© Pol J Radiol 2026; 91: e495-e505
DOI: https://doi.org/10.5114/pjr/219348

Received: 22.01.2026
Accepted: 15.03.2026
Published: 19.10.2026

Polish Journal of Radiology

http://www.polradiol.com

Original paper

Improved classification of benign and malignant hepatic tumors using combined ADC and T2-weighted radiomic features

Hong Phuong Dung Tran[A,B,C,D,E,F,#], Ngoc Thanh Hoang[A,C,D,E,F,#], Trong Binh Le[D,F], Dac Hong An Ngo[C,D], Thanh Thao Nguyen[E,F,G]

Department of Radiology, University of Medicine and Pharmacy, Hue University, Hue, Vietnam

[#]These authors contributed equally to this study.

Abstract

Purpose: To utilize machine learning algorithms to differentiate between malignant and benign hepatic lesions by integrating T2-weighted MRI-based radiomics and apparent diffusion coefficient (ADC) metrics.

Material and methods: This cross-sectional study (January 2023–December 2024) enrolled 111 patients (71 male, 40 female) aged ≥ 18 years. Hepatic tumors were classified via radiological, pathological, and biochemical assessment. Radiomic features were derived from T2-weighted images, and ADC metrics from diffusion-weighted imaging. The classification performance was evaluated using three models: Model 1 (ADC and ADC Ratios), Model 2 (Radiomic Features), and Model 3 (Combined Features). Stratified 5-fold cross-validation was performed across four algorithms: Logistic Regression (LR), Support Vector Machine (SVM), Gradient Boosting (GB), and Random Forest. The area under the receiver operating characteristic curve (AUC) was used to measure the models' discriminatory ability. All statistical analyses were conducted using Python libraries.

Results: A total of 111 hepatic lesions were analyzed: 75 malignant and 36 benign. Among 12 significant radiomic features ($p < 0.05$), firstorder_InterquartileRange, ngtdm_Contrast, and glszm_ZonePercentage showed the largest effect sizes (Cohen's $d > 0.70$). All ADC-related metrics were significantly higher in benign lesions compared to malignant ones ($p < 0.001$). Model 3 demonstrated the strongest overall performance, with both LR and SVM achieving accuracy and AUC values above 0.800. While Model 1 with SVM offered the highest sensitivity, Model 2 displayed the lowest specificity and weakest overall performance, whereas Model 3 generally showed higher classification performance than either feature set alone.

Conclusions: Combining T2w-based radiomics features with ADC and ADC ratio values may improve discrimination between malignant and benign liver tumors.

Key words: liver tumors, ADC, radiomics, classification, T2w-based.

Introduction

In recent years, the occurrence of focal hepatic lesions (FLLs) has steadily increased. Among benign liver lesions, hepatic cysts are the most common, followed by hemangiomas. Hepatocellular carcinoma (HCC) accounts for approximately 90% of primary malignant liver tumors, while metastatic deposits from extrahepatic primaries are nearly 20 times more common and usually appear as multifocal lesions [1]. Primary liver cancer is the sixth most common cancer worldwide and the third leading cause of cancer-related deaths, with men experiencing much higher rates than women, especially in Eastern Asia and parts of Africa. Most cases are HCC, mainly linked to chronic hepatitis B or C infection, aflatoxin exposure, excessive alcohol consumption, obesity, diabetes, and smoking. Precise and non-invasive characterization of these lesions is crucial for prognosis and guiding proper treatment, which can

Correspondence address:
Thanh Thao Nguyen, Department of Radiology, University of Medicine and Pharmacy, Hue University, 6 Ngo Quyen, Hue City 49135, Vietnam, e-mail: ntthao@hueuni.edu.vn

Authors' contribution:
A Study design · **B** Data collection · **C** Statistical analysis · **D** Data interpretation · **E** Manuscript preparation · **F** Literature search · **G** Funds collection

This is an Open Access journal, all articles are distributed under the terms of the Creative Commons Attribution-Noncommercial-No Derivatives 4.0 International (CC BY-NC-ND 4.0). License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

range from conservative monitoring to surgical resection or systemic therapy [2]. Although imaging techniques are central to evaluation, some lesions remain indeterminate, requiring additional diagnostic measures that can reduce patient anxiety.

Magnetic resonance imaging (MRI), especially when using gadolinium-based contrast agents, is considered the gold standard for non-invasive evaluation of FLLs. However, its diagnostic accuracy is not flawless, as some lesions show atypical enhancement features that can resemble both benign and malignant conditions [3,4]. Additionally, limitations such as contraindications to contrast use in patients with impaired kidney function, along with the risks and costs associated with invasive histopathological confirmation, highlight the urgent need for more reliable and quantitative imaging biomarkers.

Quantitative MRI techniques have emerged to address current limitations in the field. One such method, diffusion-weighted imaging (DWI), evaluates the microscopic movement of water molecules within tissues [5]. This analysis enables the calculation of the apparent diffusion coefficient (ADC), which offers important information about tissue cellularity and membrane integrity [5]. ADC values have an inverse relationship with cellular density; typically, malignant lesions show diffusion restriction and, as a result, have lower ADC values compared to benign lesions or normal liver tissue. Radiomics has become a key approach for high-throughput data analysis in medical imaging [6-8]. It focuses on extracting quantitative and reproducible information from diagnostic images. Its goal is to capture complex patterns that are difficult for the human eye to see or measure accurately. This technique systematically converts medical images into a wide range of quantitative features to describe the imaging phenotype [9,10]. In this way, radiomics acts as a non-invasive method to explore the relationship between these quantitative features and biological processes. While both ADC metrics and radiomic features help distinguish FLLs independently, combining functional data from DWI with radiomic-based morphological and textural features may produce a more comprehensive and reliable diagnostic profile.

In recent years, numerous studies have evaluated radiomic features for classifying liver lesions using different imaging protocols [11-13]. Specifically, T2-weighted (T2w) and ADC radiomic features have shown potential in distinguishing liver tumors [7]. However, although the current literature includes only a few studies assessing T2w radiomics alongside ADC diffusion characteristics, a systematic evaluation of their combined potential within a solid machine learning framework is still missing.

The present study aimed to develop and validate machine learning models to distinguish malignant from benign hepatic lesions by combining T2w radiomic features with ADC metrics, offering a more accurate and non-invasive tool to assist clinical decision-making.

Material and methods

Study population

Following institutional review board approval (01/22) and informed consent, 111 patients with FLLs were enrolled in a cross-sectional analysis (January 2023–March 2024) and underwent abdominal MRI at a single center. This dataset was previously reported for ADC-based classification only [14]; however, 4 patients were excluded here due to unsatisfactory T2w image quality. In the present study, we combined these metrics with radiomic features extracted from T2w images, hypothesizing that integrating diffusion and radiomic information would improve classification performance.

Inclusion criteria were: (1) patients aged 18 years or older; (2) presence of at least one solid focal liver lesion with a diameter greater than 10 mm identified on MRI; and (3) a definitive final diagnosis established either by histopathological analysis or by characteristic imaging features combined with at least 6 months of clinical and imaging follow-up. Exclusion criteria included: (1) unsatisfactory image quality; (2) patients with a history of ongoing chemotherapy, prior liver resection, or splenectomy; (3) contraindications to MRI; and (4) the presence of simple hepatic cysts.

Diagnostic criteria

Using imaging criteria along with biochemical and/or histopathological assessments, target lesions were classified into seven distinct types. These were then grouped into two main categories: benign lesions (hemangioma, focal nodular hyperplasia [FNH], adenoma, and focal inflammatory lesions) and malignant lesions (HCC, intrahepatic cholangiocarcinoma [ICC], liver metastases, and hepatoblastoma).

HCC was diagnosed following the 2018 clinical practice guidelines of the European Association for the Study of the Liver (EASL), while ICC or hepatoblastoma was confirmed through histopathological examination [15]. Liver metastases were identified based on characteristic imaging features in patients with a known primary malignancy, with or without histology. On MRI, metastases typically appear hypointense on non-contrast T1-weighted (T1w) images and mildly to moderately hyperintense on T2w sequences, often exhibiting a targetoid appearance on DWI. Post-contrast enhancement patterns depend on lesion vascularity: hypovascular metastases generally show persistent hypoenhancement, whereas hypervascular metastases often demonstrate peripheral rim enhancement in the arterial phase. This peripheral rim usually corresponds to viable tumor tissue, while central non-enhancement indicates necrosis or cystic degeneration, reflecting advanced metastatic changes [4,16]. Benign lesions were primarily confirmed based on radiological criteria following EASL guidelines [17],

or deemed benign if stable in size over at least six months of follow-up imaging; otherwise, percutaneous biopsy was performed for a definitive diagnosis.

Liver MRI protocol

MRI of the liver was performed using a 1.5 T system (Amira, Siemens, Erlangen, Germany). Axial DWI was acquired using a fat-suppressed single-shot echo-planar sequence with three *b*-values (50, 400, and 800 s/mm^2). Axial T2w images were obtained with a fast spin-echo sequence (detailed parameters are provided in Table 1). The institutional routine MRI protocol also included T2w fat-saturation single-shot fast spin-echo coronal and axial scans and gradient-recalled echo T1w in-phase and out-of-phase axial and coronal images acquired before and after gadolinium injections.

Imaging interpretation

DWI sequences were utilized to measure ADC values and calculate ADC ratios. T2w sequences were employed for radiomic feature extraction of hepatic lesions, as they offer superior intrinsic tissue contrast and more clearly define lesion margins.

Quantitative ADC metrics

Parametric MRI software (www.parametricmri.com, Philadelphia, PA, USA) was used for post-processing and analysis of images, and ADC was calculated based on linear and nonlinear exponential fits. Regions of interest (ROIs) were positioned for target lesions, background liver, spleen tissue, and paraspinal muscle, avoiding cystic or necrotic regions adjacent to vessels. For lesions with heterogeneous signals on DWI images, ROIs were placed in the brightest and darkest areas within the enhanced solid components. Conversely, for lesions with a homogeneous appearance, ROIs were placed at different axial slices within the lesion. ADC values at each location were recorded as the average from the ROIs. Hepatic ADC values were obtained from four ROIs in the lateral, medial, anterior, and posterior segments of the liver. For the spleen and bilateral paraspinal muscles, two ROIs were placed in each structure. The mean ADC value from all ROIs in each anatomical area was used for further analysis. All ROI measurements had a radius of 5-10 mm on DWI images. Four ADC-derived metrics were included in this study: ADC tumor (ADCt), Ratio1_liver (lesion-to-liver ADC), Ratio2_spleen (lesion-to-spleen ADC), and Ratio3_muscle (lesion-to-paraspinal muscle ADC).

Extracting T2w-based radiomic features

3D Slicer software (3D Slicer, www.slicer.org, Boston, MA, USA) was used for lesion segmentation, post-processing, and quantitative radiomic analysis. For each lesion, manual segmentation of the entire tumor volume was performed on the axial T2w images, slice by slice. The resulting three-dimensional ROIs were then used for quantitative radiomic analysis with the SlicerRadiomics extension (built on PyRadiomics), which allows extraction of first-order statistics, shape-based features, and higher-order textural features. First-order statistics quantify voxel intensity characteristics within a region of interest. Higher-order textural features in radiomics include GLCM, GLRLM, GLSZM, GLDM, and NGTDM features, which analyze complex spatial relationships and texture patterns within the image [18-20]. Original T2w-based radiomic features were used for the development of the classification model; their details and definitions are summarized in Table 2.

Table 1. T2-weighted (T2w) and diffusion-weighted imaging (DWI) acquisition parameters

Parameters	T2w	DWI
Repetition time (ms)	1200	7900
Echo time (ms)	85	70
Matrix	256	256
Voxel size	1.4 × 1.4 × 4.0	1.4 × 1.4 × 4.0
Field of view (FOV, mm)	380	380
Phase encoding direction	Anterior to posterior	Anterior to posterior
Slice	50	50
Slice thickness (mm)	4.0	4.0
Distance factor (%)	20	20
b-values	-	50, 400, 800 s/mm^2
Number of averages	1	*b* = 50 (1 average) *b* = 400 (3 averages) *b* = 800 (5 averages)
Concatenations	4	1
Fat suppression technique	-	SPAIR (spectral attenuated inversion recovery)
Parallel imaging technique	Off	GRAPPA (acceleration factor = 2)
Scan time (minutes)	1.20	4

GRAPPA – GeneRalized Autocalibrating Partial Parallel Acquisition, SPAIR – spectral attenuated inversion recovery

Feature selection and classification

The analytical methodology is designed to evaluate three different classification models using a combination of diffusion features (ADC and ratios) and T2W-based radiomics features, focusing on model stability and discriminative ability. All quantitative features are standardized with the standard scaler method. Feature selection was carried out using recursive feature elimi-

Table 2. Description of original radiomic features

Class	Features	Description	Numbers
First Order (First Order Gray Level)	original_firstorder_10Percentile, original_firstorder_90Percentile, original_firstorder_Energy, original_firstorder_Entropy, InterquartileRange, Kurtosis, Maximum, MeanAbsoluteDeviation, Mean, Median, Minimum, Range, RobustMeanAbsoluteDeviation, RootMeanSquared, Skewness, TotalEnergy, Uniformity, Variance	Describes the distribution of gray level intensity within the region of interest (ROI), disregarding spatial relationships.	18
GLCM (Gray Level Co-occurrence Matrix)	Autocorrelation, ClusterProminence, ClusterShade, ClusterTendency, Contrast, Correlation, DifferenceAverage, DifferenceEntropy, DifferenceVariance, Id, Idm, Idmn, Idn, Imc1, Imc2, InverseVariance, JointAverage, JointEnergy, JointEntropy, MCC, MaximumProbability, SumAverage, SumEntropy, SumSquares	Describes the statistical relationship regarding the simultaneous appearance of two different gray levels at neighboring positions.	24
GLDM (Gray Level Dependence Matrix)	DependenceEntropy, DependenceNonUniformity, DependenceNonUniformityNormalized, DependenceVariance, GrayLevelNonUniformity, GrayLevelVariance, HighGrayLevelEmphasis, LargeDependenceEmphasis, LargeDependenceHighGrayLevelEmphasis, LargeDependenceLowGrayLevelEmphasis, LowGrayLevelEmphasis, SmallDependenceEmphasis, SmallDependenceHighGrayLevelEmphasis, SmallDependenceLowGrayLevelEmphasis	Describes the distribution of regions with uniform gray levels based on the region's gray level.	14
GLRLM (Gray Level Run Length Matrix)	GrayLevelNonUniformity, GrayLevelNonUniformityNormalized, GrayLevelVariance, HighGrayLevelRunEmphasis, LongRunEmphasis, LongRunHighGrayLevelEmphasis, LongRunLowGrayLevelEmphasis, LowGrayLevelRunEmphasis, RunEntropy, RunLengthNonUniformity, RunLengthNonUniformityNormalized, RunPercentage, RunVariance, ShortRunEmphasis, ShortRunHighGrayLevelEmphasis, ShortRunLowGrayLevelEmphasis	Describes the length of consecutive pixel runs that share the same gray level.	16
GLSZM (Gray Level Size Zone Matrix)	GrayLevelNonUniformity, GrayLevelNonUniformityNormalized, GrayLevelVariance, HighGrayLevelZoneEmphasis, LargeAreaEmphasis, LargeAreaHighGrayLevelEmphasis, LargeAreaLowGrayLevelEmphasis, LowGrayLevelZoneEmphasis, SizeZoneNonUniformity, SizeZoneNonUniformityNormalized, SmallAreaEmphasis, SmallAreaHighGrayLevelEmphasis, SmallAreaLowGrayLevelEmphasis, ZoneEntropy, ZonePercentage, ZoneVariance	Describes the distribution of sizes of uniform gray level zones.	16
NGTDM (Neighboring Gray Tone Difference Matrix)	Busyness, Coarseness, Complexity, Contrast, Strength	Quantifies the difference between a pixel's gray level and the average gray level of its neighboring pixels.	5

nation (RFE) to reduce dimensionality and identify the most informative radiomic features [19]. RFE sequentially removes the least important features based on their contribution to model performance. In this study, four common classical machine learning algorithms – Logistic Regression (LR), SVM Support Vector Machine (SVM), GB Gradient Boosting (GB), and Random Forest – were used as estimators within the RFE framework to rank feature importance. The selection threshold was set to keep the top 20 selected radiomic features with the highest discriminative power. These selected features were then used to develop Models 2 and 3 for the classification analysis.

- Model 1 (ADC and ADC Ratios): included only lesion ADC values and ADC ratios.
- Model 2 (Radiomic Features): included only the top 20 selected radiomic features.
- Model 3 (Combined Features): integrated both ADC metrics and the top 20 selected radiomic features.

Statistics analysis

Continuous data were presented as mean ± standard deviation if the variables were normally distributed. In this case, a *t*-test (Student's) was used to determine the difference between independent samples (benign vs. malignant). If the variables were not normally distributed, the data were shown as the median and interquartile range, and the Mann-Whitney *U* test was applied accordingly. Statistical significance was considered at $p < 0.05$.

After selecting the most important radiomic features using RFE, the effect size between the two groups (benign and malignant liver tumors) was measured with Cohen's *d*. The analysis aimed to evaluate how well the selected fea-

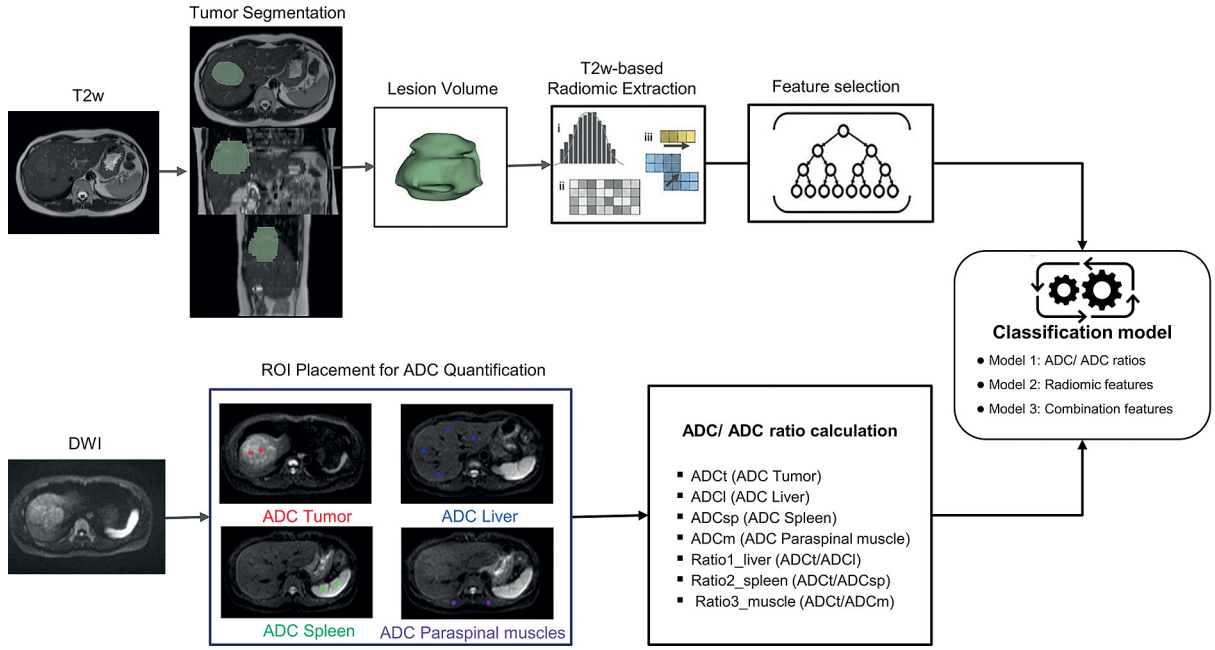


Figure 1. Study workflow illustrating apparent diffusion coefficient (ADC) and ADC ratio computation, radiomic feature extraction, feature selection, and machine-learning classification of benign and malignant lesions

tures could distinguish between the two groups. Cohen's *d* values were interpreted as follows: small ($d < 0.5$), medium ($0.5 \leq d < 0.8$), and large ($d \geq 0.8$), based on standard benchmarks. A bar plot was created to visualize the effect size distribution across the top 20 features.

To ensure a reliable estimate of generalizability, we employed a stratified 5-fold cross-validation (CV) approach. This method maintained the proportional distribution of benign and malignant lesions across folds while minimizing the risk of training and testing data leakage. Model performance was quantified by averaging accuracy, sensitivity (Se), and specificity (Sp) across five folds. Discriminatory ability was further assessed via receiver operating characteristic (ROC) analysis and the corresponding area under the curve (AUC). The AUC was used to assess discriminatory performance, which was categorized as excellent (≥ 0.900), good (0.810-0.900), fair (0.710-0.800), poor (0.610-0.700), and fail (0.510-0.600) [21]. All statistical analyses and machine learning procedures were conducted in Python (version 3.9). The research process is summarized in Figure 1.

Results

Patients' and tumor shape characteristics

A total of 111 patients were finally enrolled in this study, including 75 with malignant FLLs and 36 with benign FLLs, with median ages of 59.0 (53.0-68.5) and 45.0 (38.5-54.25), respectively. There were 71 male patients and 40 female patients. For shape metrics, the malignant group demonstrated significantly smaller least axis length ($p = 0.029$), lower sphericity ($p = 0.008$), and a reduced surface-to-volume ratio ($p = 0.005$) compared to the benign group. Other shape metrics showed no significant differences between the groups. Patient and tumor shape characteristics are summarized in Table 3.

Differences in ADC metrics among liver tumors

ADC values and ADC ratios were significantly higher in the benign group compared to the malignant group ($p < 0.001$), as summarized in Table 4.

Differences in T2w-derived radiomic features among liver tumors

Among the top 20 prominent features of T2w-based radiomics, we identified 12 features with statistically significant differences between the two groups, with $p < 0.05$ (Table 5). Notably, firstorder_InterquartileRange, ngtdm_Contrast, and glszm_ZonePercentage exhibited the strongest differences, with Cohen's *d* values of 0.88, 0.81 (large effect), and 0.70 (moderate effect), respectively ($p < 0.05$) (Figure 2).

Combined diffusion and radiomic features for liver tumor classification

The best overall performance was achieved by the LR model using Model 3, which showed the highest accuracy (0.846) and Sp (0.750), indicating good discrimination of negative cases. The SVM model trained on the same feature set achieved the highest AUC values (0.876 for mean CV and 0.873 for aggregated ROC), demonstrating the highest discriminatory performance among the tested

Table 3. Patient and shape tumor characteristics of focal liver lesions

Patient characteristics	Malignant	Benign	*p*-value
Sex			
Male	51 (68.0%)	20 (55.6%)	0.286
Female	24 (32.0%)	16 (44.4%)	0.286
Age	59.0 (53.0-68.5)	45.0 (38.5-54.25)	0.0
Shape feature			
Elongation	0.84 (0.12)	0.85 (0.15)	0.477
Flatness	0.66 (0.15)	0.64 (0.17)	0.740
Least axis length (mm)	47.2 (12.8)	14.2 (15.7)	0.029*
Major axis length (mm)	53.3 (13.8)	62.9 (20.4)	0.916
Maximum 2D diameter column (mm)	66.8 (16.1)	49.3 (22.1)	0.926
Maximum 2D diameter row (mm)	53.4 (6.63)	47.2 (19.7)	0.758
Maximum 2D diameter slice (mm)	61.7 (11.3)	49.1 (22.4)	0.768
Maximum 3D diameter (mm)	45.0 (6.75)	48.9 (19.8)	0.328
Mesh volume (mm^3)	8.40×10^5 (1.42×10^5)	5.58×10^5 (7.72×10^4)	0.723
Minor axis length (mm)	59.2 (20.6)	16.5 (1.94)	0.310
Sphericity	0.71 (0.12)	0.74 (0.06)	0.008**
Surface area (mm^2)	5.73×10^4 (1.21×10^4)	8.21×10^4 (1.33×10^4)	0.364
Surface volume ratio	0.23 (0.15)	0.31 (0.17)	0.005**
Voxel volume (mm^3)	8.41×10^5 (1.10×10^5)	6.64×10^5 (1.99×10^5)	0.454

*$p < 0.05$, **$p < 0.01$
Values are presented as median (IQR) or number (%).
IQR – interquartile range

Table 4. Differences between groups in apparent diffusion coefficient (ADC) and ADC ratios

Features	Benign (median [IQR])	Malignant (median [IQR])	*p*-value	Significance
ADCt	1744.5 [1436.0-2172.3]	1174.0 [1026.0-1404.5]	< 0.001	***
Ratio1_liver	1.79 [1.39-2.22]	1.11 [0.94-1.32]	< 0.001	***
Ratio2_spleen	2.34 [1.82-2.82]	1.49 [1.26-1.81]	< 0.001	***
Ratio3_muscle	1.21 [0.99-1.49]	0.80 [0.71-0.95]	< 0.001	***

***$p < 0.001$
ADCt – ADC tumor, Ratio1_liver – lesion-to-liver ADC, Ratio2_spleen – lesion-to-spleen ADC, Ratio3_muscle – lesion-to-paraspinal muscle ADC, IQR – interquartile range, Sig – significance

models. Meanwhile, the SVM model with Model 1 displayed the strongest Se (0.947), suggesting the highest ability to detect positive cases. Overall, the combined feature set (Model 3) provided the most consistent performance across the evaluated metrics, while Model 1 also performed well despite using fewer features. In contrast, Model 2 (the top 12 radiomic features selected by RFE) showed the weakest results, with lower accuracy and AUC values. Table 6 and Figure 3 show the classification performance of each machine learning (ML) model.

Discussion

We hypothesized that combining original radiomic features from T2w imaging with ADC values and ADC ratios derived from DWI of liver tumors can improve the classification of benign and malignant lesions. Theoretically, radiomic features from T2w capture tumor shape, texture, and heterogeneity, providing microstructural information that is difficult to discern visually [13,22]. Meanwhile, ADC values/ratios from DWI reflect water diffusion within tissue, which is closely related to cellular density and tumor biology [6,23]. Integrating these two sources of information allows for a simultaneous assessment of both the morphological and functional characteristics of tumors, potentially enhancing classification accuracy compared with using either modality alone. Recent studies have also shown that combining structural and functional imaging often achieves higher predictive performance, particularly in lesions with overlapping clini-

Table 5. Differences between groups in T2-weighted-based original radiomic features

Features	Benign (median [IQR])	Malignant (median [IQR])	*p*-value	Significance
ngtdm_Contrast	0.062 [0.024-0.122]	0.020 [0.013-0.034]	< 0.001	***
glcm_Idn	0.907 [0.888-0.923]	0.932 [0.914-0.948]	< 0.001	***
glcm_Idmn	0.974 [0.962-0.982]	0.987 [0.979-0.992]	< 0.001	***
glszm_ZonePercentage	0.062 [0.029-0.107]	0.035 [0.024-0.044]	< 0.05	**
firstorder_InterquartileRange	64.125 [30.000-102.062]	36.000 [24.000-45.000]	< 0.05	**
glrlm_ShortRunEmphasis	0.781 [0.749-0.861]	0.731 [0.699-0.785]	< 0.05	**
oglrlm_RunPercentage	0.719 [0.666-0.812]	0.656 [0.613-0.707]	< 0.05	**
gldm_SmallDependenceEmphasis	0.067 [0.041-0.105]	0.043 [0.033-0.053]	< 0.05	**
glrlm_RunLengthNonUniformityNormalized	0.587 [0.531-0.705]	0.522 [0.470-0.589]	< 0.05	**
glcm_Id	0.614 [0.522-0.682]	0.676 [0.631-0.735]	< 0.05	**
firstorder_Uniformity	0.177 [0.143-0.298]	0.256 [0.209-0.362]	< 0.05	*
gldm_DependenceNonUniformityNormalized	0.072 [0.059-0.096]	0.059 [0.052-0.072]	< 0.05	*

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$
IQR – interquartile range

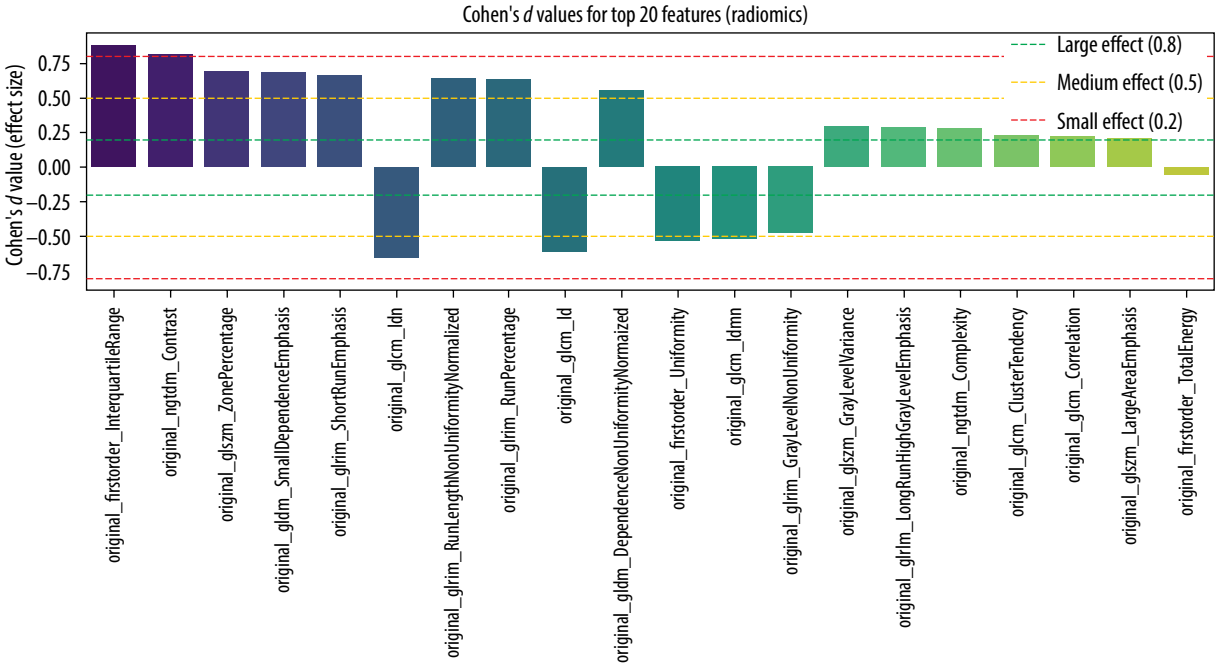


Figure 2. Cohen's *d* values for the top 20 T2-weighted-based radiomic features from recursive feature elimination. Especially, firstorder_InterquartileRange, ngtdm_Contrast, and glszm_ZonePercentage exhibited the largest effect sizes

cal and imaging features between benign and malignant types. To our knowledge, this is the first study to apply this integrated set of features for liver tumor classification.

Shape characteristic

In our study, malignant liver tumors showed significantly lower sphericity and a surface-to-volume ratio compared with benign ones, indicating that malignant tumors were less spherical. However, unexpectedly, they had a smaller relative surface area. According to previous studies, malignant tumors typically display irregular, spiculated, and complex shapes with lower sphericity and compactness but higher surface area and surface-to-volume ratio, whereas benign tumors exhibit the opposite pattern [13]. Tumor volume may explain this finding: malignant tumors tend to be larger, and a larger volume reduces the surface-to-volume ratio. Therefore, despite their irregular shapes, their size decreases the overall ratio, reflecting a volume effect rather than smoother margins.

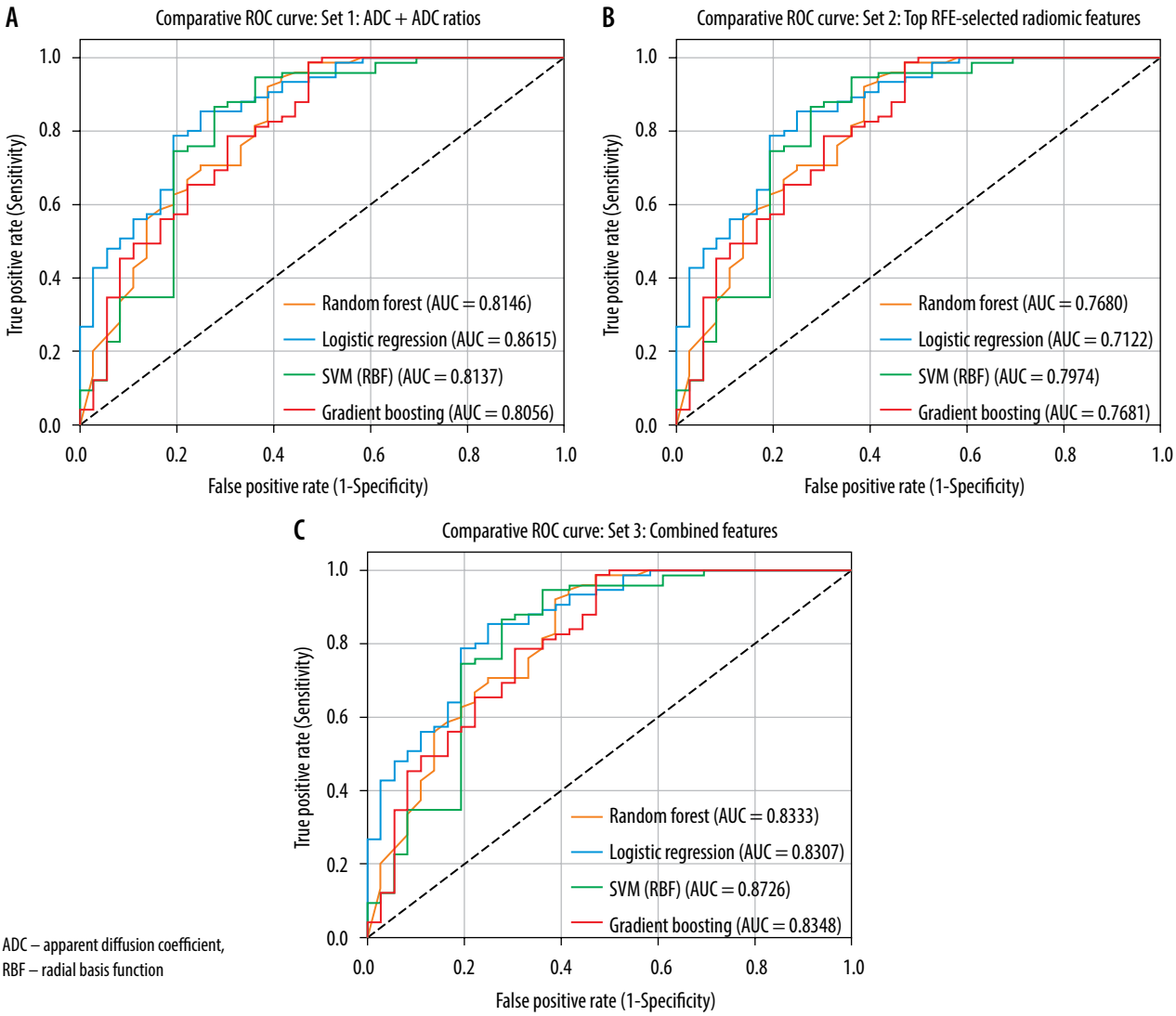
Diffusion and T2w radiomic features differentiate benign from malignant liver tumors

Both ADC values and ADC ratios showed significant differences between the two groups ($p < 0.05$), consistent

Table 6. Classification performance of each model

Feature set	Model	Acc (mean CV)	Se (mean CV)	Sp (mean CV)	AUC (mean CV)	AUC (aggregated ROC)
Model 3 (Combined features)	SVM	0.819	0.920	0.611	0.876	0.873
Model 1 (ADC and ADC ratios)	Logistic regression	0.811	0.893	0.643	0.872	0.862
Model 3 (Combined features)	Gradient boosting	0.811	0.893	0.639	0.842	0.835
Model 3 (Combined features)	Random forest	0.838	0.907	0.696	0.829	0.833
Model 3 (Combined features)	Logistic regression	0.846	0.893	0.750	0.844	0.831
Model 1 (ADC and ADC ratios)	Random forest	0.793	0.880	0.614	0.828	0.815
Model 1 (ADC and ADC ratios)	SVM	0.838	0.947	0.611	0.830	0.814
Model 1 (ADC and ADC ratios)	Gradient boosting	0.748	0.800	0.643	0.817	0.806
Model 2 (Radiomic features)	SVM	0.757	0.907	0.446	0.804	0.797
Model 2 (Radiomic features)	Gradient boosting	0.785	0.893	0.561	0.772	0.768
Model 2 (Radiomic features)	Random forest	0.775	0.920	0.475	0.780	0.768
Model 2 (Radiomic features)	Logistic regression	0.738	0.867	0.471	0.732	0.712

Acc – averaging accuracy, ADC – apparent diffusion coefficient, AUC – area under the ROC curve, CV – cross validation, RFE – recursive feature elimination, ROC – receiver operating characteristic, Se – sensitivity, Sp – specificity



ADC – apparent diffusion coefficient, RBF – radial basis function

Figure 3. The receiver operating characteristic (ROC) curve was used to evaluate the area under the curve (AUC), averaging accuracy (Acc), sensitivity, and specificity (Sp) for three distinct models (Model 1, Model 2, and Model 3). Each model was implemented using four classical machine learning algorithms: Logistic Regression (LR), Support Vector Machine (SVM), Gradient Boosting (GB) and Random Forest (RF). Overall, the LR algorithm within Model 3 achieved the highest accuracy (Acc; 0.8462) and specificity (Sp; 0.7500) among the tested models

with findings from previous studies [7,23,24]. Additionally, the use of ADC indices has demonstrated potential in distinguishing benign from malignant tumors, as well as in tumor staging and treatment monitoring [25].

For T2w radiomic features, group comparisons of the top 20 selected features revealed 12 with significant differences ($p < 0.05$). The most notable differences were observed for firstorder_InterquartileRange, ngtdm_Contrast, and glszm_ZonePercentage, all with Cohen's $d > 0.70$ and $p < 0.05$ (Table 5, Figure 2). These findings may reflect differences in the underlying tissue characteristics of benign and malignant tumors. These differences were captured through radiomic features derived from T2w images, a pulse sequence considered particularly sensitive to pathological changes within lesions. Additionally, these differences may reflect heterogeneity within each group, as the benign group included various tumor types (hemangioma, adenoma, FNH, inflammatory lesions), while the malignant group included HCC, cholangiocarcinoma, metastases, and hepatoblastoma. In this study, both ADC indices and T2w-based radiomic features showed significant differences between the two groups. This raises the question of how adding these features may affect classification performance.

Combined diffusion and radiomic features for classifying liver tumors

This study compared four common ML models in the binary classification of benign and malignant FLLs. Model 3, using 12 features with LR and SVM, achieved the best overall performance regarding accuracy and AUC; however, Model 1 (ADC and ADC Ratios) with SVM showed the highest Se. Conversely, Model 2 (Radiomic Features) showed the poorest performance, especially in Sp, suggesting that radiomic features alone were less effective than the ADC and ratio features. Models 1 and 3 achieved AUCs greater than 0.800 across all tested ML models. The combination of selected radiomic features and ADC indices was associated with improvement in classification performance, indicating potential usefulness for further diagnostic evaluation. In this study, the SVM model with a radial basis function (RBF) kernel achieved the highest AUC among the tested models. This may reflect the RBF kernel's ability to model non-linear relationships between radiomic features and ADC indices, which are often complex in biological tissues. The RBF kernel effectively maps input data into a higher-dimensional space, which can facilitate separation between benign and malignant tumors. Moreover, SVM can be well suited for datasets with relatively small sample sizes and high-dimensional feature spaces, common characteristics of radiomics-based studies.

Zhang *et al.* [26] integrated two clinical indicators with DWI radiomic features and ADC values, and the proposed model exhibited robust predictive performance in patients with endometrial cancer. Notably, the AUCs ranged between 0.852 and 0.980 across the training, testing, and independent external validation cohorts. Another study, by Martin *et al.* [11], used automated T2w radiomic features to differentiate solid benign and malignant liver lesions in cirrhotic patients. Their results indicated that T2w radiomics showed potential for differentiation, achieving AUCs between 0.700 and 0.800. These findings suggest that combining T2w-based radiomic characteristics with ADC features could improve model performance, although specific evidence remains inconclusive. Research on other organs has already demonstrated the potential complementary value of radiomic features and ADC values in distinguishing benign from malignant lesions [27,28]. Consistent with this, our study showed that combining T2w radiomics with ADC values was associated with improved classification performance across various machine learning models. Notably, AUC values were consistently above 0.800 when these features were integrated. Given the use of radiomic features, future research should focus on explainable artificial intelligence tools, including SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME), to provide clearer insights into feature importance within the model's decision-making process [11].

Limitations

This study has several limitations. First, selection bias may arise because some lesions were diagnosed based on radiologic features in accordance with current guidelines, which may not have been pathologically confirmed. Second, the sample size was small and limited to a single institution; larger multicenter studies are needed to validate our model. Third, three-dimensional lesion segmentation was performed manually, a process that is inherently time-consuming. Segmentation of tissues or lesions is a crucial step in the radiomics workflow and can be done manually, semi-automatically, or fully automatically. However, manual segmentation is often limited by its time consumption, susceptibility to inter-observer variability, and difficulties with larger datasets. Therefore, semi-automatic or automatic segmentation may improve efficiency and reproducibility [22]. Future research should focus on developing fully automated methods with higher reliability and reproducibility. Fourth, the use of 4-mm slice thickness in T2w imaging results in anisotropic voxels and increased partial volume effects. This may reduce the reliability of three-dimensional shape descriptors and wavelet-derived features, particularly in small tumors. To mitigate this issue, these features were excluded from the predictive modeling. Last but not least, calculating ADC metrics using ROIs may not fully capture the cellular and diffusion properties of localized liver tumors, potentially limiting the accuracy of the classification model. To address this, we carefully positioned ROIs in regions that clearly

reflected pathological features, avoiding areas with fluid, necrosis, or large blood vessels. Additionally, ADC values from healthy liver tissue, spleen, and paraspinal muscle were included to minimize the impact of noise inherent in DWI sequences. This study is the first to combine ADC values/ratios with T2w-based radiomic features of liver masses for distinguishing benign from malignant lesions. However, we only used original radiomic features and did not include higher-order statistical radiomic features. Incorporating these higher-order statistics could expand the dataset, requiring thorough evaluation using a broader range of machine learning or deep learning models to ensure statistical power and robustness.

Conclusions

This study indicates that T2w-based radiomics features, combined with ADC and ADC ratio values, may help distinguish malignant from benign liver tumors. Although the results are encouraging, the model requires validation in larger, multicenter studies to confirm its effectiveness.

Disclosures

1. Institutional review board statement: Te study was approved by the Ethics Committee of University of Medicine and Pharmacy, Hue University, with approval number 01/22.
2. Assistance with the article: None.
3. Financial support and sponsorship: None.
4. Conflicts of interest: None.
5. AI use statement: None.

References

1. Salem S, Houseni M, Zidan L, Kandil A. The added value of PET/Ce-CT/DW-MRI fusion in assessment of hepatic focal lesions: PET/Ce-CT/DW-MRI fusion in hepatic focal lesion. Nucl Med Biol 2015; 42: 637-642.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.
3. Crowe PM, Olliff JF. MRI of focal liver lesions. Imaging 1998; 10: 59-72.
4. Maino C, Vernuccio F, Cannella R, Cortese F, Franco PN, Gaetani C, et al. Liver metastases: the role of magnetic resonance imaging. World J Gastroenterol 2023; 29: 5180-5197.
5. Bammer R. Basic principles of diffusion-weighted imaging. Eur J Radiol 2003; 45: 169-184.
6. Koh DM, Collins DJ. Diffusion-weighted MRI in the body: applications and challenges in oncology. Am J Roentgenol 2007; 188: 1622-1635.
7. Nalaini F, Shahbazi F, Mousavinezhad SM, Ansari A, Salehi M. Diagnostic accuracy of apparent diffusion coefficient (ADC) value in differentiating malignant from benign solid liver lesions: a systematic review and meta-analysis. Br J Radiol 2021; 94: 20210059. DOI: 10.1259/bjr.20210059.
8. Sharma S, Dalal V, Prem Kumar G, Malhotra P. Diffusion MRI with quantification of ADC value in characterization of benign and malignant hepatic lesions and their correlation with cyto-histopathology. IP Int J Med Paediatr Oncol 2020; 6: 131-135.
9. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to radiomics. J Nucl Med 2020; 61: 488-495.
10. Baishya NK, Baishya K, Baishya K, Sarma R, Ray S. MRI radiomics in imaging of focal hepatic lesions: a narrative review. Cureus 2024; 16: e62570. DOI: 10.7759/cureus.62570.
11. Starmans MPA, Miclea RL, Vilgrain V, Ronot M, Purcell Y, Verbeek J, et al. Automated assessment of T2-weighted MRI to differentiate malignant and benign primary solid liver lesions in noncirrhotic livers using radiomics. Acad Radiol 2024; 31: 870-879.
12. Yang C, Turgunov B, Zhu F, Xia Y, Zhou J, Zhao Z. A T2 weighted imaging-based radiomics nomogram for the classification of hepatic blood-rich lesions: hepatocellular carcinoma and benign liver lesions. Discov Oncol 2025; 16: 1223. DOI: 10.1007/s12672-025-02868-7.
13. Wu J, Liu A, Cui J, Chen A, Song Q, Xie L. Radiomics-based classification of hepatocellular carcinoma and hepatic haemangioma on precontrast magnetic resonance images. BMC Med Imaging 2019; 19: 23. DOI: 10.1186/s12880-019-0321-9.
14. Tran HPD, Hoang NT, Dang CN. Benign versus malignant solid liver lesions: how can apparent diffusion coefficient value be added to the differentiation? 2025; 29: 216-224.
15. Association E, Montagnese S, Rautou P, Romero-gómez M, Larsen FS, Shawcross DL, et al. EASL Clinical Practice Guidelines on the management of hepatocellular carcinoma. J Hepatol 2022; 77: 807-824.
16. Ozaki K, Higuchi S, Kimura H, Gabata T. Liver metastases: correlation between imaging features and pathomolecular environments. Radiographics 2022; 42: 1994-2013.
17. Colombo M, Forner A, Ijzermans J, Paradis V, Reeves H, Vilgrain V, et al. EASL Clinical Practice Guidelines on the management of benign liver tumours. J Hepatol 2016; 65: 386-398.
18. Perniciano A, Loddo A, Di Ruberto C, Pes B. Insights into radiomics: impact of feature selection and classification. Multimed Tools Appl 2025; 84: 31695-31721.
19. Zhang YP, Zhang XY, Cheng YT, Li B, Teng XZ, Zhang J, et al. Artificial intelligence-driven radiomics study in cancer: the role of feature engineering and modeling. Mil Med Res 2023; 10: 22. DOI: 10.1186/s40779-023-00458-8.

20. Saini SK, Thakur N, Juneja M. Radiomics based diagnosis with medical imaging: a comprehensive study. Wirel Pers Commun 2023; 130: 481-514.
21. Çorbacıoğlu ŞK, Aksel G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: a guide to interpreting the area under the curve value. Turkish J Emerg Med 2023; 23: 195-198.
22. Maino C, Vernuccio F, Cannella R, Franco PN, Giannini V, Dezio M, et al. Radiomics and liver: where we are and where we are headed? Eur J Radiol 2024; 171: 111297. DOI: 10.1016/j.ejrad.2024.111297.
23. Caraiani C, Chiorean L, Fenesan DI, Lebovici A, Feier D, Gersak M, et al. Diffusion weighted magnetic resonance imaging for the classification of focal liver lesions as benign or malignant. J Gastrointest Liver Dis 2015; 24: 309-317.
24. Sutherland T, Steele E, van Tonder F, Yap K. Solid focal liver lesion characterisation with apparent diffusion coefficient ratios. J Med Imaging Radiat Oncol 2014; 58: 32-37.
25. Loy LM, How GY, Low HM, Pua U, Hwee Quek LH, Tan CH. DWI/ADC in response assessment after local-regional treatment of HCC – pearls and pitfalls. Eur J Radiol 2025; 188: 112156. DOI: 10.1016/j.ejrad.2025.112156.
26. Zhang M, Wang X, Li Z, Liu W, Li X, Jin X, et al. A radiomics model based on diffusion-weighted imaging developed using machine learning enables prediction of microsatellite instability in endometrial cancer. BMC Med Imaging 2025; 25: 405. DOI: 10.1186/s12880-025-01944-2.
27. Li C, Yin J. Radiomics based on T2-weighted imaging and apparent diffusion coefficient images for preoperative evaluation of lymph node metastasis in rectal cancer patients. Front Oncol 2021; 11: 671354. DOI: 10.3389/fonc.2021.671354.
28. Wu S, Jiao Y, Zhang Y, Ren X, Li P, Yu Q, et al. Imaging-based individualized response prediction of carbon ion radiotherapy for prostate cancer patients. Cancer Manag Res 2019; 11: 9121-9131.