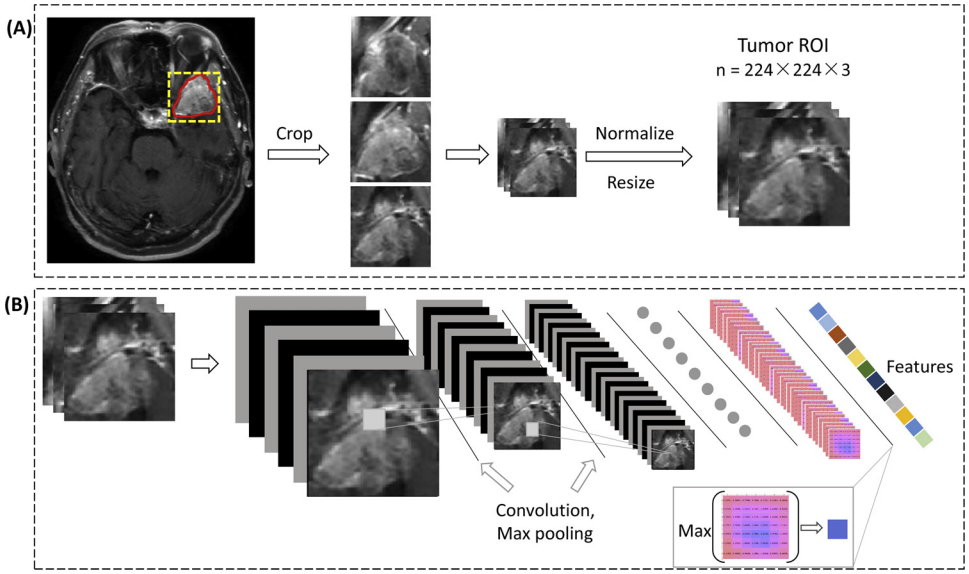

Fig. 1. (A) The extraction process of ROI. The tumor ROI was used as the inputs of neural net. (B) Deep learning feature extraction process.

Random forest algorithm was used to calculate the feature importance and the features with importance value above 0.001 were retained. Random forest algorithm constructed a multitude of decision trees and used the average of results to improve the predictive accuracy and control over-fitting. During building a Random forest, if the accuracy rate is greatly reduced when a feature is added with random noise, it indicates that this feature has relatively high importance and plays a significant role in the classification results. Furthermore, sequential backward selection (SBS, it begins with candidate features and sequentially eliminates the least important feature at each step until the desired number of features remain) method was performed to reduce the feature dimension [24]. This feature selection method started from the whole feature set and acquired the final feature set by removing the feature with the minimum decrease of F-measure in each step. F-measure was defined as a weighted average of the precision and recall, which was beneficial for our imbalanced data. The methods of random forest classifier and SBS method were performed on the primary cohort for feature selection.

Finally, classification was performed with linear discriminant analysis (LDA) method. Note that, bagging method was used for feature selection and model tuning on the primary cohort considering our imbalanced data, which built an average model based on multiple subsets sampling of the primary cohort [25]. The flow chart of DLR model is shown in Fig. 2.

2.6. Building of HCR model and combined model

We built a HCR model for comparison with DLR model. Similarly, the development of HCR model included a couple of steps: ROI acquirement, feature extraction, feature selection and model construction.

A total of 899 HCR features were extracted from manually segmented ROI using an open source python toolkit, pyradiomics (version 2.01) [26]. Considering our imbalanced data, oversampling technique was adopted in the training cohort by the method of synthetic minority over-sampling technique (SMOTE), which was benefit for feature selection and model building. A two-stage feature selection was performed for selecting the most effective features. First, with univariate analysis (Mann-Whitney U tests), features with $p < 0.05$ were selected. Then gradient boosting method was performed to select the most predictive features. At last, similar to the radiomics model proposed before [27–29], our radiomics model was built by the method of support vector machine (SVM) using three-fold cross-validation to avoid overfitting [30]. HCR feature extraction and selection are described in details in Supplementary Material 1.

Furthermore, we built a combined model using a SVM classifier by combining DLR signature and HCR features and explored whether feature fusion strategy could improve the prognosis performance.

2.7. Evaluation of performance and statistical analysis

The relationship between features and meningioma grades was assessed using univariate analysis. Mann-Whitney U test or independent *t*-test was used to assess the difference in continuous variables between the patient groups and Fisher's exact test or chi-square test was used to assess the difference in categorical variables. Receiver-operating characteristic (ROC) curve analysis was performed and the area under curve (AUC) was obtained to evaluate the predictive performance of models. Sensitivity and specificity values were used to assess the accuracy of each model. The calibration curve of the DLR signature for the predicted risks of high-grade meningiomas was plotted with the Hosmer-Lemeshow test. The comparison of ROC curves and stratified analysis were carried out by Delong test.

The statistical analysis was performed using Python software (version 3.6, https://www.python.org). Scikit-learn, statsmodels and scipy packages were used for statistical analysis [31]. Statistical significance was decided when the two-sided *P* value less than 0.05.

3. Results

No differences were observed in gender ($\chi^2 = 0.557$, $P = 0.455$) and the pathological grades ($\chi^2 = 0.793$, $P = 0.373$) between the primary and validation cohorts. Age of patient, however, was different ($P < 0.001$) (Table 1), which might be caused by regional differences and small-scale cohorts. In order to evaluate the influence of age, we performed stratified analysis for the subgroups split by the mean age of 53.3.

3.1. The diagnostic performance of DLR model

In feature dimension reduction, 316 features with importance values above 0.001 calculated by random forest algorithm were retained. Furthermore, 39 features were retained by SBS method. Then, LDA transformed 39 features into a DLR signature.

DLR signature was significantly correlated with meningioma

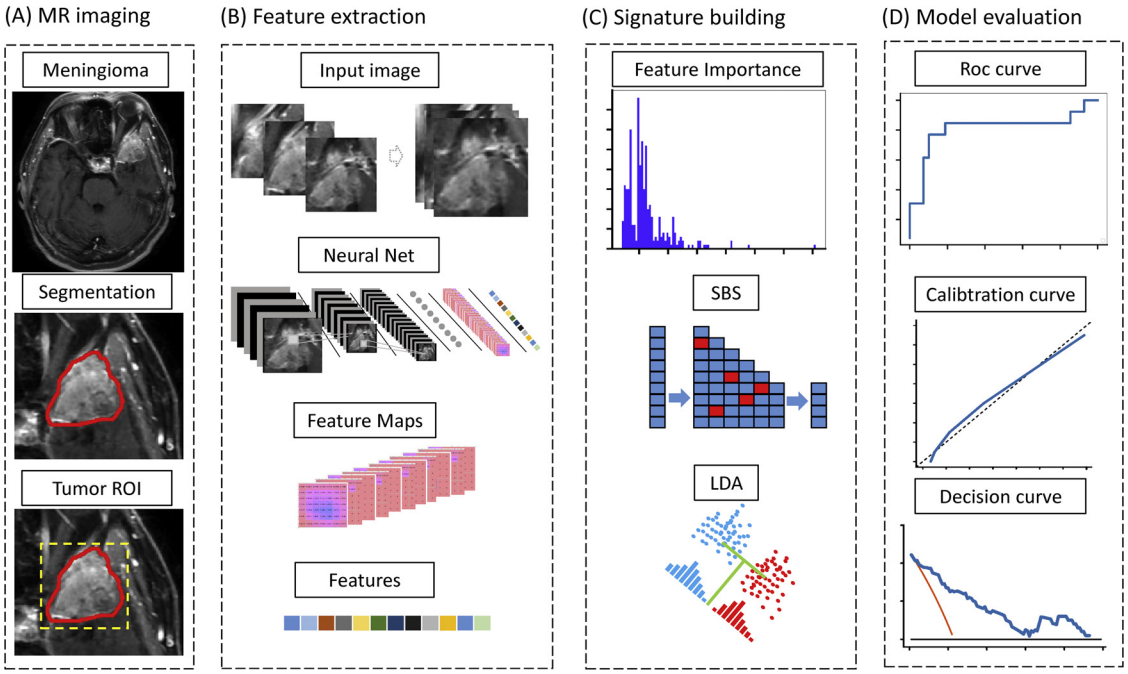

Fig. 2. DLR model workflow in this study. DLR features were extracted from Xception net, and random forest algorithm, sequential backward selection (SBS) method and linear discriminant analysis (LDA) were used for feature dimensionality reduction.

Table 1
Characteristics of patients with meningioma in primary and validation cohorts.

Characteristics	Primary cohort (n = 99)	Validation cohort (n = 82)	*P* value
Age, mean ± SD, years	50.00 ± 11.87	57.37 ± 12.30	< 0.001*
Female, No (%)	65/99 (65.6)	59/82(72.0)	0.455
Grade I, No (%)	77/99 (77.8)	69/82 (84.1)	0.373
Grade II, No (%)	17/99 (17.2)	10/82 (12.2)	
Grade III, No (%)	5/99 (5.0)	3/82 (3.7)	

The pathological grading was categorized according to the 2016 WHO classification of Tumors of the Central Nervous System. *P* values were calculated between primary and validation cohorts. Age was calculated by Student's *t*-test and gender and WHO grade were calculated by Fisher's exact test (chi-squared test).

* $P < 0.05$.

pathology with an AUC of 0.891 (95% CI 0.818–0.963, $p < 0.001$) in the primary cohort, which was then confirmed in the validation cohort with an AUC of 0.811 (95% CI 0.635–0.986, $p < 0.001$). The sensitivity and specificity results were 0.769 and 0.898 respectively in the validation cohort as well as 0.826 and 0.857 respectively in the primary cohort. The detailed diagnostic performance of DLR signature is shown in Fig. 3A and Table 2.

The calibration curve of DLR signature for the predicted risk of high-grade meningioma demonstrated good agreements between predictions and the observed outcomes of high-grade meningioma (Fig. 3B, C), and the Hosmer Lemeshow test showed good similarity ($p = 0.737$).

3.2. Stratified analysis of DLR model

Stratified analysis for the subgroups split by mean age: young and old subgroups got AUCs of 0.877 and 0.815, respectively (Delong test p-value: 0.629 and 0.653 compared with the results on the overall cohort).

Stratified analysis on version of MR system: patients scanned using GE MR system and SIEMENS MR system got AUCs of 0.847 and 0.788, respectively (Delong test p-value: 0.427 and 0.997 compared with the results on the overall cohort).

The results of above stratified analyses showed that our model was not affected by age and scanning parameters.

3.3. Comparison of DLR model with HCR model

The HCR model showed good performances in discriminating between high-grade and low-grade meningimas in the primary and validation cohorts with AUCs of 0.716 (95% CI 0.597–0.834, $p = 0.002$) and 0.678 (95% CI 0.520–0.835, $p = 0.038$), respectively. The results are shown in Fig. 3A and Table 2.

The DLR model could significantly increase the efficiency of diagnosis compared with HCR model (AUC = 0.811 vs 0.678, sensitivity = 0.769 vs 0.615, specificity = 0.898 vs 0.710).

The construction of HCR model included a two-stage feature selection. First, with univariate analysis, 618 features with $p < 0.05$ were selected as potential and informative features. Then, six HCR features were remained after gradient boosting (range AUC = 0.59 to 0.77; one shape feature and five second-order features calculated from the gray level size zone matrix (GLSZM), gray level co-occurrence matrix (GLCM), and gray level run length matrix (GLRLM)). HCR feature selection process and violin plots for the selected features are shown in Fig. 4.

3.4. Combination of DLR signature and HCR features

Based on the combination of HCR features and DLR signature, the diagnosis efficiency was slightly elevated with an AUC value from 0.811 to 0.816, a sensitivity value from 0.769 to 0.846 and a specificity value from 0.898 to 0.797. The detailed diagnostic performance of the combined model is shown in Fig. 3A and Table 2. The ROC curves for DLR model and combined model were compared using Delong test. The results showed no significant difference with a *P* value of 0.503 between two models.

4. Discussion

In this retrospective multicenter research, a DLR model and a HCR model were built. For assessing meningioma grading, DLR model

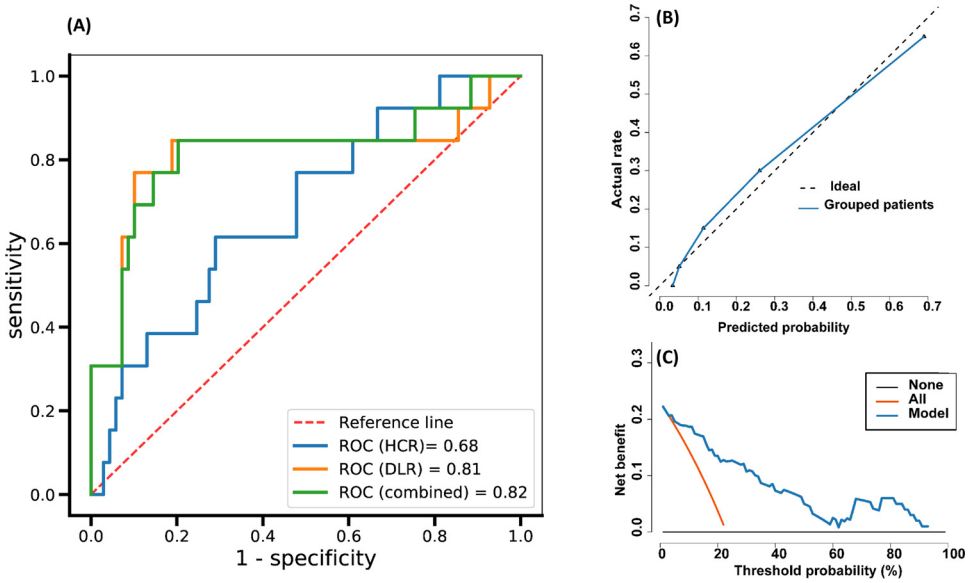

Fig. 3. Results of ROC curve analysis for three models and evaluation of HCR model. (A) The ROC curves for HCR signature, DLR signature and combined signature. (B, C) Calibration curve and decision curve of DLR model.

Table 2
The results of three models for meningioma grading.

Model	AUC(95% CI)		Sensitivity	Specificity	*P* value
	Primary	Validation			
DLR	0.891(0.818, 0.963)	0.811(0.635, 0.986)	0.769	0.898	< 0.001*
HCR	0.716(0.597,0.834)	0.678(0.520,0.835)	0.615	0.710	0.038
Combined	0.900(0.831, 0.969)	0.816(0.654, 0.978)	0.846	0.797	< 0.001*

Abbreviation: CI, confidence interval.
The combined model integrated HCR features and DLR signature. Sensitivity and specificity were used to assess the accuracy in the validation cohort. *P* values were calculated between high-grade and low-grade groups using predict probability.
* $P < 0.05$.

demonstrated significant improvements compared with the HCR model. The presurgical meningioma grading may facilitate making clinical decisions and promote prognosis, thus facilitating counseling in the earlier stage of clinical care.

In our study, age was significantly different between the primary and validation cohorts in univariate analysis, which was then proved to have no significant difference in subgroup analysis. The information of mean age, gender proportion and WHO grading were consistent with those in the previous studies [32,33]. As to scanning parameters, stratified analysis showed no significance as well.

Our study showed that DLR signature could characterize intratumor imaging heterogeneity and differentiate meningioma grades well. Deep learning networks comprising hundreds of self-learning units had advantages in quantifying the prognostic features of meningioma grading that could not be manually defined. In other studies, deep learning in brain tumors image analyses achieved excellent performances. For instance, DLRs could accurately and automatically detect and segment brain tumors in despite of diverse scanner data [34,35]. Also, the learned features from DLR played important roles in accurately predicting types of brain tumor and the survival time of patients with brain tumor [36,37]. Our results kept consistence with above studies, which demonstrated that deep learning model was able to quantify meningioma heterogeneity and differentiated meningioma grades. Considering that the clinical manifestations of other intra-axial tumors were also closely related to the heterogeneity in imaging phenotype [38,39], this suggested the potential of deep learning model in clinical diagnosis of more intra-axial tumors.

On the other hand, HCR features were also useful in diagnosis of this clinical problem though with a poorer performance than DLR features. According to the experimental results in our study, radiomic features with lower values were shown to have a close connection with histologic high-grade meningioma to some extent. For one, high-grade meningiomas had lower sphericity values (representing irregularly shaped tumors) than low-grade cases with an obvious statistical difference ($p = 0.001$). For another, lower LGLRE values (indicating a sparse distribution of low gray-level values) in the image with a significant difference ($p = 0.03$) may represent that heterogeneous meningiomas with necrosis and/or hemorrhage were most likely belonging to high grade. The good results of HCR model could guide better model designs in the future, such as feature fusion model or more appropriate neural network model.

A preliminary experiment of fusion model was carried out in our study. We found that the combination of the two models had a slight improvement with an AUC of 0.816 (sensitivity 0.846, specificity 0.797). So it made sense to adopt better fusion method or introduce other types of variables. A research conducted by Coroller TP et al combines semantic and HCR features using random forest classifiers and the combination of the two improved diagnostic performance very well, which showed that both semantic and HCR features were complementary [8]. In our study, DLR and HCR features captured high-order and low-order features of tumor image respectively. Maybe HCR features and DLR features are both derived from quantitative measurement for tumor, thus giving only a slight effect of feature complementation. The study results also reflected the efficacy differences between the DLR and the HCR in quantifying tumor phenotypes. Although we did not carry out additional optimization for the network,

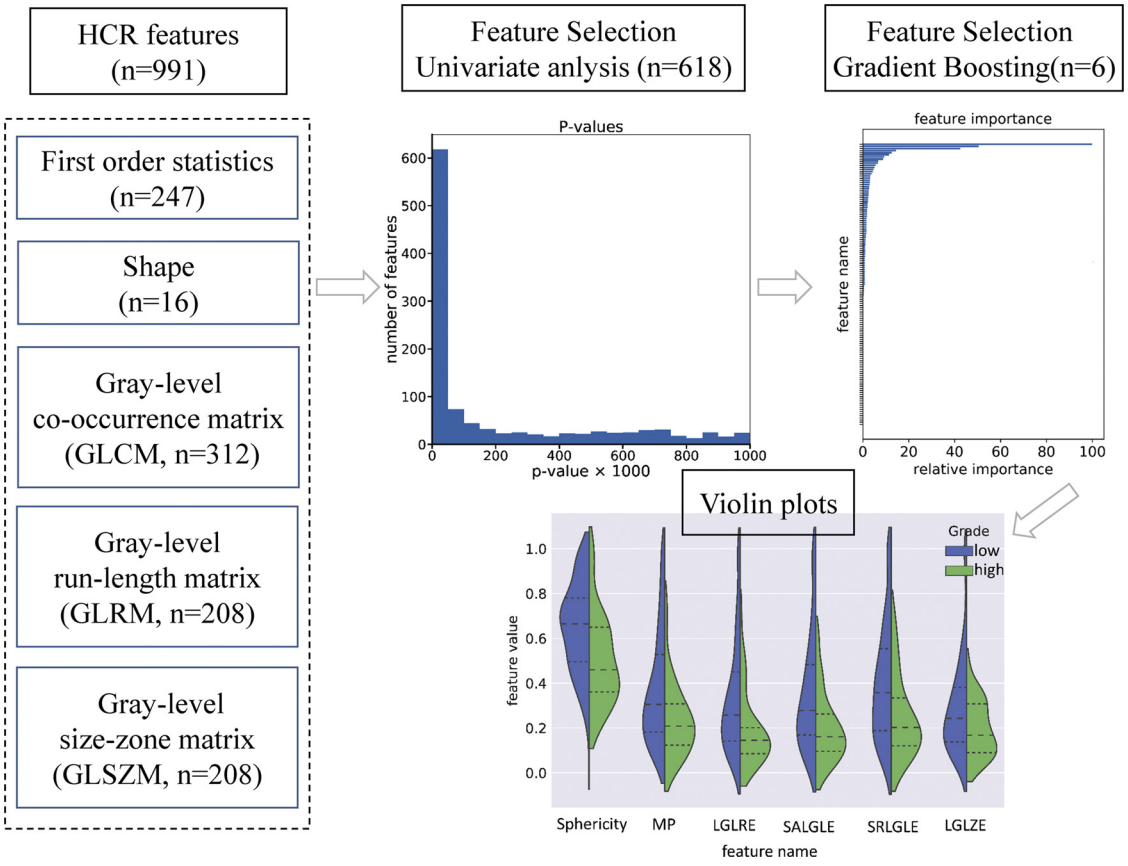

Fig. 4. HCR feature selection process and violin plots for the selected features.

the deep learning features still achieved a quantization capability superior to the hand-crafted features, indicating that the DLR could dig out more image features related to the tumor pathology. Considering the development of DLR, its clinical application value in various types of tumors will be further explored.

Meningioma grading by radiological images is practicable. A research conducted by Okuchi S et al. showed that Tl-SPECT had a higher diagnostic capability of meningioma grades compared with FDG-PET [6]. Meningioma grading could also be done by high-resolution post-contrast T1WI and ADC maps [9,10]. Our study was based on routine enhanced T1WI, which was more convenient than high-resolution scanning. Therefore, it was more practicable in daily medical work.

Our research had some limitations. Its retrospective nature and potential bias came first. Secondly, clinical characteristics were not used in experiments due to the incomplete clinical information. What is more, features were extracted from three consecutive slices rather than all lesion slices, which might not be able to reflect tumor characteristics comprehensively. Furthermore, we had a limited amount of data in this study. In order to prevent the model from overfitting, there was no further refinement of the deep learning classification model. More training data would strengthen diagnostic performance in the future.

In conclusion, we employed transfer learning strategy and extracted deep learning features to build a DLR for preoperative grading in meningioma. For comparison, a radiomic model based on hand-crafted features was built. Our study demonstrated that, based on routine post-contrast T1WI, meningioma pathological grades can be well defined by DL model, which had better performance than HCR model.

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by the National Key R&D Program of China [grant numbers 2017YFC1308700, 2017YFA0205200, 2017YFC1309100, 2017YFC0114300, 2018YFC0115604]; National Natural Science Foundation of China [grant numbers 81771924, 81501616, 81227901, 81671854, 81772005, 81271629]; the Beijing Natural Science Foundation [grant number L182061]; the Bureau of International Cooperation of Chinese Academy of Sciences [grant number 173211KYSB20160053]; the Instrument Developing Project of the Chinese Academy of Sciences [grant number YZ201502]; the Youth Innovation Promotion Association CAS [grant number 2017175]; and Natural Science Foundation of Heilongjiang Province [grant number F201216].

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejrad.2019.04.022.

References

[1] A. Drevelegas, Extra-axial brain tumors, Eur. Radiol. 15 (2005) 453–467.
[2] D.N. Louis, A. Perry, G. Reifenberger, A. Von Deimling, D. Figarella-Branger, W.K. Cavenee, et al., The 2016 World Health Organization classification of tumors of the central nervous system: a summary, Acta Neuropathol. 131 (2016) 803–820.
[3] Perry A. Meningiomas, Practical Surgical Neuropathology: A Diagnostic Approach, Elsevier, 2018, pp. 259–298.
[4] C. Champeaux, L. Dunn, World Health Organization grade II meningioma: a 10-year retrospective study for recurrence and prognostic factor assessment, World Neurosurg. 89 (2016) 180–186.
[5] L.B. Nabors, J. Portnow, M. Ammirati, J. Baehring, H. Brem, N. Butowski, et al., NCCN guidelines Insights: central nervous system cancers, version 1. 2017, J. Natl. Compr. Cancer Netw. 15 (2017) 1331–1345.
[6] S. Okuchi, T. Okada, A. Yamamoto, M. Kanagaki, Y. Fushimi, T. Okada, et al.,

Grading meningioma: a comparative study of thallium-SPECT and FDG-PET, Medicine 94 (2015).
[7] D.N. Louis, H. Ohgaki, O.D. Wiestler, W.K. Cavenee, P.C. Burger, A. Jouvet, et al., The 2007 WHO classification of tumours of the central nervous system, Acta Neuropathol. 114 (2007) 97–109.
[8] W.L. Bi, T. Corroller, N.F. Greenwald, E. Huynh, M. Abedalthagafi, A. Aizer, et al., Radiographic prediction of meningioma grade and genomic profile, J. Neurol. Surg. Part B Skull Base 78 (2017) A109.
[9] T.P. Coroller, W.L. Bi, E. Huynh, M. Abedalthagafi, A.A. Aizer, N.F. Greenwald, et al., Radiographic prediction of meningioma grade by semantic and radiomic features, PLoS One 12 (2017) e0187908.
[10] Y. Lu, L. Liu, S. Luan, J. Xiong, D. Geng, B. Yin, The diagnostic value of texture analysis in predicting WHO grades of meningiomas based on ADC maps: an attempt using decision tree and decision forest, Eur. Radiol. (2018) 1–11.
[11] H. Kolles, I. Niedermayer, C. Schmitt, W. Henn, R. Feld, W. Steudel, et al., Triple approach for diagnosis and grading of meningiomas: histology, morphometry of Ki-67/Feulgen stainings, and cytogenetics, Acta Neurochir. 137 (1995) 174–181.
[12] H.J. Aerts, The potential of radiomic-based phenotyping in precision medicine: a review, JAMA Oncol. 2 (2016) 1636–1642.
[13] J.D. Song, J.Y. Shi, D. Dong, M.J. Fang, W.Z. Zhong, K. Wang, et al., A new approach to predict progression-free survival in stage IV EGFR-mutant NSCLC patients with EGFR-TKI therapy, Clin. Cancer Res. 24 (15) (2018) 3583–3592.
[14] D. Dong, D. Tang, Z.Y. Li, M.J. Fang, J.B. Gao, X.H. Shan, et al., Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer, Ann. Oncol. 30 (2019) 431–438.
[15] B. Zhang, J. Tian, D. Dong, D.S. Gu, Y.H. Dong, et al., Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma, Clin. Cancer Res. 23 (2017) 4259–4269.
[16] D.S. Kermany, M. Goldbaum, W. Cai, C.C. Valentim, H. Liang, S.L. Baxter, et al., Identifying medical diagnoses and treatable diseases by image-based deep learning, Cell 172 (2018) 1122–1131 e9.
[17] S. Wang, J.Y. Shi, Z.X. Ye, D. Dong, D.D. Yu, M. Zhou, et al., Predicting EGFR mutation status in lung adenocarcinoma on CT image using deep learning, Eur. Respir. J. 53 (3) (2019) 1800986.
[18] S. Wang, M. Zhou, Z. Liu, Z. Liu, D. Gu, Y. Zang, et al., Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation, Med. Image Anal. 40 (2017) 172–183.
[19] R.F. Thompson, G. Valdes, C.D. Fuller, C.M. Carpenter, O. Morin, S. Aneja, et al., Artificial intelligence in radiation oncology: a specialty-wide disruptive transformation? Radiother. Oncol. (2018).
[20] K. Chaudhary, O.B. Poirion, L. Lu, L.X. Garmire, Deep learning–based multi-omics integration robustly predicts survival in liver cancer, Clin. Cancer Res. 24 (2018) 1248–1259.
[21] R. Poplin, A.V. Varadarajan, K. Blumer, Y. Liu, M.V. McConnell, G.S. Corrado, et al., Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning, Nat. Biomed. Eng. 2 (2018) 158.
[22] A.P. Dhawan, Medical Image Analysis, John Wiley & Sons, 2011.
[23] F. Chollet, Xception: deep learning with depthwise separable convolutions, arXiv preprint 1610 (2017) 02357.
[24] S.J. Reeves, Z. Zhe, Sequential algorithms for observation selection, IEEE Trans. Signal Process. 47 (1999) 123–132.
[25] L. Breiman, Bagging predictors, Mach. Learn. 24 (1996) 123–140.
[26] J.J. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, et al., Computational radiomics system to decode the radiographic phenotype, Cancer Res. 77 (2017) e104–e107.
[27] J. Wang, C.-J. Wu, M.-L. Bao, J. Zhang, X.-N. Wang, Y.-D. Zhang, Machine learning-based analysis of MR radiomics can help to improve the diagnostic performance of PI-RADS v2 in clinically relevant prostate cancer, Eur. Radiol. 27 (2017) 4082–4090.
[28] M. Yuan, Y.-D. Zhang, X.-H. Pu, Y. Zhong, H. Li, J.-F. Wu, et al., Comparison of a radiomic biomarker with volumetric analysis for decoding tumour phenotypes of lung adenocarcinoma with different disease-specific survival, Eur. Radiol. 27 (2017) 4857–4865.
[29] L. Yang, D. Dong, M. Fang, Y. Zhu, Y. Zang, Z. Liu, et al., Can CT-based radiomics signature predict KRAS/NRAS/BRAF mutations in colorectal cancer? Eur. Radiol. 28 (2018) 2058–2067.
[30] A. Statnikov, C.F. Aliferis, Are random forests better than support vector machines for microarray-based cancer classification? AMIA Annual Symposium Proceedings: American Medical Informatics Association, (2007), p. 686.
[31] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, et al., Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[32] K. Gao, H. Ma, Y. Cui, X. Chen, J. Ma, J. Dai, Meningiomas of the cerebellopontine angle: radiological differences in tumors with internal auditory canal involvement and their influence on surgical outcome, PLoS One 10 (2015) e0122949.
[33] G. Kaur, E.T. Sayegh, A. Larson, O. Bloch, M. Madden, M.Z. Sun, et al., Adjuvant radiotherapy for atypical and malignant meningiomas: a systematic review, Neuro-oncology 16 (2014) 628–636.
[34] K.R. Laukamp, F. Thiele, G. Shakirin, D. Zopfs, A. Faymonville, M. Timmer, et al., Fully automated detection and segmentation of meningiomas using deep learning on routine multiparametric MRI, Eur. Radiol. 29 (2019) 124–132.
[35] Z. Akkus, A. Galimzianova, A. Hoogi, D.L. Rubin, B.J. Erickson, Deep learning for brain MRI segmentation: state of the art and future directions, J. Digit. Imaging 30 (2017) 449–459.
[36] H. Mohsen, E.-S.A. El-Dahshan, E.-S.M. El-Horbaty, A.-B.M. Salem, Classification using deep learning neural networks for brain tumors, Future Comput. Inform. J. 3 (2018) 68–71.
[37] D. Nie, H. Zhang, E. Adeli, L. Liu, D. Shen, 3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients, International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 212–220.
[38] H.J. Aerts, E.R. Velazquez, R.T. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, et al., Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, Nat. Commun. 5 (2014) 4006.
[39] K. Chang, H.X. Bai, H. Zhou, C. Su, W.L. Bi, E. Agbodza, et al., Residual convolutional neural network for the determination of IDH status in low-and high-grade gliomas from MR imaging, Clin. Cancer Res. 24 (2018) 1073–1081.