RESEARCH

Deep learning–based automatic segmentation of meningioma from T1-weighted contrast-enhanced MRI for preoperative meningioma differentiation using radiomic features

Liping Yang[1†], Tianzuo Wang[2†], Jinling Zhang[3], Shi Kang[4], Shichuan Xu[5*] and Kezheng Wang[1*]

Abstract

Background This study aimed to establish a dedicated deep-learning model (DLM) on routine magnetic resonance imaging (MRI) data to investigate DLM performance in automated detection and segmentation of meningiomas in comparison to manual segmentations. Another purpose of our work was to develop a radiomics model based on the radiomics features extracted from automatic segmentation to differentiate low- and high-grade meningiomas before surgery.

Materials A total of 326 patients with pathologically confirmed meningiomas were enrolled. Samples were randomly split with a 6:2:2 ratio to the training set, validation set, and test set. Volumetric regions of interest (VOIs) were manually drawn on each slice using the ITK-SNAP software. An automatic segmentation model based on SegResNet was developed for the meningioma segmentation. Segmentation performance was evaluated by dice coefficient and 95% Hausdorff distance. Intra class correlation (ICC) analysis was applied to assess the agreement between radiomic features from manual and automatic segmentations. Radiomics features derived from automatic segmentation were extracted by pyradiomics. After feature selection, a model for meningiomas grading was built.

Results The DLM detected meningiomas in all cases. For automatic segmentation, the mean dice coefficient and 95% Hausdorff distance were 0.881 (95% CI: 0.851–0.981) and 2.016 (95% CI:1.439–3.158) in the test set, respectively. Features extracted on manual and automatic segmentation are comparable: the average ICC value was 0.804 (range, 0.636–0.933). Features extracted on manual and automatic segmentation are comparable: the average ICC value was 0.804 (range, 0.636–0.933). For meningioma classification, the radiomics model based on automatic segmentation performed well in grading meningiomas, yielding a sensitivity, specificity, accuracy, and area under the curve (AUC) of 0.778 (95% CI: 0.701–0.856), 0.860 (95% CI: 0.722–0.908), 0.848 (95% CI: 0.715–0.903) and 0.842 (95% CI: 0.807–0.895) in the test set, respectively.

---

†Liping Yang and Tianzuo Wang contributed equally to this work.

*Correspondence:
Shichuan Xu
69744792@qq.com
Kezheng Wang
wangkezheng9954001@163.com
Full list of author information is available at the end of the article

Yang *et al. BMC Medical Imaging*   (2024) 24:56

Page 2 of 12

Conclusions The DLM yielded favorable automated detection and segmentation of meningioma and can help deploy radiomics for preoperative meningioma differentiation in clinical practice.

Keywords Meningiomas, Magnetic resonance imaging, Radiomics, Deep learning, Segmentation

Introduction

Meningioma, originated from the cell of the middle layer of meninges and the arachnoid, is the most common histopathological type of adult central nervous system (CNS) tumor [1]. The 2021 World Health Organization (WHO) classification system of CNS tumors (5th edition) has not substantially changed the management of meningioma patients, and the WHO grades still guide its treatment decisions, which affects the patients' survival prognosis to a certain extent [2]. Compared with low-grade meningiomas (WHO grade I), high-grade meningiomas (WHO II/III grade) exhibit more aggressive biological behavior, a more pronounced tendency to recur, and a poorer clinical prognosis [3–5]. According to statistical reports, the 10-year survival rate of malignant (high-grade) meningiomas is only 60%, while for non-malignant (low-grade) meningiomas, it is approximately 83.4% [6]. In clinical practice, non-invasive and accurate identification of the grade of meningioma is of great significance.

Magnetic resonance imaging (MRI) technology has been widely used in the preoperative diagnosis of meningiomas due to its excellent soft tissue resolution [7]. Especially, the T1-weighted contrast-enhanced (T1W-CE) sequence could provide a large amount of blood supply information to display tumor tissue as clearly as possible [8]. However, it has to be admitted that the imaging manifestations of different grades of meningiomas are largely overlapped, leading to inaccurate grading and improper treatment measures [9]. So far, there is still a lack of a widely accepted method to accurately predict histopathological grading in clinical practice.

Recently, radiomics and deep learning (DL), two main categories of machine learning (ML), have rapidly developed into a research hotspot in medical image analysis, enabling the extraction of high-throughput quantitative imaging features from medical image [10–12]. It captures relationships between image voxels that may not be perceived by the naked eye of physicians-even experienced radiologists, which can contribute to the diagnostic and predictive accuracy of the disease. A previous study had proven that radiomics is a valid tool for grading meningioma, and it has outperformed the subjective diagnosis of experienced doctors [13]. Tumor segmentation is the first and major step in radiomics analysis, but the manual segmentation of tumor lesions is a laborious ordeal and time-consuming, and high inter- and intra-reader variability is not negligible [14]. It is worth considering whether the DL-based automatic segmentation in the radiomic analysis of meningioma differentiation can replace time-consuming manual segmentation. The neural network with a U-shape architecture is a promising tool for automatic segmentation even under the condition of limited sample size, as has been reported in glioma patients [15–17]. To the best of our knowledge, few studies attempted to develop an ML pipeline for meningiomas grading using automatic segmentation. Therefore, we decided to establish a dedicated deep-learning model (DLM) on routine MRI data to investigate DLM performance in automated detection and segmentation of meningiomas in comparison to manual segmentations. Another purpose of our work was to develop a radiomics model based on the radiomics features extracted from automatic segmentation to differentiate low- and high-grade meningiomas before surgery.

Material and methods

Study population

Ethical approval was obtained for this retrospective study, and the need for written informed consent was waived. Specific inclusion criteria were listed as follows: (1) patients diagnosed with meningioma by histopathology and with definite WHO grading [18]; (2) previously untreated solitary primary tumor before MRI scans; (3) available axial T1W-CE images; (4) satisfactory image quality and no artifacts for each patient. The exclusion criteria were as follows: (1) A history of relevant treatment before the preoperative MRI examinations; (2) multiple lesions; (3) Patients with metal foreign bodies or claustrophobia. The flowchart of patient selection is displayed in Fig. 1. The flowchart of ML pipeline is presented in Fig. 2.

MRI acquisition

The examinations were performed by using three 3.0 T MRI scanners (GE Signa HDi, GE Signa HDxt, and GE Discovery MR750w). All MRI examinations were performed within two weeks before surgery. Detailed scanning parameters were: TR = 1665 ms, TE = 25 ms, FOV = 25 cm, slice thickness = 5 mm and matrix = 256 × 256. Each patient was given a dose of 0.1 mmol/kg of gadopentetate dimeglumine at a rate of 3.5 ml/s. Subsequently, the T1W-CE scanning was performed within 250 s.
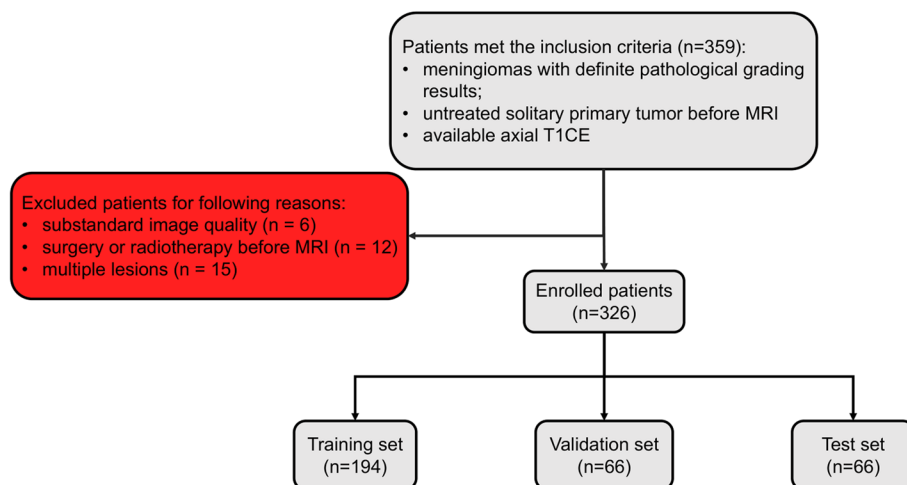
Yang *et al. BMC Medical Imaging*   (2024) 24:56

Page 3 of 12
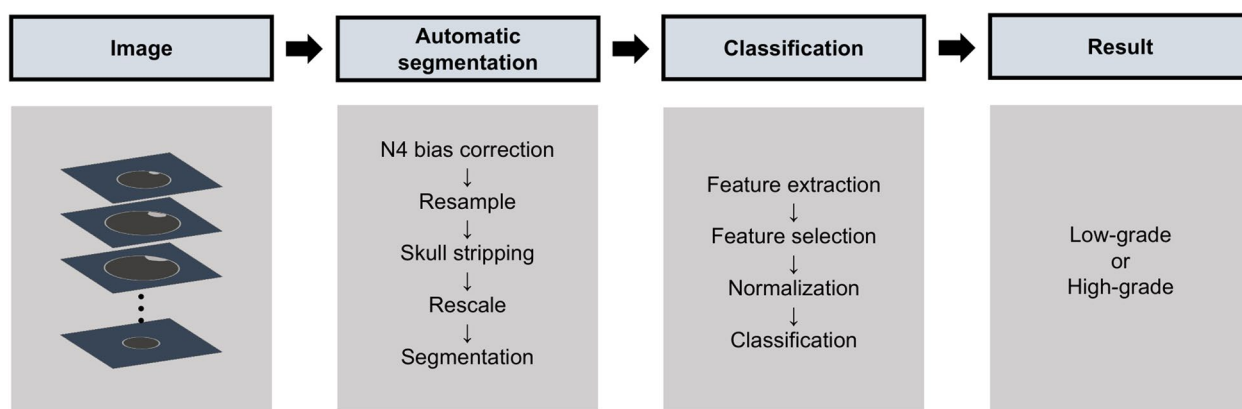

Fig. 1 Flowchart of patient selection

Fig. 2 Flowchart of the ML pipeline

Image pre-processing

Before tumor segmentation, the image preprocessing step is first performed, including: (1) resampling to a $1{\times}1{\times}1$ mm^3 resolution; (2) N4 bias correction using SimpleITK software (version 2.0, https:// www. simpl eitk. org/), it could correct low frequency intensity non-uniformity present in MRI image data known as a bias or gain field; (3) skull stripping was performed using the multi-contrast brain STRipping (MONSTR); (4) resize to $256{\times}256{\times}16$; (5) rescale (range 0–1).

Manual segmentation

Volumetric regions of interest (VOIs) of training and validation sets were manually segmented by using dedicated ITK-SNAP software (version 3.8.0; www.itksn ap.org). VOIs were delineated along the boundary of

tumor lesions slice by slice on axial images. The adjacent invasion, necrosis, and peritumoral edema were excluded from the VOIs. Two radiologists (both of who have 15 years of experience in brain MRI interpretation), blinded to the pathological results, drew VOIs jointly. Next, an expert radiologist with 15 years of experience in brain oncology reviewed the results.

Automatic segmentation

SegResNet, a 3D U-net-like network with a ResNet-like block, was applied to develop the automatic segmentation model, whose code was available on GitHub (https://github.com/Project-MONAI/MONAI) [18]. The architecture of this algorithm is shown in Fig. 3. Manual segmentation is used as the reference standard, and the loss function we use is Dice Loss, which represents
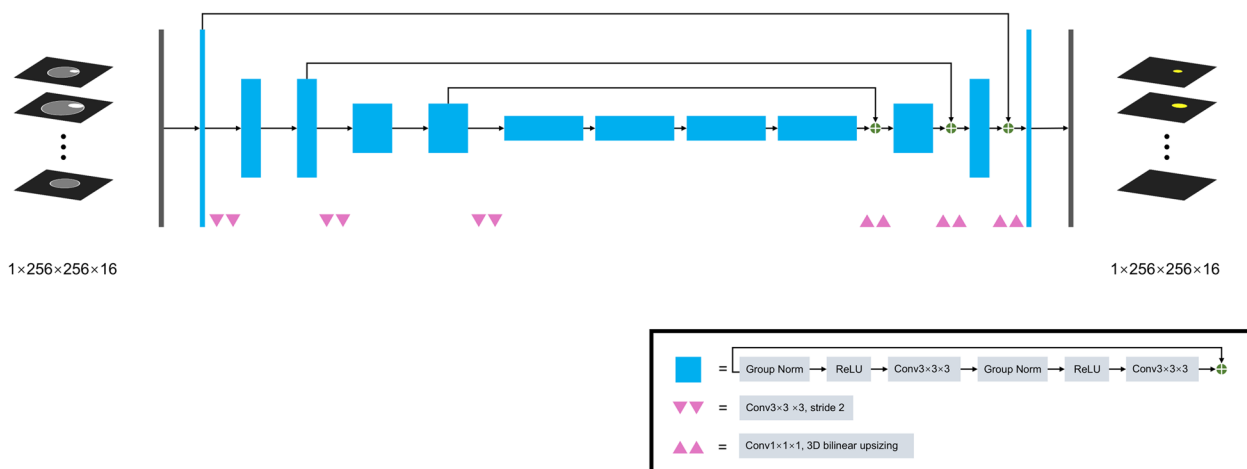
Fig. 3 SegResNet architecture

the overlap between the two groups and can reduce the impact of imbalanced data.

The equation is displayed as follows:

$$\text{Dice Loss} = 1 - \frac{2|A \cap B|}{|A| + |B|}$$

where $A \cap B$ is the intersection of two sets, $|\ |$ represents the number of elements in the set. The batch size was set as 7. We used an AdamW optimizer with an initial learning rate of 0.01 and weight decay of 0.00001. A cosine scheduler with warmup was utilized to dynamically adjust the learning rate according to the epoch. The total epoch and warmup steps were 400 and 40, respectively. The data of the training set were fed into the network to train the model. During training, we implemented data augmentation by randomly mirror flipping and rotating (degree range -45 to 45), which was able to improve the generalization ability. The random probability was 0.5. According to the mean of the Dice coefficient in the validation set, the best model was chosen for automatic segmentation. Our network was developed using the python package "Pytorch" and trained using RTX-2080Ti in the cloud computation platform "AI-Galaxy" (http://www.ai-galaxy.cn/) [19].

Radiomics analysis

All of the study patients were categorized into low-grade (grade I) and high-grade (grades II and III) groups. Since there were relatively few high-grade meningiomas in this study, we integrated the training and validation sets for radiomics analysis. Radiomics features were extracted from by using the python package "Pyradiomics" (https://pyradiomics.Readthedocs.io/). The scheme of $\mu \pm 3\sigma$ (μ: mean; standard deviation) was adapted to reduce noise disturbance [20]. A total of 1688 radiomics features

were extracted from each VOIs, including seven categories: first-order, shape, gray-level cooccurrence matrix (GLCM), gray-level dependence matrix (GLDM), gray-level size zone matrix (GLSZM), gray-level run-length matrix (GLRLM), and neighboring gray-tone difference matrix (NGTDM). The filter we used included exponential, square, squareroot, gradient, lbp, logarithm and wavelet. To avoid adverse influence from the different value scales and outliers of the radiomics features, all of them were transformed using robust standardization by the median, 25th percentiles, and 75th percentiles of each feature in training data. The equation is listed as follows: we integrated

$$\text{value} = \frac{x - \text{median}}{p75 - p25},$$

where x represents the feature value, while p75 and p25 are the 75th percentile and 25th percentile, respectively.

The intra- and inter-class correlation coefficients (ICCs) was calculated to screen out the stable radiomics features. For sixty cases of meningiomas (30 low-grade and 30 high-grade) were selected randomly for ICC calculation. Two neuroradiologists (neuroradiologists A and B) extracted the radiomics features independently, and neuroradiologist A re-extracted features two weeks later. Finally, the radiomics features with both inter-ICC and intra-ICC are greater than 0.8 were chosen for further analysis. The optimal radiomics feature set was determined using a four-step feature selection method. Firstly, the features with zero variance were discarded. Secondly, the one-way variance of analysis (ANOVA) *P*-value between labels and features was calculated for classification tasks and the features with $P > 0.05$ were excluded. Next, the mutual information value of each feature was calculated by measuring the dependency between feature

Yang *et al. BMC Medical Imaging*   (2024) 24:56

Page 5 of 12

and target, and the features ranked in the top 100 were retained. Finally, the most effective combination of the radiomics features was selected using recursive feature elimination with tenfold cross-validation (RFECV) based on the AUC mean. Due to the low proportion of high-grade meningioma patients, the synthetic minority oversampling technique (SMOTE) was employed to obtain smoother data for training the model after the preliminary feature selection process in training set. The SMOTE technique is an effective oversampling method commonly employed in medical applications to address the issue of class-imbalanced data. It works by augmenting the number of data instances in the minority class through the generation of synthetic data points from its nearest neighbors using Euclidean distance [21, 22]. Furthermore, EasyEnsemble classifier was also applied to improve the predictive performance considering the imbalance of data distribution, this approach effectively solves the problem of unbalanced data types and reduces the loss of information due to undersampling. The classification models based on lightGBM and EasyEnsembleClassifier algorithm were developed using the optimal feature combination.

Pipeline

When a preprocessed image was inputted into the DL-based segmentation model, the VOIs could be delineated automatically. Subsequently, a radiomics model for grading meningioma was developed using the radiomics features extracted from automatic segmentation. It was used to predict meningioma grade according to automatic segmentation images. For comparison, a radiomics model for grading meningioma was also constructed based on the radiomics features extracted from manual segmentation.

Performance evaluation

For the automatic segmentation model, dice coefficient and 95% Hausdorff distance (95HD) were utilized to evaluate the predictive performance of the developed models. Besides, ICC analysis was applied to assess the agreement between radiomic features from manual and automatic segmentations in the test set. The diagnostic performance of the radiomics models was evaluated based on the receiver operating characteristics (ROC) curve. The sensitivity, specificity and accuracy were calculated in the training, validation, and test cohorts, respectively.

Statistics

In this study, all the statistical analyses were achieved with the scikit-learn package in Python (version 3.8, https://www.python.org/) [23, 24]. The chi-squared test and ANOVA test were used to evaluate the difference among different sets for categorical and continuous variables, respectively. For automatic segmentation model, dice coefficient and 95% Hausdorff distance (95HD) were utilized to evaluate its predictive performance. Besides, ICC analysis was applied to assess the agreement between radiomics features extracted from manual segmentation and automatic segmentation in the test set. The ROC curve was used to evaluate the discriminative performance of the predictive model. The sensitivity, specificity and accuracy were calculated in the training, validation, and test cohorts, respectively. The calibration curve was applied to assess the agreement between the prediction results of the radiomics models and the actual clinical findings, and decision curve analysis (DCA) was used to validate the clinical usefulness of the radiomics models. Comparison of the different radiomics models based on manual segmentation and automatic segmentation using DeLong test. A two-sided *P* value < 0.01 was used as the criterion to indicate a statistically significant difference.

Results

Patients' characteristics

A total of 326 patients who met inclusion criteria from January 2017 to February 2022 were consecutively collected in this retrospective study, composed of 93 males and 233 females (ranging from 34 to 72 years old, median 56). Forty-three patients were assigned to high-grade meningiomas, and 283 patients were confirmed with low-grade meningiomas. All patients were randomly divided into training set, validation set and test set in a 6:2:2 ratio. The training set consisted of 194 patients with meningiomas (170 low-grade, 24 high-grade), the validation set consisted of 66 patients with meningiomas (56 low-grade, 10 high-grade) and the test set consisted of 66 patients with meningiomas (57 low-grade, 9 high-grade). There was no significant difference in age, gender and WHO grade among different sets (both *P* values > 0.01). The independent test group was not used for feature selection and hyperparameter tuning. The baseline characteristics of all patients are listed in Table 1.

Segmentation performance

The DLM detected meningiomas in all cases. For automatic segmentation, the mean dice coefficient and 95% Hausdorff distance were 0.951 and 0.953 for the training set, respectively; 0.866 and 3.139 for the validation set, respectively; 0.881 and 2.016 for the test set, respectively. Boxplots for the Dice coefficient and 95HD of each automatic segmentation are shown in Fig. 4. Two representative cases for automatic segmentation were displayed in Fig. 5. In the

Table 1 The baseline characteristics of all patients

Characteristic	Training set (*n* = 194)	Validation set (*n* = 66)	Test set (*n* = 66)	All patients (*n* = 326)	*P* value
Age, mean ± std	55.4 ± 9.7	53.9 ± 10.0	55.2 ± 9.5	55.0 ± 9.8	> 0.01
Gender, (%)					> 0.01
Male	59 (30.4%)	17 (25.8%)	17 (25.8%)	93 (28.5%)	
Female	135 (69.6%)	49 (74.2%)	49 (74.2%)	233 (71.5%)	
WHO grade, (%)					> 0.01
Low-grade	170 (87.6%)	56 (84.8%)	57 (84.4%)	283 (86.8%)	
High-grade	24 (12.4%)	10 (15.2%)	9 (13.6%)	43 (13.2%)	

first case, the meningioma was perfectly segmented, while the segmentation result in the second case was not ideal. Overall, the developed deep learning–based segmentation method enables automatic and accurate extraction of meningiomas in the vast majority of cases. Features extracted on manual and automatic segmentation are comparable: the average ICC value was 0.804 (range, 0.636–0.933), and the corresponding results are listed in Table 2. All ICC values were listed in Supplementary material 1.

Radiomics performance

In total, 1496 radiomics features demonstrated both inter-ICC and intra-ICC values exceeding 0.80 (listed in Supplementary material 2). Among these features, ten features were finally selected to construct a radiomics model for grading meningiomas. The detailed name of the selected features is shown in Table 2. For meningioma classification, the radiomics model based on the radiomics features extracted from automatic segmentation constructed performed well in grading meningiomas, yielding a sensitivity, specificity, accuracy, and AUC of 0.824 (95% CI: 0.738–0.910), 0.898 (95% CI: 0.760–0.944), 0.888 (95% CI: 0.753–0.921), and 0.930 (95% CI: 0.896–0.952) in the training set, respectively, while these indexes were 0.778 (95% CI: 0.701–0.856), 0.860 (95% CI: 0.722–0.908), 0.848 (95% CI: 0.715–0.903) and 0.842 (95%

CI: 0.807–0.895) in the test set, respectively. The radiomics model based on the radiomics features extracted from manual segmentation had a sensitivity, specificity, accuracy, and AUC of 0.941 (95% CI: 0.845–0.969), 0.872 (95% CI: 0.755–0.939), 0.881 (95% CI: 0.750–0.913), and 0.961 (95% CI: 0.904–0.969) in the training set, respectively, while they were 0.778 (95% CI: 0.718–0.859), 0.842 (95% CI: 0.716–0.904), 0.833 (95% CI: 0.709–0.898) and 0.813 (95% CI: 0.799–0.862) in the test set, respectively. The performance of meningioma differentiation on the training and test sets was listed in Table 3, and the corresponding ROC curves of the radiomics models based on automatic segmentation and manual segmentation for meningioma grading were shown in Fig. 6A and B, respectively. The Delong test of radiomics model based on automatic segmentation with radiomics model based on manual segmentation detected no significant differences (*P* = 0.65). Calibration curves (Fig. 6C, D) showed that the predicted probabilities of the radiomics models based on automatic segmentation and manual segmentation were closely aligned with the actual clinical observation in the training set. The agreement between the prediction results of the radiomics models and the actual clinical findings was moderate. The decision curve demonstrated that the results predicted by our radiomics models exhibited favorable clinical usefulness.
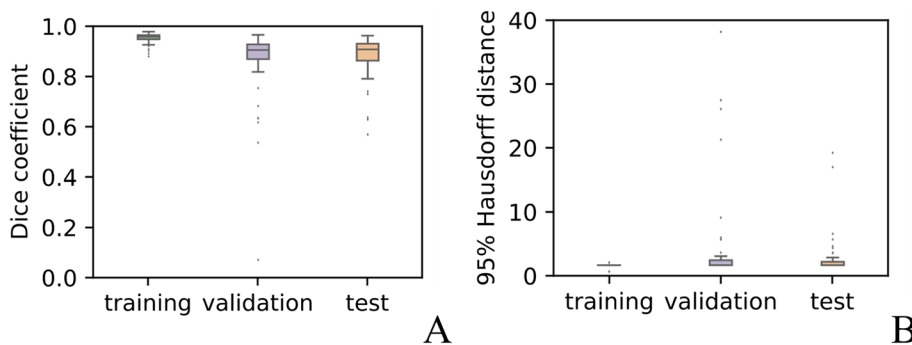

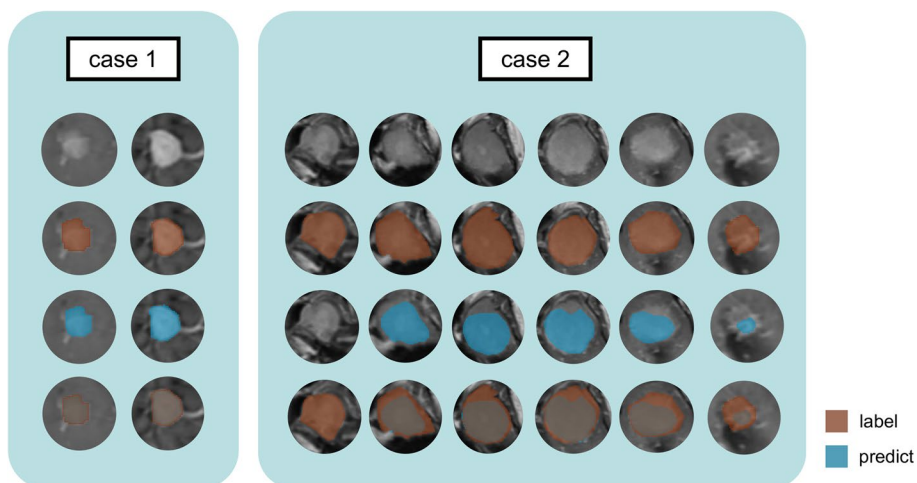**Fig. 4** Boxplots for Dice coefficient and 95HD of each automatic segmentation

Yang *et al. BMC Medical Imaging*   (2024) 24:56

Page 7 of 12

Fig. 5 Two representative cases for DLM-based automatic segmentation

Table 2 Corresponding ICCs of ten selected features between automatic segmentation and manual segmentation

Class	Feature	ICC
Shape	original_shape_Maximum2DDiameterSlice	0.933
	original_shape_MinorAxisLength	0.928
GLCM	wavelet-LLH_glcm_JointAverage	0.930
	wavelet-LLH_glcm_JointEnergy	0.838
	wavelet-LHL_glcm_ClusterTendency	0.910
GLSZM	wavelet-LHH_glszm_ZoneEntropy	0.665
	wavelet-HLL_glszm_ZoneEntropy	0.636
	wavelet-HHL_glszm_GrayLevelNonUniformity	0.703
	wavelet-LLL_glszm_SmallAreaHighGrayLevelEmphasis	0.694
	wavelet-LLL_glszm_HighGrayLevelZoneEmphasis	0.803

Discussion

The selection of therapeutic strategy and survival prognosis are extremely different for low- and high-grade meningiomas due to their distinct biological characteristics. Given this clinical context, an accurate prediction of meningioma grade has crucial role in guiding treatment decisions. In the present study, we developed a DL-based model for automatic meningioma segmentation in the radiomic analysis of meningioma grading, yielding an AUC of 0.930 and 0.842, respectively, in the training set and test set. Features extracted on manual and automatic segmentation are comparable, and there was a large overlap between these selected radiomics features for the manual and automatic segmentation approaches in the radiomic analysis of meningioma grading.

In the past decades, with the rapid development of ML and computer vision technology, medical image segmentation and classification have made great progress. So far, methods used to segment brain tumors mainly include conventional imaging algorithms, ML-based approaches, and techniques using DL networks. Currently, several (semi-)automatic methods for meningioma segmentation based on brain MR images have been developed. In a study conducted by Hsieh et al., meningiomas were segmented from non-contrast-enhanced MRI images using fuzzy clustering and the region-growing method [25]. Tsai et al. utilized the estimation maximization clustering technique to infer the location of meningiomas and peri-tumor oedemas from T2 axial images

Table 3 Performance of meningioma differentiation on the training and test sets

Dataset	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)	AUC (95% CI)
Training set (automatic)	0.824 (0.738–0.910)	0.898 (0.760–0.944)	0.888 (0.753–0.921)	0.930 (0.896–0.952)
Test set (automatic)	0.778 (0.701–0.856)	0.860 (0.722–0.908)	0.848 (0.715–0.903)	0.842 (0.807–0.895)
Training set (manual)	0.941 (0.845–0.969)	0.872 (0.755–0.939)	0.881 (0.750–0.913)	0.961 (0.904–0.969)
Test set (manual)	0.778 (0.718–0.859)	0.842 (0.716–0.904)	0.833 (0.709–0.898)	0.813 (0.799–0.862)

Yang *et al. BMC Medical Imaging*   (2024) 24:56

Page 8 of 12

[26]. The above traditional approaches yielded moderate efficacy, and their accuracy cannot meet the requirements of clinical routine for meningioma segmentation. To improve performance and to develop an automated segmentation model, DL procedures have been used resulting in significant improvements, notably with the recent developments of convolutional neural networks (CNNs) [27–29]. To perform a pixel-wise segmentation, DL procedures can learn rich features and present a computation advantage over atlas-based strategies [30]. For instance, previous studies have demonstrated that using a DL-based approach to meningiomas greatly improved the segmentation accuracy compared with the traditional technique [31, 32]. Thus, the DL technique could have a big influence on encouraging outcomes on target lesion segmentation and classification, which enables an image-based diagnosis to be more automated. To the best of our knowledge, relatively few studies have been conducted regarding the fully automated detection and segmentation of meningiomas to date. In this study, a DLM was built for automatic segmentation of meningioma based on MRI images, the automatic segmentation model enabled accurate extraction of meningiomas and generate radiomics features that are highly consistent with those obtained using manual segmentation [33]. Furthermore, the radiomics model constructed on features from automatic segmentation can assist in accurately differentiating high-grade meningiomas from low-grade ones in clinical practice.

Compared with other routine sequences in clinical practice, T1W-CE images could show more vivid lesions and boundaries of meningiomas due to abundant blood supply, which facilitated the segment of the tumor to some degree [34]. Laukamp et al. built a dedicated meningioma DLM based on T1W-CE data and evaluated its performance for automated tumor segmentation. Of the 56 meningiomas in the validation group, 55 were detected by the DLM. In these patients, the comparison of the DLM and manual segmentations revealed average dice coefficients of 0.82 ± 0.12 for total lesion volume [31]. In the study by Kang et al., the Sorensen-Dice similarity coefficients of the U-Net for small meningiomas less than 1 cm3 were 0.769 and 0.780 with the internal validation set and external validation set, respectively

[35]. We guess that imperfect automation and disappointing performance for small meningiomas of previous automated tools limit their use in routine clinical practice. Although our results demonstrated clinically applicable performance for meningiomas segmentation using SegResNet developed, our prediction performance still needs to be further improved. In the present study, we considered that several challenges still exist in automatic meningiomas segmentation algorithms in the following aspects: First of all, anatomical variations and different MRI equipment resulted in varying imaging data and inconsistent scanning parameters. In addition, the imaging manifestation of different grades and subtypes of meningiomas varies greatly. Finally, large amounts of training data are needed for deep convolutional neural networks to extract complex feature hierarchies through self-learning capabilities because that DLM must work with multiple processing layers and abstraction levels.

Recently, radiomics models based on MRI images have been developed for grading meningiomas in previous works. A meta-analysis summarized eight related radiomics studies, where the pooled AUC of studies employed a test group achieved 0.84 (95% CI: 0.78–0.90), suggesting that radiomics could serve as an effective tool in grading meningiomas [17]. In the present study, the radiomics model constructed on features from automatic segmentation exhibited favorable performance in meningiomas grading on the testing set, yielding a sensitivity, specificity, accuracy, and AUC of 0.778, 0.860, 0.848, and 0.842 in the test set, respectively. The achieved good performance and high reliability of the test set in the present study demonstrated the potential of applying radiomics to assist in accurately differentiating high-grade meningiomas from low-grade ones in clinical practice. The calibration curves for the test set are not good enough, possibly due to limited sample size and imbalanced data. Enlarged datasets are needed to further test the generalizability and clinical usefulness of the constructed ML pipeline. Of note, Verma and his colleagues firstly created a radiomics risk score using radiomic features obtained from different sub-compartments of the tumor habitat (enhancing tumor, peritumoral edema and non-enhancing tumor region, and tumor

(See figure on next page.)

Fig. 6 ROC and calibration curves of the radiomics model in the training and test sets. **A** ROC curve of radiomics model based on the radiomics features extracted from automatic segmentation; **B** ROC curve of radiomics model based on the radiomics features extracted from manual segmentation; **C** Calibration curve of radiomics model based on the radiomics features extracted from automatic segmentation; **D** Calibration curve of radiomics model based on the radiomics features extracted from manual segmentation. **E** Decision curve analysis of radiomics model based on the radiomics features extracted from automatic segmentation; **F** Decision curve analysis of radiomics model based on the radiomics features extracted from manual segmentation
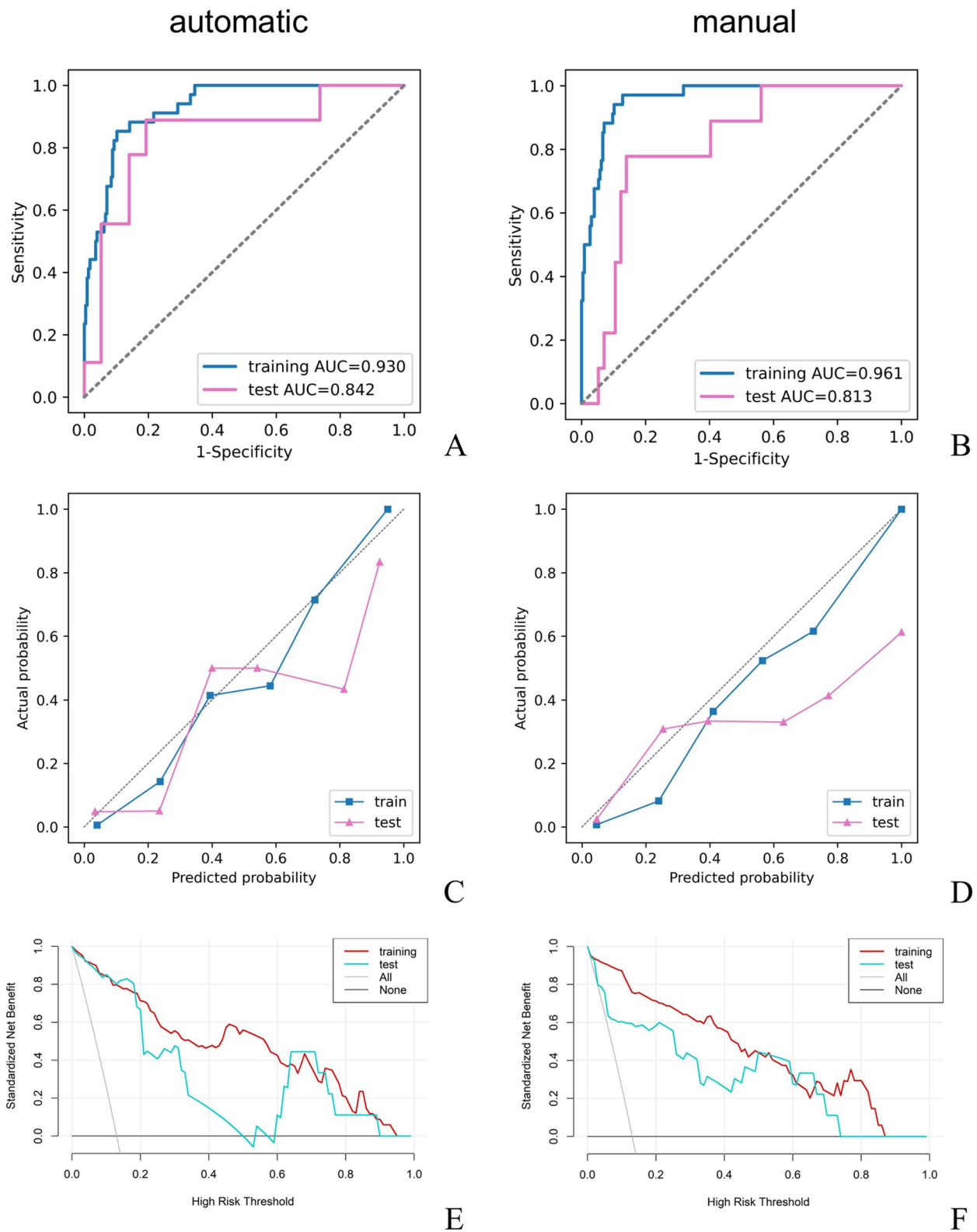
automatic   **manual**

Fig. 6 (See legend on previous page.)

necrosis regions) to classify patients as low and high risk for poor progression-free survival in response to treatment. Preliminary findings revealed significant associations of prognostic radiomics features with disease-specific histologic attributes [36]. In our analysis, we mainly emphasized the importance of radiomics features derived from the tumor itself, but overlooked the morphologic associations of radiomics features obtained from peritumoral edema with the underlying pathophysiologic processes that drive tumor behavior. In subsequent analysis, we will incorporate radiomics features obtained from peritumoral edema to alleviate this impact.

Among the retained radiomics features, the first-order features describe the distribution of voxel intensities in images. The GLCM features quantify the second-order joint probabilities of images which quantifies the intensity distribution of the gray level at a given offset to extract information about tone homogeneity, linear connection, contrast, and boundaries adjacent to gray zones, as well as complicacy of distribution [37]. The GLSZM features describe gray-level runs in an image. Skewness, as one of the simple parameters, represents the asymmetric distribution of gray levels in the histogram that describes the heterogeneity of lesions [38]. The above features describe the patterns or spatial distribution of voxel intensities within the ROI, which serve as recognized parameters to capture tumor heterogeneity [39]. Indirectly, our findings confirmed that the selected features were all closely related to high-dimensional space information that can hardly be understood by naked-eye examination, which may potentially assist in the differential diagnosis.

Likewise, our study still had several limitations. Firstly, potential selection bias might exist because of the retrospective nature, and the conduct of prospective studies may alleviate this impact; Secondly, the VOIs were delineated along the boundary of meningiomas, suggesting that it might exclude the potential information of peritumoral edema; Thirdly, our data are not collected using scanners from one the same vendor; Fourth, the radiomics model is solely established on the axial T1W-CE images, other sequences, such as T2WI, FLAIR, DWI, ADC images, should be included in subsequent analysis; Finally, thorough validation of this radiomics model will ultimately require application to an external, multi-institutional dataset with a larger cohort of patients. We will add an external validation to provide more sufficient evidence for clinical application in the near future.

Conclusions

In summary, the developed SegResNet-based segmentation model allowed for automatic and accurate meningioma segmentation from T1W-CE MRI images. The DLM with automatic segmentation demonstrated performance comparable to that of the model with manual segmentation. With respect to meningioma differentiation, our automatic segmentation approach will likely enable the efficient implementation of radiomics for grading meningiomas before surgery and facilitate its clinical application.

Abbreviations

WHO	World Health Organization
DL	Deep Learning
ICC	Intra Class Correlation
DLM	Deep Learning Model
WHO	World Health Organization
MRI	Magnetic Resonance Imaging
T1W-CE	T1-weighted contrast-enhanced
ML	Machine Learning
VOIs	Volumetric Regions of Interest
ANOVA	One-Way Variance of Analysis
95HD	95% Hausdorff Distance
ROC	Receiver Operating Characteristics
SMOTE	Synthetic Minority Oversampling Technique
AUC	Area Under the Curve

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12880-024-01218-3.

> **Additional file 1: Supplementary Material 1.** All ICCs between features extracted on manual and automatic Segmentation.
>
> **Additional file 2: Supplementary Material 2.** All radiomics features with both inter-ICC and intra-ICC values exceeding 0.80.

Availability of data and materials
The datasets generated during and analyzed during the current study are not publicly available due to patient privacy concerns but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Institutional Review Board and Ethical Committee of The Harbin Medical University Cancer Hospital. The need for informed consent was waived by the ethics committee of The Harbin Medical University Cancer Hospital, because of the retrospective nature of the study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of PET-CT, Harbin Medical University Cancer Hospital, Harbin 150001, China. [2]Medical Imaging Department, Changzheng Hospital of Harbin City, Harbin, China. [3]Medical Imaging Department, The Second Affiliated Hospital of Harbin Medical University, Harbin, China. [4]Medical Imaging Department, The Second Hospital of Heilongjiang Province, Harbin, China. [5]Department of Medical Instruments, Second Hospital of Harbin, Harbin 150001, China.

References

1. Wiemels J, Wrensch M, Claus EB. Epidemiology and etiology of meningioma. J Neurooncol. 2010;99:307–14.
2. Alexandru D, Glantz MJ, Kim L, Chamberlain MC, Bota DA. Pulmonary metastases in patients with recurrent, treatment-resistant meningioma: prognosis and identification by [111]Indium-octreotide imaging. Cancer. 2011;117:4506–11.
3. Spille DC, Adeli A, Sporns PB, Heß K, Streckert EMS, Brokinkel C, et al. Predicting the risk of postoperative recurrence and high-grade histology in patients with intracranial meningiomas using routine preoperative MRI. Neurosurg Rev. 2021;44:1109–17.
4. Lam Shin Cheung V, Kim A, Sahgal A, Das S. Meningioma recurrence rates following treatment: a systematic analysis. J Neurooncol. 2018;136:351–61.
5. Biczok A, Jungk C, Egensperger R, von Deimling A, Suchorska B, Tonn C, et al. Microscopic brain invasion in meningiomas previously classifiedas WHO grade I is not associated with patient outcome. J Neurooncol. 2019;145:469–77.
6. Goldbrunner R, Minniti G, Preusser M, Jenkinson MD, Sallabanda K, Houdart E, et al. EANO guidelines for the diagnosis and treatment of meningiomas. Lancet Oncol. 2016;17:e383–91.
7. Willis J, Smith C, Ironside JW, Erridge S, Whittle IR, Everington D. The accuracy of meningioma grading: a 10-year retrospective audit. Neuropathol Appl Neurobiol. 2005;31:141–9.
8. Bertero L, Dalla Dea G, Osella-Abate S, Botta C, Castellano I, Morra I, et al. Prognostic Characterization of Higher-Grade Meningiomas: A Histopathological Score to Predict Progression and Outcome. J Neuropathol Exp Neurol. 2019;78:248–56.
9. Ly KI, Hamilton SR, Rostomily RC, Rockhill JK, Mrugala MM. Improvement in Visual Fields After Treatment of Intracranial Meningioma With Bevacizumab. J Neuroophthalmol. 2015;35:382–6.
10. Fiani B, Jarrah R, Shields J, Durrani S, Panico N, Mualem W, et al. A Comprehensive Overview of Pediatric Neoplasms at the Craniocervical Junction: Meningiomas, Schwannomas, and Chordomas. Cureus. 2022;14: e31083.
11. Zhu Y, Man C, Gong L, Dong D, Yu X, Wang S, et al. A deep learning radiomics model for preoperative grading in meningioma. Eur J Radiol. 2019;116:128–34.
12. Huang RY, Bi WL, Griffith B, Kaufmann TJ, la Fougère C, Schmidt NO, et al. International Consortium on Meningiomas Imaging and diagnostic advances for intracranial meningiomas. Neuro Oncol. 2019;21:i44–61.
13. Spille DC, Sporns PB, Heß K, Stummer W, Brokinkel B. Prediction of High-Grade Histology and Recurrence in Meningiomas Using Routine Preoperative Magnetic Resonance Imaging: A Systematic Review. World Neurosurg. 2019;128:174–81.
14. Ke C, Chen H, Lv X, Li H, Zhang Y, Chen M, et al. Differentiation Between Benign and Nonbenign Meningiomas by Using Texture Analysis From Multiparametric MRI. J Magn Reson Imaging. 2020;51:1810–20.
15. Radakovich N, Nagy M, Nazha A. Machine learning in haematological malignancies. Lancet Haematol. 2020;7:e541–50.
16. Martín Noguerol T, Paulano-Godino F, Martín-Valdivia MT, Menias CO, Luna A. Strengths, Weaknesses Opportunities, and Threats Analysis of Artificial Intelligence and Machine Learning Applications in Radiology. J Am Coll Radiol. 2019;16:1239–47.
17. Ugga L, Perillo T, Cuocolo R, Stanzione A, Romeo V, Green R, et al. Meningioma MRI radiomics and machine learning: systematic review, quality score assessment, and meta-analysis. Neuroradiology. 2021;63:1293–304.
18. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classificationof Tumors of the Central Nervous System: a summary. Acta Neuropathol. 2016;131:803–20.
19. Novac OC, Chirodea MC, Novac CM, Bizon N, Oproescu M, Stan OP, et al. Analysis of the Application Efficiency of TensorFlow and PyTorch in Convolutional Neural Network. Sensors (Basel). 2022;22:8872.
20. Fernández Patón M, Cerdá Alberich L, Sangüesa Nebot C, de Las Martínez Heras B, Veiga Canuto D, Cañete Nieto A, et al. MR Denoising Increases Radiomic Biomarker Precision and Reproducibility in Oncologic Imaging. J Digit Imaging. 2021;34:1134–45.
21. Wang J. Prediction of postoperative recovery in patients with acoustic neuroma using machine learning and SMOTE-ENN techniques. Math Biosci Eng. 2022;19:10407–23.
22. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inform Decis Mak. 2019;19:281.
23. Le Glaz A, Haralambous Y, Kim-Dufor DH, Lenca P, Billot R, Ryan TC, et al. Machine Learning and Natural Language Processing in Mental Health: Systematic Review. J Med Internet Res. 2021;23: e15708.
24. Nakamura M, Kajiwara Y, Otsuka A, Kimura H. LVQ-SMOTE- Learning Vector Quantization based Synthetic Minority Over-sampling Technique for biomedical data. BioData Min. 2013;6:16.
25. Hsieh TM, Liu YM, Liao CC, Xiao F, Chiang IJ, Wong JM. Automatic segmentation of meningioma from non-contrasted brain MRI integrating fuzzy clustering and region growing. BMC Med Inform Decis Mak. 2011;11:54.
26. Tsai YF, Chiang IJ, Lee YC, Liao CC, Wang KL. Automatic MRI meningioma segmentation using estimation maximization. Conf Proc IEEE Eng Med Biol Soc. 2005;2005:3074–7.
27. Currie G, Hawk KE, Rohren E, Vial A, Klein R. Machine Learning and Deep Learning in Medical Imaging: Intelligent Imaging. J Med Imaging Radiat Sci. 2019;50:477–87.
28. Rezaeijo SM, Chegeni N, Baghaei Naeini F, Makris D, Bakas S. Within-Modality Synthesis and Novel Radiomic Evaluation of Brain MRI Scans. Cancers (Basel). 2023;15:3565.
29. Bouget D, Pedersen A, Hosainey SAM, Solheim O, Reinertsen I. Meningioma Segmentation in T1-Weighted MRI Leveraging Global Context and Attention Mechanisms. Front Radiol. 2021;1: 711514.
30. Neromyliotis E, Kalamatianos T, Paschalis A, Komaitis S, Fountas KN, Kapsalaki EZ, et al. Machine Learning in Meningioma MRI: Past to Present A Narrative Review. J Magn Reson Imaging. 2022;55:48–60.
31. Laukamp KR, Thiele F, Shakirin G, Zopfs D, Faymonville A, Timmer M, et al. Fully automated detection and segmentation of meningiomas using deep learning on routine multiparametric MRI. Eur Radiol. 2019;29:124–32.
32. Zhang H, Mo J, Jiang H, Li Z, Hu W, Zhang C, et al. Deep Learning Model for the Automated Detection and Histopathological Prediction of Meningioma. Neuroinformatics. 2021;19:393–402.
33. Scalco E, Rizzo G, Mastropietro A. The stability of oncologic MRI radiomic features and the potential role of deep learning: a review. Phys Med Biol. 2022;19;67.
34. Chen H, Li S, Zhang Y, Liu L, Lv X, Yi Y, et al. Deep learning-based automatic segmentation of meningioma from multiparametric MRI for preoperative meningioma differentiation using radiomic features: a multicentre study. Eur Radiol. 2022;32:7248–59.

35. Kang H, Witanto JN, Pratama K, Lee D, Choi KS, Choi SH, et al. Fully Automated MRI Segmentation and Volumetric Measurement of Intracranial Meningioma Using Deep Learning. J Magn Reson Imaging. 2023;57:871–81.

36. Verma R, Correa R, Hill VB, Statsevych V, Bera K, Beig N, et al. Tumor Habitat-derived Radiomic Features at Pretreatment MRI That Are Prognostic for Progression-free Survival in Glioblastoma Are Associated with Key Morphologic Attributes at Histopathologic Examination: A Feasibility Study. Radiol Artif Intell. 2020;2: e190168.

37. Li X, Guindani M, Ng CS, Hobbs BP. Spatial Bayesian modeling of GLCM with application to malignant lesion characterization. J Appl Stat. 2018;46:230–46.

38. Defeudis A, Mattia CD, Rizzetto F, Calderoni F, Mazzetti S, Torresin A, et al. Standardization of CT radiomics features for multi-center analysis: impact of software settings and parameters. Phys Med Biol. 2020;65:195012–195010pp.

39. Reis SP, Sutphin PD, Singal AG, Grzybowski R, Fisher S, Ball C, et al. Tumor Enhancement and Heterogeneity Are Associated With Treatment Response to Drug-Eluting Bead Chemoembolization for Hepatocellular Carcinoma. J Comput Assist Tomogr. 2017;41:289–93.

Publisher's Note