

Radiomics-based differentiation of hepatocellular carcinoma and dysplastic nodules using noncontrast abbreviated MRI

Jun Young Park^{1, #}, Yoonseok Lee^{2, #}, Young Jae Kim³, Seung Kak Shin² and Kwang Gi Kim⁴

Abstract

Background: Noncontrast abbreviated magnetic resonance imaging (NC-AMRI) has emerged as a cost-effective alternative to ultrasound for hepatocellular carcinoma (HCC) surveillance; however, its diagnostic performance for differentiating dysplastic nodules (DNs) from HCC remains limited in the absence of contrast enhancement. This study investigated whether machine learning (ML)-based radiomics using noncontrast AMRI can improve discrimination between DN and HCC.

Methods: This retrospective study included 189 patients with histopathologically confirmed hepatic nodules (41 DN and 148 HCC). NC-AMRI, defined in this study as a contrast-free abbreviated protocol excluding dynamic contrast-enhanced phases, was performed using 3.0 T MRI systems and included a limited set of four noncontrast imaging sequences: in-phase, opposed-phase, T2-weighted imaging (T2WI), and diffusion-weighted imaging (DWI). Model performance in differentiating HCC and DN was assessed using 5-fold stratified cross-validation on the full dataset. Radiomic features were extracted from manually segmented regions of interest using an Image Biomarker Standardization Initiative (ISBI) compliant PyRadiomics pipeline. Feature selection was performed using recursive feature elimination after multicollinearity filtering. Logistic regression, support vector machine, random forest, and extreme gradient boosting models were trained.

Results: In terms of diagnostic performance, the machine learning-based models for differentiating HCC from DN showed AUC values of 0.798 for logistic regression, 0.790 for support vector machine, 0.781 for random forest, and 0.758 for extreme gradient boosting, with corresponding sensitivities of 0.712, 0.669, 0.698, and 0.741 and specificities of 0.756, 0.805, 0.756, and 0.610, respectively. Notably, the machine learning-based models also demonstrated good performance in distinguishing high-grade DN from HCC, achieving AUC values of 0.783 for logistic regression, 0.751 for support vector machine, 0.800 for random forest, and 0.742 for extreme gradient boosting, with corresponding sensitivities of 0.799, 0.835, 0.727, and 0.799 and specificities of 0.700, 0.500, 0.700, and 0.600, respectively.

Conclusions: Radiomics-based machine learning enhances the diagnostic performance of NC-AMRI for differentiating HCC from DN, complementing conventional imaging and addressing key limitations of AMRI surveillance.

¹Department of Translational-Clinical Medicine, Gachon University, Incheon, Republic of Korea

²Department of Internal Medicine, Gachon University Gil Medical Center, Gachon University College of Medicine, Incheon, Republic of Korea

³Gachon Biomedical & Convergence Institute, Gachon University Gil Medical Center, Incheon, Republic of Korea

⁴Medical Devices R&D Center, Gachon University Gil Medical Center, Incheon, Republic of Korea

[#]Jun Young Park, Yoonseok Lee, co-author contributed equally.

*Seung Kak Shin, Kwang Gi Kim, co-corresponding author contributed equally.

Corresponding authors:

Seung Kak Shin, Gil Medical Center, Gachon University College of Medicine 38-13, Dokjeom-ro 3beon-gil, Namdong-gu, Incheon, Republic of Korea.
Email: puressk@gilhospital.com

Kwang Gi Kim, Medical Center, Gachon University College of Medicine, 38-13, Dokjeom-ro 3beon-gil, Namdong-gu, Incheon, Republic of Korea.
Email: kimkg@gachon.ac.kr



Keywords

Hepatocellular carcinoma, dysplastic nodules, radiomics, machine learning

Received: 2 September 2025; accepted: 6 March 2026

Introduction

The prevalence of hepatocellular carcinoma (HCC) is increasing worldwide and is quickly becoming one of the leading causes of death.¹ More than 80% of patients with HCC are predisposed to underlying cirrhosis. Hepatitis B virus (HBV) infection, hepatitis C virus (HCV) infection, alcohol consumption, and nonalcoholic fatty liver disease (NAFLD) are major risk factors for HCC.² While most of these factors promote hepatocarcinogenesis through cirrhosis, HBV can also lead to HCC in the absence of cirrhosis. The standard international guidelines recommend biannual surveillance with ultrasound and alpha fetoprotein (AFP) testing in at-risk patients.^{2,3} However, ultrasound has a low sensitivity and specificity for detecting early-stage HCC. As such, there is a need for better imaging modalities.⁴ One proposed alternative to ultrasound is the abbreviated magnetic resonance imaging (AMRI), a shortened MRI protocol designed to reduce examination time and cost.⁵ AMRI comprises both noncontrast abbreviated magnetic (NC-AMRI) and contrast-enhanced AMRI protocols. Among these, NC-AMRI has attracted particular interest for HCC surveillance because it does not require contrast administration, thereby avoiding contrast-related risks while offering the shortest acquisition time and favorable cost-effectiveness. NC-AMRI, which typically includes T2-weighted, T1-weighted, and diffusion-weighted images, has been evaluated as a potential HCC screening tool due to its short examination time and good cost-effectiveness.^{6–8} In previous studies, AMRI has shown high diagnostic accuracy compared to ultrasound.^{6,7} For example, Park et al. reported that ultrasound-based screening achieved a sensitivity of 27.9% and an accuracy of 93.6%, whereas the AMRI-based approach achieved a sensitivity of 86.0% and an accuracy of 95.2%.⁶ Similarly, a meta-analysis by Gupta et al. found that the sensitivity of ultrasound was 53%, while AMRI demonstrated a sensitivity ranging from 82% to 86%. Notably, NC-AMRI achieved diagnostic performance comparable to that of contrast-enhanced AMRI, with both showing a sensitivity of 86% and a specificity of 94%.⁷

A dysplastic nodule (DN) is a regenerative nodule with atypical cells that are distinct from the surrounding liver parenchyma.⁹ DN is a premalignant lesion that can progress to HCC. Therefore, it is difficult to differentiate between high-grade DN (HGDN) and early HCC on radiologic imaging.¹⁰ Contrast-enhanced dynamic CT or MRI improves the differentiation of DNs from HCC.¹¹ The typical imaging finding of HCC is arterial phase contrast

uptake followed by delayed phase washout.^{12,13} When using a hepatocyte-specific contrast agent, such as gadoteric acid (gadolinium ethoxybenzyl dimeglumine; Primovist) in MRI, features of the hepatobiliary phase can significantly increase the diagnostic accuracy of HCC.^{14,15} However, these features are not available in nonenhanced MR, making it more difficult to differentiate DN from HCC.

Recent studies have proposed generative deep learning-based abbreviated MRI protocols that synthesize contrast-enhanced images from noncontrast MRI, demonstrating diagnostic performance for hepatocellular carcinoma comparable to that of conventional contrast-enhanced MRI while substantially reducing scan time.¹⁶ Deep learning-accelerated noncontrast abbreviated liver MRI has been shown to improve image quality and maintain high sensitivity for detecting malignant focal hepatic lesions, including hepatocellular carcinoma, compared with standard abbreviated MRI protocols.¹⁷ Deep learning models applied to whole-slide histopathologic images have demonstrated high accuracy in classifying hepatocellular nodular lesions, including the challenging differentiation between high-grade dysplastic nodules and well-differentiated hepatocellular carcinoma.¹⁸ An explainable deep learning model trained on gadoteric acid-enhanced MRI achieved high diagnostic accuracy for hepatocellular carcinoma and improved radiologists' sensitivity by assisting with automated lesion classification and imaging feature interpretation.¹⁹

Although several studies have investigated radiologic or radiomics-based approaches for differentiating DN from HCC, most prior work has relied on contrast-enhanced imaging, and evidence regarding the diagnostic value of radiomics on nonenhanced MRI remains limited. Furthermore, the ability of radiomics to distinguish HGDN from HCC using noncontrast MRI has not been sufficiently explored.

Radiomics is an emerging field in medical imaging that enables the extraction of high-throughput, quantitative features such as texture, shape, and intensity from medical images.²⁰ These features, which are often imperceptible to the human eye, provide mathematical and statistical insights that can complement conventional image interpretation. Quantitative data obtained through radiomics can complement traditional image assessments.^{21,22} This study aimed to determine whether radiomics can be used to successfully differentiate between DNs and HCC on NC-AMRI. The primary aim of this study was to evaluate the potential of radiomics-based machine learning models using nonenhanced MRI to differentiate DN, particularly HGDN, from HCC and to explore the diagnostic applicability of this approach in a noncontrast MRI setting.

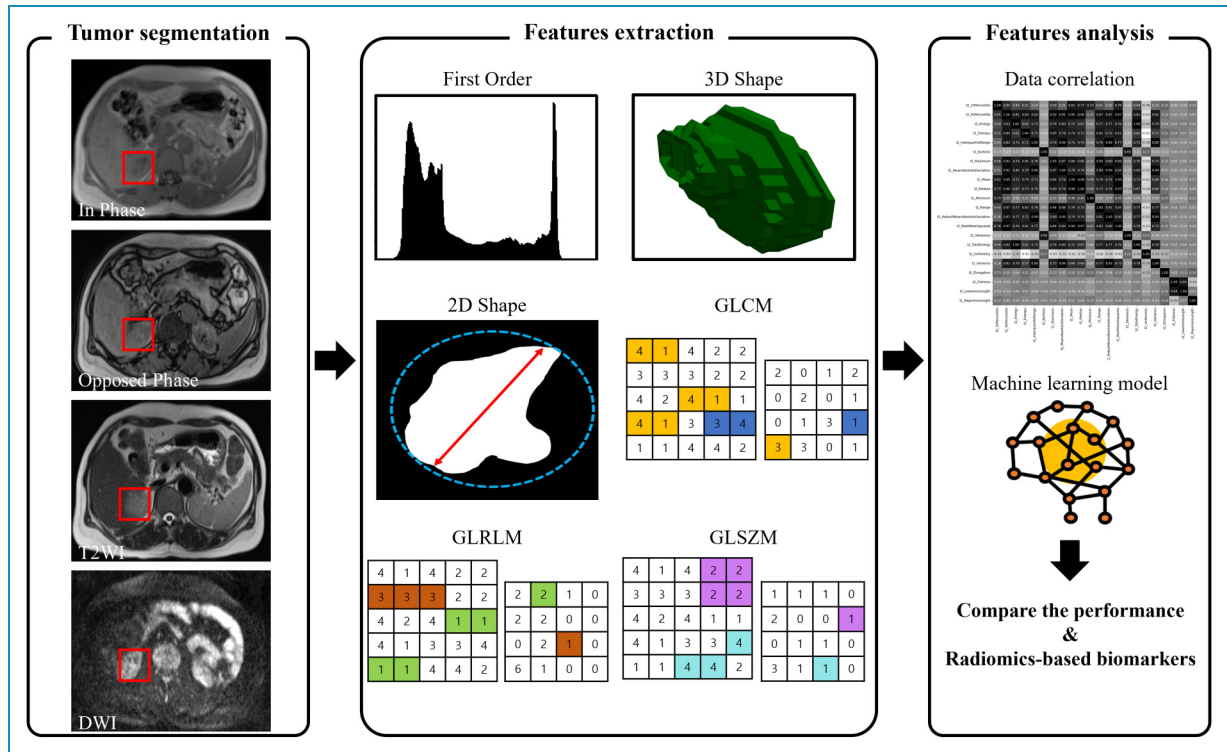


Figure 1. Radiomics feature extraction using four phases of MRI images.

Materials and methods

Participants and data collection

A total of 189 patients with hepatic nodules detected on NC-AMRI and histologically confirmed HCC (n = 148) or DN (n = 41) between February 2009 and May 2021 were enrolled. All AMRI examinations were performed using NC-AMRI protocols.

The NC-AMRI protocol consisted of four noncontrast imaging sequences: in-phase and opposed-phase T1-weighted imaging, T2-weighted imaging (T2WI), and diffusion-weighted imaging (DWI). This protocol was intentionally limited to a small number of noncontrast sequences and did not include dynamic contrast-enhanced imaging, in order to reduce acquisition time and reflect abbreviated MRI strategies proposed for HCC surveillance.

Among the 41 patients with DN, 28 were diagnosed with HGDN, and 13 were diagnosed with low-grade dysplastic nodule (LGDN). Patients with more than one hepatic nodule were excluded from the data collection. Patients with multiple hepatic nodules were excluded to avoid ambiguity in lesion correspondence during image segmentation and feature extraction. Histopathologic confirmation was obtained for all cases and served as the reference standard. The classification of nodules into LGDN, HGDN, and HCC was based on standard histopathologic criteria, including cellularity, nuclear atypia, and architectural patterns. Region of interest (ROI) segmentation was performed

manually. All cases were reviewed by an abdominal radiologist and a hepatologist independently. Due to the small sample size, we conducted 5-fold stratified cross-validation using the entire dataset without a separate hold-out test set. In each fold, 80% of the data was used for model training and 20% for validation, ensuring internal performance assessment across all samples.

Feature extraction

Radiomics involves the extraction of quantitative features from medical images, which are inherently mathematical and statistical in nature, to represent quantitative values.²⁰ In addition, radiomics was employed to capture subtle texture- and intensity-based differences between DNs and HCC that are not readily appreciable by visual inspection. Image features are extracted from the HCC and nodule regions. As shown in Figure 1, we extracted 19 first-order features based on the histograms, 26 2D shape features that extract features from the size and shape of the region of interest, and 56 s-order texture features capturing spatial relationships between adjacent voxels in the image. The second-order features include 24 Gray Level Co-occurrence Matrix (GLCM) features to analyze texture by identifying the relationship between mutual pixels in the image, 16 Gray Level Run Length Matrix (GLRLM) features to extract statistical information about the run length of pixels for gray level, and 16 Gray Level Size Zone

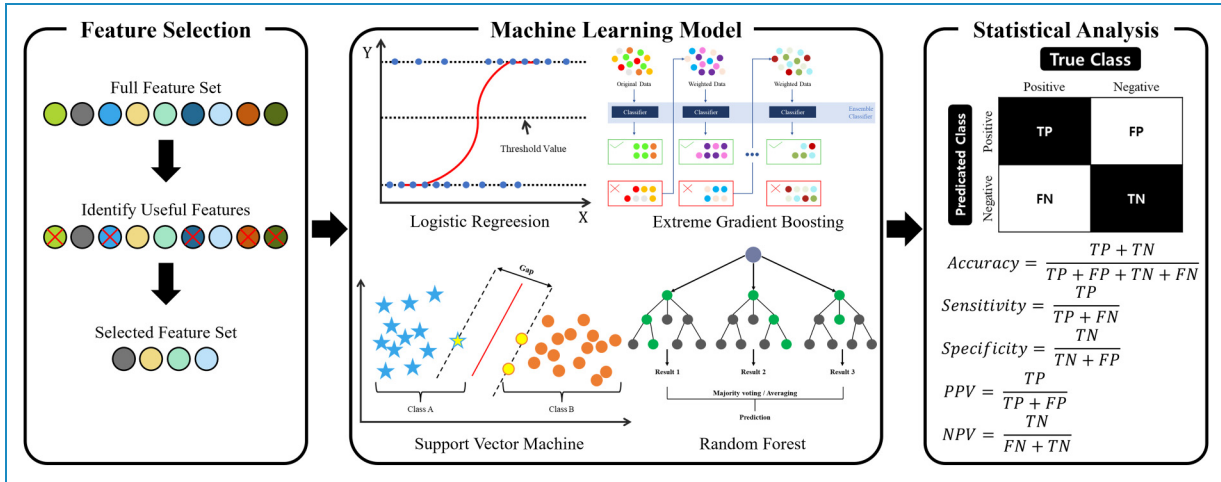


Figure 2. Workflow of feature selection and subsequent machine learning model training.

Matrix (GLSZM) features to extract detailed features of the image by highlighting information about the size of pixel values in the image.^{23–25} We extracted all features supported by the Pyradiomics library. Among all features, only meaningful features were selected using variance inflation factor (VIF) and feature selection. The extracted features were equally divided into four phases: in-phase (very low intensity), opposed phase (low intensity), T2WI (very high intensity), and DWI (high intensity), and 101 features were extracted from each phase, totaling 404 features. Radiomic features were extracted using PyRadiomics, an open-source library compliant with the Image Biomarker Standardization Initiative (IBSI) guidelines. Features were computed independently for each MRI sequence, and all feature extraction settings were fixed and uniformly applied across the entire dataset.

Machine learning model

Radiomics features underwent multicollinearity filtering ($VIF > 10$) and normalization using a robust scaler. Recursive feature elimination (RFE) was used for feature selection. The RFE process iteratively removed features with the lowest importance scores based on model coefficients until optimal performance was achieved. The stopping criterion was defined as the feature subset achieving the highest mean cross-validated area under the curve (AUC) across five folds. Four machine learning models logistic regression (LR), which assumes a linear relationship between independent and dependent variables and estimates regression coefficients from the data, support vector machine (SVM), which determines the hyperplane with the largest margin between each class; random forest (RF), which forms multiple decision trees as an ensemble ML model and selects features randomly; and extreme gradient boosting (XGB), which is a tree-based ensemble ML model as a variant of the gradient boosting algorithm were trained to differentiate HCC from dysplastic nodules

(Figure 2). Model hyperparameters were optimized using grid search with five-fold stratified cross-validation. For each model, parameter ranges were selected based on commonly recommended values in the machine learning literature and pilot experiments. Regularization strength (C) for LR and SVM was explored across logarithmic scales; ensemble parameters (number of estimators, tree depth) for RF and XGB were tested across a range of values to balance model complexity and performance; learning rates for XGB followed standard recommendations. The optimal hyperparameter set for each model was selected to maximize mean cross-validated AUC. Final performance metrics were derived from cross-validated predictions.²⁶ The final settings to classify HCCs and DNs are as follows: for LR, L2 regularization was applied with $C = 7$; for SVM, the kernel was set to “linear” with $C = 262$; for RF, the number of estimators was set to 731 with a maximum depth of 189; for XGB, the a learning rate of 0.1 with 145 estimators and a maximum depth of 5; The final settings to classify HCCs and HGDN are as follows: for LR, L2 regularization was applied with $C = 997$; for SVM, the kernel was set to “linear” with $C = 998$; for RF, the number of estimators was set to 709 with a maximum depth of 190; for XGB, the learning rate was set to 0.05 with 100 estimators and a maximum depth of 11; The feature selection methods and machine learning models used in this study were implemented using standard algorithms available in the scikit-learn library. Model training and validation were performed on a conventional workstation without the use of graphics processing unit (GPU) or other specialized hardware.

Statistical analysis

To evaluate the classification performance of HCC and DNs in liver MRI images, this study compared the pathology results with the prediction results of the ML models. HCC was defined as the positive class for binary classification.

Table 1. Patient characteristics.

	DN (n = 41)	HCC (n = 148)	P value
Age (mean $\pm$ SD, years)	58.95 $\pm$ 9.28	57.72 $\pm$ 11.08	0.375
Male (%)	28 (68.3%)	117 (79.1%)	0.011
Underlying liver disease (%)			0.002
HBV	21 (51.2%)	110 (74.3%)	
HCV	7 (17.1%)	16 (10.8%)	
HBV + HCV	0 (0%)	1 (0.6%)	
Alcohol	8 (20%)	16 (10.8%)	
Others	5 (12%)	1 (0.6%)	
LC (%)	34 (82.9%)	126 (85.1%)	0.498
Tumor size (mean $\pm$ SD, cm)	1.76 $\pm$ 0.84	2.42 $\pm$ 1.65	0.062
AFP (median (IQR) ng/mL)	10.3 (4.6–13.5)	11.6 (5.21–70.5)	0.070

The AUC, accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated using the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values obtained through comparison. AUC was calculated by performing receiver operating characteristic (ROC) curve analysis to calculate the area under the ROC curve. ROC analysis was used to comprehensively evaluate the trade-off between sensitivity and specificity across different decision thresholds. Multiple performance metrics were reported to reflect different aspects of diagnostic performance, particularly in the context of imbalanced classification between malignant and premalignant lesions.

Results

Clinical characteristics of enrolled patients

Table 1 summarizes the patient characteristics. The mean lesion size was 1.76 cm for DNs and 2.42 cm for HCC, with no significant difference between the two groups. The proportion of patients with underlying cirrhosis was 82.9% in the DN group and 85.1% in the HCC group, which did not differ significantly. In addition, the proportion of male patients was 68.3% in the DN group and 79.1% in the HCC group, showing a statistically significant difference, indicating that the proportion of male patients was significantly higher in the HCC group. In terms of underlying liver disease, HBV, HCV, and alcohol-related liver disease were observed in 51.2%, 17.1%, and 20.0% of patients with DNs, respectively, and in 74.3%, 10.8%,

and 10.8% of patients with HCC, respectively. This difference was statistically significant, indicating that HBV infection was more common in the HCC group than in the DN group.

Performance of machine learning models in differentiating between HCC and DNs

We used the feature selection method RFE and the ML models LR, SVM, RF, and XGB to differentiate between HCC and DNs on NC-AMRI. Figure 3 shows the ROC curves of the four models: LR, SVM, RF, and XGB had an AUC of 0.798, 0.790, 0.781, and 0.758, respectively, indicating that LR had the highest AUC performance.

Table 2 shows the diagnostic performance of LR, SVM, RF, and XGB using RFE-based feature selection. LR achieved the highest AUC (0.798) and accuracy (0.722). XGB showed the highest sensitivity (0.741), while SVM had the highest specificity (0.805) and PPV (0.921). The highest NPV (0.437) was observed in LR. Overall, LR showed balanced performance, and each model had strengths in different metrics.

Performance of machine learning models in differentiating between HCC and HGDNs

Radiomics also showed excellent AUC values for differentiating HGDN from HCC, which are more difficult to distinguish using imaging. As a result, LR, SVM, RF, and XGB were found to have an AUC of 0.783, 0.751, 0.800, and 0.742, respectively (Figure 4). The AUC, accuracy,

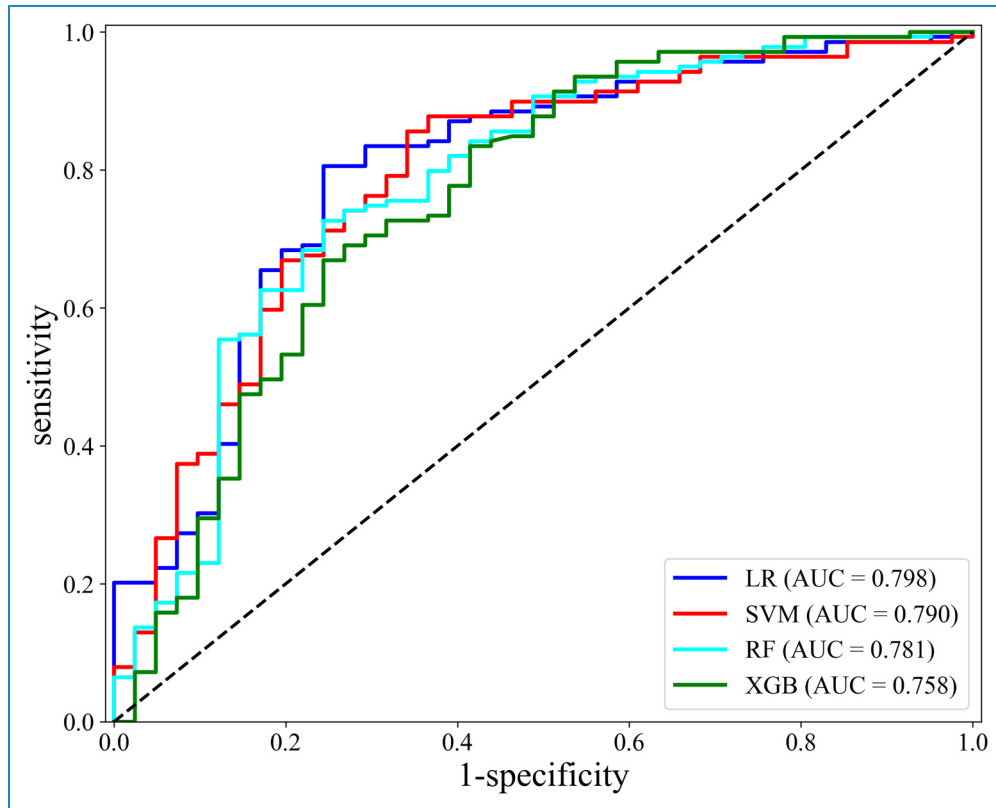


Figure 3. ROC curve for diagnostic performance in differentiating between HCC and DNs. HCC, hepatocellular carcinoma; DN, dysplastic nodule; LR, logistic regression; SVM, support vector machine; RF, random forest; XGB, extreme gradient boosting.

Table 2. AUC, accuracy, sensitivity, specificity, PPV, and NPV for the classification performance of the feature selection method RFE and the machine learning models LR, SVM, RF, and XGB.

	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
LR	0.798 (0.732–0.854)	0.722 (0.657–0.788)	0.712 (0.637–0.787)	0.756 (0.625–0.888)	0.908 (0.854–0.962)	0.437 (0.321–0.552)
SVM	0.790 (0.723–0.847)	0.700 (0.633–0.767)	0.669 (0.591–0.747)	0.805 (0.684–0.926)	0.921 (0.868–0.973)	0.418 (0.309–0.526)
RF	0.781 (0.713–0.839)	0.711 (0.645–0.777)	0.698 (0.622–0.774)	0.756 (0.625–0.888)	0.907 (0.851–0.962)	0.425 (0.311–0.538)
XGB	0.758 (0.689–0.819)	0.711 (0.645–0.777)	0.741 (0.668–0.814)	0.610 (0.460–0.759)	0.866 (0.804–0.927)	0.410 (0.286–0.533)

sensitivity, specificity, PPV, and NPV of each model are summarized in Table 3.

Feature importance analysis for predicting HCC and DNs

Figure 5 shows the feature importance values for each combination of the feature selection method, RFE, and machine-

learning model. Using the RFE feature selection method, the four features with the highest relative performance were selected. The four selected features were ShortRunLowGrayLevelEmphasis (SRLGLE) and GrayLevelNonUniformity.1 (GLN.1) for the T2WI phase and Inverse Variance and Run Variance for DWI. The feature importance of the LR model was 0.157 for GLN.1, 0.088 for Inverse Variance, 0.051 for SRLGLE, and 0.010 for Run Variance. The feature importance of the

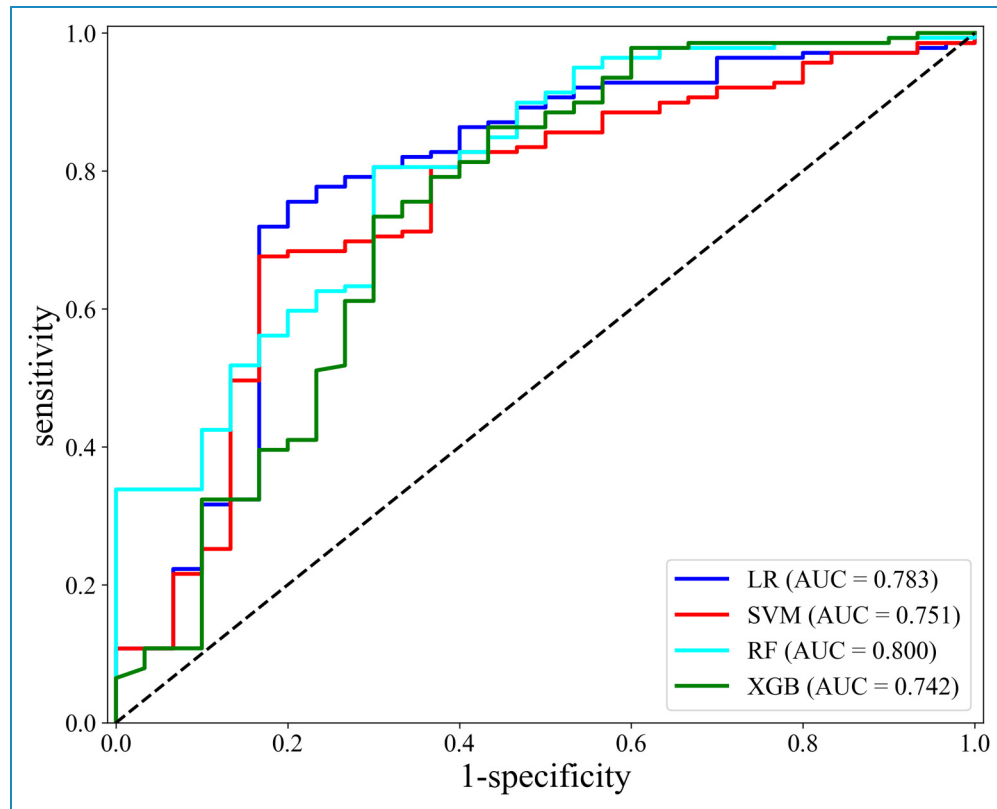


Figure 4. ROC curve for differentiation between HCC and HGDNs. HCC, hepatocellular carcinoma; HGDN, high grade dysplastic nodule. LR, logistic regression; SVM, support vector machine; RF, random forest; XGB, extreme gradient boosting.

Table 3. AUC, accuracy, sensitivity, specificity, PPV, and NPV for the classification performance of the feature selection method RFE and the machine learning models LR, SVM, RF, and XGB.

	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
LR	0.783 (0.713–0.842)	0.781 (0.719–0.843)	0.799 (0.732–0.865)	0.700 (0.536–0.864)	0.925 (0.878–0.972)	0.429 (0.290–0.567)
SVM	0.751 (0.679–0.814)	0.775 (0.712–0.838)	0.835 (0.773–0.896)	0.500 (0.321–0.679)	0.885 (0.831–0.940)	0.395 (0.239–0.550)
RF	0.800 (0.731–0.857)	0.722 (0.654–0.789)	0.727 (0.653–0.801)	0.700 (0.536–0.864)	0.918 (0.867–0.969)	0.356 (0.234–0.478)
XGB	0.742 (0.669–0.806)	0.763 (0.699–0.827)	0.799 (0.732–0.865)	0.600 (0.425–0.775)	0.902 (0.850–0.955)	0.391 (0.250–0.532)

SVM model was as follows: GLN.1, 0.145; Inverse Variance, 0.092; SRLGLE, 0.036; Run Variance, -0.002 . The feature importance of the RF model was: Inverse Variance, 0.065; SRLGLE, 0.032; GLN.1, 0.026; Run Variance, -0.004 . For the XGB model, the feature importance was: Inverse Variance, 0.111; SRLGLE, 0.066; GLN.1, 0.027; Run Variance, 0.022. When averaging the importance of each model, Inverse Variance was the most

important (0.089), followed by GLN.1 (0.088), SRLGLE (0.046), and Run Variance (0.006).

Feature importance analysis for predicting HCC and HGDN

Figure 6 shows the feature importance values for each combination of the feature selection method, RFE, and machine-

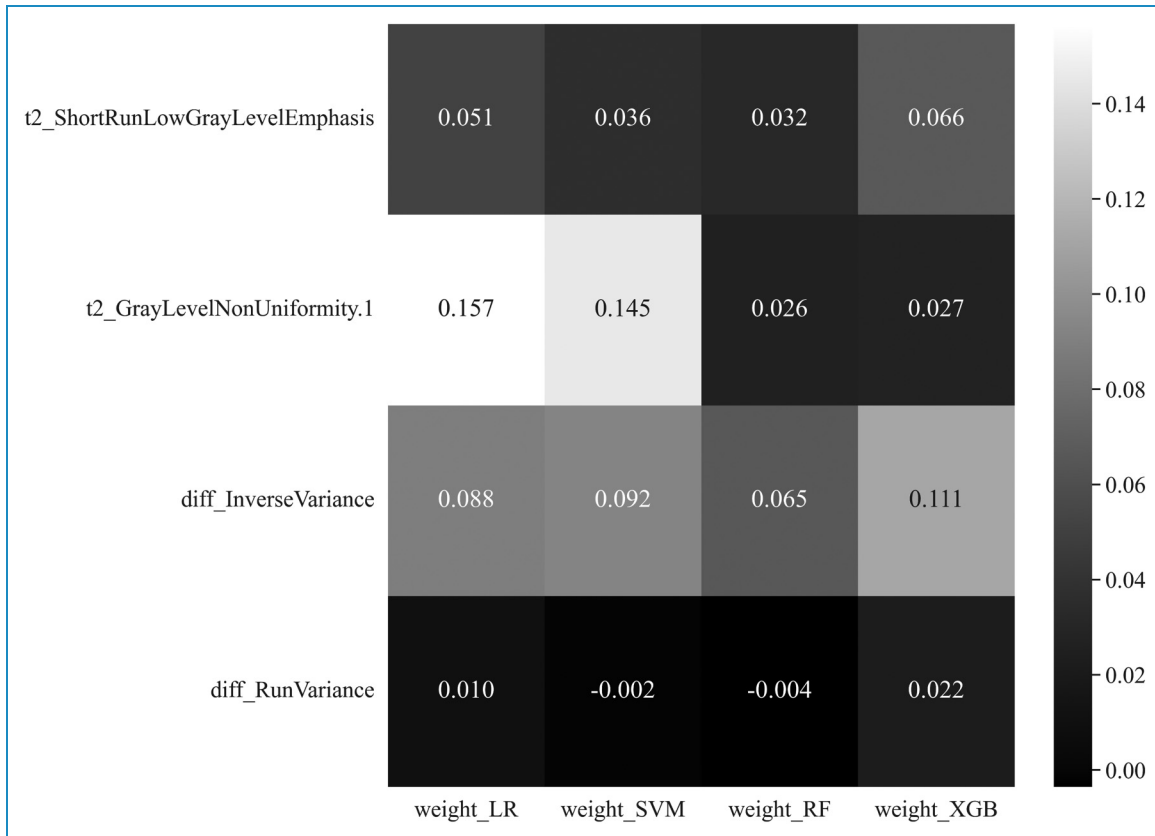


Figure 5. Feature importance graph using a combination of the feature selection method RFE and the machine learning models LR, SVM, RF, and XGB using a heatmap. Higher values are shown in white, indicating the importance of the model in classifying hepatocellular carcinoma and DN.

learning model. Using the RFE feature selection method, the four features with the highest relative performance were selected. The four selected features were the total energy, SmallAreaHighGrayLevel Emphasis (SAE) for the T2WI phase, and Flatness, Inverse Variance for DWI. The feature importance of the LR model had a Total Energy of 0.155, an SAE of 0.007, a flatness of 0.079, and an Inverse Variance of 0.057. The feature importance of the SVM model was a Total Energy of 0.173, an SAE of 0.015, a flatness of 0.058, and an Inverse Variance of 0.060. The feature importance of the RF model was a Total Energy of 0.058, an SAE of 0.017, a flatness of 0.060, and an Inverse Variance of 0.083. For the XGB model, the feature importance had a Total Energy of 0.057, an SAE of -0.002 , a flatness of 0.067, and an Inverse Variance of 0.066. When averaging the importance of each model, Total Energy was the most important (0.111), followed by Inverse Variance (0.067), flatness (0.066), and SAE (0.009).

Discussion

NC-AMRI has attracted attention due to its cost-effectiveness in the surveillance of HCC in individuals at

high risk of HCC. However, its diagnostic accuracy for HCC is inferior to that of complete MRI.^{7,27} In particular, DNs, commonly observed in patients with cirrhosis, are challenging to differentiate from HCC using noncontrast MRI.^{28,29} Therefore, the purpose of this study was to explore the feasibility of overcoming the limitations of NC-AMRI through a radiomics-based machine learning approach. From a practical perspective, the proposed NC-AMRI-based radiomics approach may offer advantages over conventional contrast-enhanced MRI. By eliminating the need for contrast agent administration, this approach can reduce examination time and avoid additional costs associated with contrast materials and related procedures, while maintaining diagnostic utility.

HCC is known to exhibit hyperintensity on T2WI and diffusion-weighted imaging restrictions on MRI. However, due to its inconsistency, it is currently used as a supplement to image findings in the arterial phase, delayed phase, and hepatobiliary phase.^{30–32} In this study, we extracted and analyzed radiomic features in only four phases, except for the arterial phase, delayed phase, and hepatobiliary phase (in-phase, opposed phase, T2WI, and DWI) of MR images, but successfully differentiated between HCC and DNs with AUCs of 0.758–0.798. In

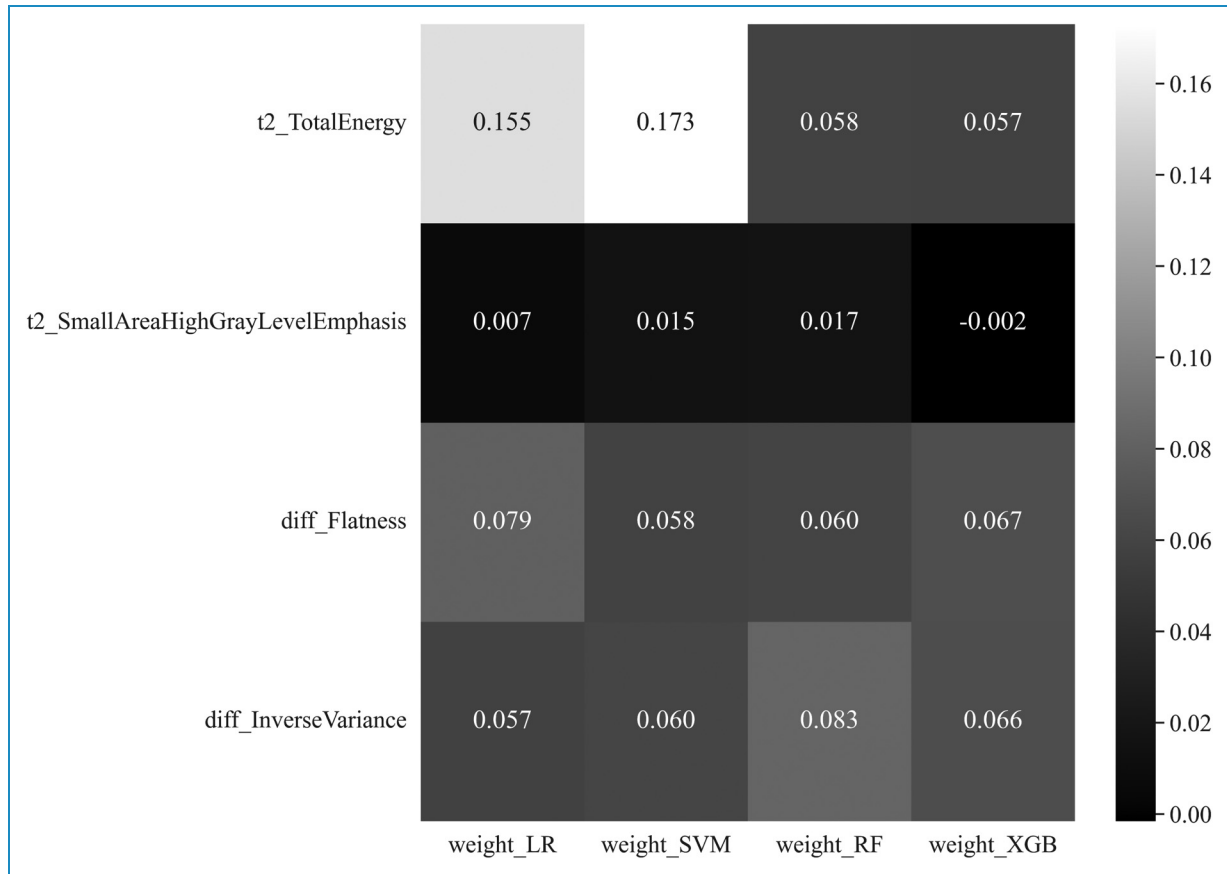


Figure 6. Feature importance graph using a combination of the feature selection method RFE and the machine learning models LR, SVM, RF, and XGB using a heatmap. Higher values are shown in white, indicating the importance of the model in classifying hepatocellular carcinoma and HGDN.

addition, the radiomics-based ML model successfully differentiated between HGDN and HCC, with AUCs of 0.742–0.800, which are traditionally more challenging to differentiate.³³

Most radiomics-based studies in liver MRI have focused on differentiating HCC from non-HCC lesions or from other primary hepatic malignancies, such as intrahepatic cholangiocarcinoma.^{34–36} While these studies demonstrate the overall feasibility of radiomics and machine learning for liver tumor classification, they address different diagnostic tasks and are therefore not directly comparable to the present study.

Recent systematic reviews summarize the expanding application of radiomics and artificial intelligence in HCC imaging, largely focusing on tumor detection, prognostication, and treatment-related outcomes.^{37–40} Within this broader context, radiomics-based differentiation between dysplastic nodules and HCC remains relatively underexplored.

Although the pathophysiological basis of individual radiomics features is not always clearly interpretable, our findings suggest that these quantitative markers capture image-based heterogeneity reflective of underlying histologic changes. For example, SRLGLE and GLN.1 from T2-weighted images, as well as Inverse Variance and Run

Variance from DWI, were among the most informative features for differentiating HCC from DNs. These texture-based features represent subtle signal variations that are often imperceptible to the human eye during conventional radiologic assessment. This highlights the potential of radiomics to complement visual interpretation by detecting quantitative patterns associated with tumor biology. In this study, a deliberately parsimonious feature selection strategy was adopted. RFE was used to reduce the high-dimensional radiomics feature space while minimizing the risk of overfitting in a relatively small cohort. The final set of four features corresponded to the smallest subset at which cross-validated performance remained stable, favoring model robustness and interpretability over marginal performance gains. We acknowledge that systematic optimization of feature number, such as learning curve or ablation analyses, may provide additional insight, but such analyses were beyond the scope of this proof-of-concept study.

The prominence of features from T2WI and DWI may reflect their sensitivity to tissue microstructure and water diffusivity, which differ between HCC and DN. T2WI captures changes in cellularity and stroma, while DWI reflects restricted diffusion in hypercellular tumors. In contrast,

in-phase and opposed-phase images mainly assess fat content, offering limited discriminatory value in this context.

From a clinical implementation perspective, several practical considerations should be acknowledged. First, manual lesion segmentation was used in this study, and formal inter-reader variability was not assessed. However, ROIs were drawn for lesions that were clinically suspicious and subsequently biopsied, which may limit ambiguity in lesion boundaries. Future studies incorporating automated or semi-automated segmentation methods will be important to improve reproducibility and scalability. Second, radiomics feature extraction and machine learning-based prediction are performed offline after image acquisition and do not prolong MRI examination time. Once lesion segmentation is completed, computational processing can be completed within a short time frame, and prediction itself is near-instantaneous. Third, while integration into routine clinical workflows was not directly evaluated in this study, the present results provide a foundation for future investigations aimed at incorporating radiomics-based decision support into NC-AMRI surveillance protocols, particularly as AMRI adoption for HCC surveillance and postoperative monitoring continues to expand. Finally, NC-AMRI is inherently cost-effective compared with contrast-enhanced MRI, and the addition of machine learning analysis does not increase scanning time or require additional imaging resources. By potentially reducing unnecessary biopsies and supporting diagnostic decision-making in settings with limited subspecialty radiology expertise, radiomics-based approaches may further enhance the cost-effectiveness of NC-AMRI. Formal cost-effectiveness analyses, however, were beyond the scope of this study.

This study has several limitations. Although patient enrollment and imaging data acquisition were completed in 2021, the present study was designed to address a persistent and clinically relevant diagnostic challenge—differentiating DNs from HCC on noncontrast MRI—rather than to benchmark rapidly evolving machine learning architectures. Accordingly, recent methodological advances such as transformer-based models or federated learning approaches for multi-center validation were not evaluated and remain important directions for future research.

This was a single-center retrospective study without external validation. Although five-fold stratified cross-validation was used to mitigate this limitation, future studies with external validation are necessary to confirm generalizability. In addition, direct comparison with radiologist interpretations or conventional contrast-enhanced MRI was not performed in this cohort, and the incremental clinical value of the proposed model relative to standard imaging protocols remains to be determined. Given the high dimensionality of radiomics features relative to the sample size, there remains a potential risk of model overfitting despite internal validation procedures. Although dimensionality reduction was performed through recursive feature elimination, residual overfitting cannot be entirely excluded given the sample size. Independent multi-

institutional validation with larger cohorts is therefore essential to establish robustness and clinical applicability.

Furthermore, ROIs were manually segmented and independently reviewed by an abdominal radiologist and a hepatologist. However, formal inter-reader agreement or reproducibility analyses were not conducted. Because radiomics features can be sensitive to segmentation variability, the potential influence of reader-dependent variation cannot be entirely excluded. In addition, manual segmentation limits scalability and reproducibility in routine clinical practice. Future studies incorporating automated or semi-automated segmentation and formal reproducibility assessment are warranted.

There was a significant difference in the distribution of underlying liver diseases between the DN and HCC groups. While chronic hepatitis B was the most common etiology in both groups, alcoholic liver disease was more frequent in the DN group (20% vs. 10.8%). Alcohol-related cirrhosis may exhibit imaging characteristics distinct from viral hepatitis-related cirrhosis, raising the possibility of confounding effects on radiomic feature extraction. The higher prevalence of alcoholic liver disease in the DN group likely reflects the real-world epidemiologic characteristics of biopsy-confirmed dysplastic nodules rather than arbitrary sampling bias. Because DNs were identified exclusively through histopathologic confirmation, such differences in etiology distribution were difficult to avoid. We did not perform etiology-adjusted or stratified analyses due to the limited number of histopathologically confirmed DNs and concerns regarding model instability in high-dimensional radiomics analyses. Accordingly, the findings should not be interpreted as demonstrating etiology-independent discrimination, and the potential influence of underlying liver disease etiology remains an important limitation.

The HCC group included tumors with diverse histopathologic grades, and distinguishing early-stage HCC based on imaging alone remains a diagnostic challenge.

And several performance-related limitations warrant consideration. The relatively low negative predictive value (~0.40) reflects the low DN prevalence (22%) in our cohort and should be interpreted accordingly, as NPV is highly dependent on disease prevalence. Our model is intended for risk stratification to identify high-risk DNs warranting intensified surveillance rather than as a rule-out test, and thus sensitivity and positive predictive value are more clinically relevant metrics. Additionally, we did not perform decision curve analysis or cost-effectiveness evaluation, which would provide further evidence for clinical implementation but require validated threshold probabilities not yet standardized for DN surveillance strategies. Future studies should incorporate decision-analytic and health economic modeling to determine optimal risk thresholds and evaluate clinical utility in real-world settings.

Conclusion

This study suggests that radiomics-based analysis of NC-AMRI offers a quantitative framework for

characterizing imaging features that are conventionally assessed visually when differentiating DNs from HCC in high-risk patients. Radiomic features extracted from in-phase, opposed-phase, T2-weighted, and diffusion-weighted images demonstrate the feasibility of capturing lesion heterogeneity without contrast enhancement. These results indicate that radiomics may complement standard visual interpretation in noncontrast AMRI, warranting further validation in larger, multicenter cohorts.

Acknowledgments

None.

ORCID iD

Kwang Gi Kim  <https://orcid.org/0000-0001-9714-6038>

Ethics approval statement

This study was approved by the Institutional Review Board of Gachon University Gil Medical Center (GDIRB2020-421), and the requirement for written informed consent was waived due to the retrospective nature of the study.

Author contribution statement

Jun Young Park developed the code, performed data analysis, prepared all figures and tables, and wrote the main manuscript draft. Yoonseok Lee contributed to study design, clinical data interpretation, and manuscript revision. Young Jae Kim provided statistical and technical support, assisted in algorithm implementation, and contributed to manuscript editing. Kwang Gi Kim supervised the overall project, secured funding, and guided the study design and manuscript development. Seung Kak Shin provided clinical insights, supported data acquisition, and contributed to critical manuscript revisions.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (Grant number: RS-2024-00360687), and the Gachon University research fund of 2025 (Grant number: GCU-202509060001) and by the Digital Medical Products Development Based on Medical Data Synthesis and AI Technologies Program (RS-2025-02305698, Development of On-Device AI Digital Medical Products Utilizing Synthetic Technology and Synthetic Data for Atypical Medical Data) funded by the Ministry of Trade, Industry & Energy (MOTIE) of Korea.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability statement

The data used to support the findings of this study are available upon request from the corresponding authors.

Abbreviations

NC-AMRI	Noncontrast abbreviated magnetic resonance imaging
MRI	Magnetic resonance imaging
HCC	Hepatocellular carcinoma
ML	Machine learning
DN	Dysplastic nodule
DNs	Dysplastic nodules
T2WI	T2-weighted imaging
DWI	Diffusion-weighted imaging
AUC	Area under the curve
PPV	Positive predictive value
NPV	Negative predictive value
ROC	Receiver operating characteristic
LR	Logistic regression
SVM	Support vector machine
RF	Random forest
XGB	Extreme gradient boosting
HGDN	High grade dysplastic nodule
LGDN	Low grade dysplastic nodule
GLCM	Gray level co-occurrence matrix
GLRLM	Gray level run length matrix
GLSZM	Gray level size zone matrix
VIF	Variance inflation factor
RFE	Recursive feature elimination
ROI	Region of interest

References

1. Kulik L and El-Serag HB. Epidemiology and management of hepatocellular carcinoma. *Gastroenterology* 2019; 156: 477–491. e471.
2. Marrero JA, Kulik LM, Sirlin CB, et al. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American Association for the Study of Liver Diseases. *Hepatology* 2018; 68: 723–750.
3. Llovet J, Mazzaferro V, Piscaglia F, et al. European association for the study of the liver electronic address: easloffice@easloffice. European association for the study of the liver EASL clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 2018; 69: 182–236.
4. Tzartzeva K, Obi J, Rich NE, et al. Surveillance imaging and alpha fetoprotein for early detection of hepatocellular carcinoma in patients with cirrhosis: a meta-analysis. *Gastroenterology* 2018; 154: 1706–1718. e1701.
5. Kim YK, Kim YK, Park HJ, et al. Noncontrast MRI with diffusion-weighted imaging as the sole imaging modality for detecting liver malignancy in patients with high risk for hepatocellular carcinoma. *Magn Reson Imaging* 2014; 32: 610–618.
6. Park HJ, Kim SY, Singal AG, et al. Abbreviated magnetic resonance imaging vs ultrasound for surveillance of

- hepatocellular carcinoma in high-risk patients. *Liver Int* 2022; 42: 2080–2092.
7. Gupta P, Soundararajan R, Patel A, et al. Abbreviated MRI for hepatocellular carcinoma screening: a systematic review and meta-analysis. *J Hepatol* 2021; 75: 108–119.
 8. Brunsing RL, Fowler KJ, Yokoo T, et al. Alternative approach of hepatocellular carcinoma surveillance: abbreviated MRI. *Hepatoma Research* 2020; 6: 59.
 9. Llovet JM, Pinyol R, Kelley RK, et al. Molecular pathogenesis and systemic therapies for hepatocellular carcinoma. *Nature Cancer* 2022; 3: 386–401.
 10. Krinsky GA, Lee VS, Theise ND, et al. Hepatocellular carcinoma and dysplastic nodules in patients with cirrhosis: prospective diagnosis with MR imaging and explanation correlation. *Radiology* 2001; 219: 445–454.
 11. Van Beers BE, Pastor CM and Hussain HK. Primovist, eovist: what to expect?. *J Hepatol* 2012; 57: 421–429.
 12. Bruix J and Sherman M. Management of hepatocellular carcinoma: an update. *Hepatology* 2011; 53: 1020–1022.
 13. Sangro B, Argemi J, Ronot M, et al. EASL clinical practice guidelines on the management of hepatocellular carcinoma. *J Hepatol* 2025; 82: 315–374.
 14. Wang R, Xu H, Chen W, et al. Gadoteric acid-enhanced MRI with a focus on LI-RADS v2018 imaging features predicts the prognosis after radiofrequency ablation in small hepatocellular carcinoma. *Front Oncol* 2023; 13: 975216.
 15. Kim DH, Choi SH, Kim SY, et al. Gadoteric acid-enhanced MRI of hepatocellular carcinoma: value of washout in transitional and hepatobiliary phases. *Radiology* 2019; 291: 651–657.
 16. Zhang Y, Sheng R, Qian X, et al. Deep learning empowered gadolinium-free contrast-enhanced abbreviated MRI for diagnosing hepatocellular carcinoma. *JHEP Reports* 2025; 7: 101392.
 17. Park SH, Choi MH, Kim B, et al. Deep learning-accelerated non-contrast abbreviated liver MRI for detecting malignant focal hepatic lesions: dual-center validation. *Korean J Radiol* 2025; 26: 333.
 18. Cheng N, Ren Y, Zhou J, et al. Deep learning-based classification of hepatocellular nodular lesions on whole-slide histopathologic images. *Gastroenterology* 2022; 162: 1948–1961. e1947.
 19. Li M, Zhang Z, Chen Z, et al. Interactive explainable deep learning model for hepatocellular carcinoma diagnosis at gadoteric acid-enhanced MRI: a retrospective, multicenter, diagnostic study. *Radiol: Imag Cancer* 2025; 7: e240332.
 20. Parekh V and Jacobs MA. Radiomics: a new application from established techniques. *Exp Rev Precis Med Drug Dev* 2016; 1: 207–226.
 21. Xu X, Zhang H-L, Liu Q-P, et al. Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma. *J Hepatol* 2019; 70: 1133–1144.
 22. Wei J, Jiang H, Gu D, et al. Radiomics in liver diseases: current progress and future opportunities. *Liver Int* 2020; 40: 2050–2063.
 23. Yang Y, Yan LF, Zhang X, et al. Optimizing texture retrieving model for multimodal MR image-based support vector machine for classifying glioma. *J Magn Reson Imaging* 2019; 49: 1263–1274.
 24. Liu Q, Jiang P, Jiang Y, et al. Prediction of aneurysm stability using a machine learning model based on PyRadiomics-derived morphological features. *Stroke* 2019; 50: 2314–2321.
 25. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012; 30: 1234–1248.
 26. Chan JY-L, Leow SMH, Bea KT, et al. Mitigating the multicollinearity problem and its machine learning approach: a review. *Mathematics* 2022; 10: 1283.
 27. An JY, Pena MA, Cunha GM, et al. Abbreviated MRI for hepatocellular carcinoma screening and surveillance. *Radiographics* 2020; 40: 1916–1931.
 28. Chou CT, Chou JM, Chang TA, et al. Differentiation between dysplastic nodule and early-stage hepatocellular carcinoma: the utility of conventional MR imaging. *World J Gastroenterol* 2013; 19: 7433–7439.
 29. Ronot M. Improving HCC surveillance with abbreviated MRI: a call to integrate and innovate? *J Hepatol* 2024; 81: 376–378.
 30. Yu MH, Kim JH, Yoon J-H, et al. Small (≤ 1 -cm) hepatocellular carcinoma: diagnostic performance and imaging features at gadoteric acid-enhanced MR imaging. *Radiology* 2014; 271: 748–760.
 31. Renzulli M, Biselli M, Brocchi S, et al. New hallmark of hepatocellular carcinoma, early hepatocellular carcinoma and high-grade dysplastic nodules on Gd-EOB-DTPA MRI in patients with cirrhosis: a new diagnostic algorithm. *Gut* 2018; 67: 1674–1682.
 32. An C, Rakhmonova G, Choi JY, et al. Liver imaging reporting and data system (LI-RADS) version 2014: understanding and application of the diagnostic algorithm. *Clin Mol Hepatol* 2016; 22: 296–307.
 33. Park HJ, Choi BI, Lee ES, et al. How to differentiate borderline hepatic nodules in hepatocarcinogenesis: emphasis on imaging diagnosis. *Liver Cancer* 2017; 6: 189–203.
 34. Oestmann PM, Wang CJ, Savic LJ, et al. Deep learning-assisted differentiation of pathologically proven atypical and typical hepatocellular carcinoma (HCC) versus non-HCC on contrast-enhanced MRI of the liver. *Eur Radiol* 2021; 31: 4981–4990.
 35. Liu X, Khalvati F, Namdar K, et al. Can machine learning radiomics provide pre-operative differentiation of combined hepatocellular cholangiocarcinoma from hepatocellular carcinoma and cholangiocarcinoma to inform optimal treatment planning? *Eur Radiol* 2021; 31: 244–255.
 36. Hu R, Li H, Hornig H, et al. Automated machine learning for differentiation of hepatocellular carcinoma from intrahepatic cholangiocarcinoma on multiphasic MRI. *Sci Rep* 2022; 12: 7924.

37. Bo Z, Song J, He Q, et al. Application of artificial intelligence radiomics in the diagnosis, treatment, and prognosis of hepatocellular carcinoma. *Comput Biol Med* 2024; 173: 108337.
38. Mohd Haniff NS, Ng KH, Kamal I, et al. Systematic review and meta-analysis on the classification metrics of machine learning algorithm based radiomics in hepatocellular carcinoma diagnosis. *Heliyon* 2024; 10: e36313.
39. Heo S, Park HJ and Lee SS. Prognostication of hepatocellular carcinoma using artificial intelligence. *Korean J Radiol* 2024; 25: 550–558.
40. Jin J, Jiang Y, Zhao YL, et al. Radiomics-based machine learning to predict the recurrence of hepatocellular carcinoma: a systematic review and meta-analysis. *Acad Radiol* 2024; 31: 467–479.