

RESEARCH

Open Access



Gd-EOB-DTPA-enhanced MRI radiomics and deep learning models to predict microvascular invasion in hepatocellular carcinoma: a multicenter study

Zhu Zhu¹, Kaiying Wu¹, Jian Lu², Sunxian Dai³, Dabo Xu¹, Wei Fang¹, Yixing Yu^{4*} and Wenhao Gu^{1*}

Abstract

Background Microvascular invasion (MVI) is an important risk factor for early postoperative recurrence of hepatocellular carcinoma (HCC). Based on gadolinium-ethoxybenzyl-diethylenetriamine pentaacetic acid (Gd-EOB-DTPA)-enhanced magnetic resonance imaging (MRI) images, we developed a novel radiomics model. It combined bi-regional features and two machine learning algorithms. The aim of this study was to validate its potential value for preoperative prediction of MVI.

Methods This retrospective study included 304 HCC patients (training cohort, 216 patients; testing cohort, 88 patients) from three hospitals. Intratumoral and peritumoral volumes of interest were delineated in arterial phase, portal venous phase, and hepatobiliary phase images. Conventional radiomics (CR) and deep learning radiomics (DLR) features were extracted based on FeAture Explorer software and the 3D ResNet-18 extractor, respectively. Clinical variables were selected using univariate and multivariate analyses. Clinical, CR, DLR, CR-DLR, and clinical-radiomics (Clin-R) models were built using support vector machines. The predictive capacity of the models was assessed by the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, and specificity.

Results The bi-regional CR-DLR model showed more gains and gave better predictive performance than the single-regional models or single-machine learning models. Its AUC, accuracy, sensitivity, and specificity were 0.844, 76.9%, 87.8%, and 69.1% in the training cohort and 0.740, 73.9%, 50%, and 84.5% in the testing cohort. Alpha-fetoprotein (odds ratio was 0.32) and maximum tumor diameter (odds ratio was 1.270) were independent predictors. The AUCs of the clinical model and the Clin-R model were 0.655 and 0.672, respectively. There was no significant difference in the AUCs between all the models ($P > 0.005$).

Conclusion Based on Gd-EOB-DTPA-enhanced MRI images, we focused on developing a radiomics model that combines bi-regional features and two machine learning algorithms (CR and DLR). The application of the new model

*Correspondence:

Yixing Yu
yuyixing@163.com
Wenhao Gu
tyyxxkgu@163.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

will provide a more accurate and non-invasive diagnostic solution for medical imaging. It will provide valuable information for clinical personalized treatment, thereby improving patient prognosis.

Clinical trial number Not applicable.

Keywords Hepatocellular carcinoma, Microvascular invasion, Conventional radiomics, Deep learning radiomics, Magnetic resonance imaging, Gd-EOB-DTPA

Introduction

Hepatocellular carcinoma (HCC) is the sixth most common malignant tumor and the fourth most lethal tumor worldwide [1]. The recurrence and metastasis rates are as high as 50–70% at 5 years after liver cancer resection and liver transplantation [2, 3]. Microvascular invasion (MVI) is defined as the invasion of tumor cells into the endothelial cell space of blood vessels [4]. It is an important risk factor for early recurrence and poor prognosis after treatment for HCC [5]. The incidence rate of MVI in HCC is approximately 50%. Patients with HCC combined with MVI exhibit a markedly inferior prognosis compared to those without MVI [6]. However, the precise identification of MVI necessitates a postoperative histopathological examination [7]. Therefore, preoperative identification of MVI is tremendously essential to guide individualized treatment decision-making [8].

Conventional radiomics (CR) employs machine learning algorithms to transform medical images into high-dimensional and qualitative features. From these, qualitative disease diagnosis models will be constructed [9]. Tian Y [10] developed CR models to predict preoperative MVI status with promising results. Deep learning is a type of algorithm that is capable of processing images and learning through an iterative process. Deep learning radiomics (DLR) deep mines potential imaging features through multitasking convolutional neural networks [11]. He X [12] developed an intratumoral DLR model based on enhanced computed tomography images, also with good MVI predictive power. Moreover, Huynh's study [13] concluded that high-performance and cross-agency generalization of models can be achieved by combining CR and DLR models. Nevertheless, the extent to which the combination of CR and DLR models can enhance predictive efficacy in the MVI domain remains to be elucidated.

Moreover, the peritumoral environment has significant potential. It may provide valuable insights into the clinical assessment of tumor aggressiveness [14]. The peritumoral CR model established by Chong [15] had demonstrated efficacy in predicting MVI. This lends support to the aforementioned perspective. However, few attempts have been made to evaluate the performance of DLR mining peritumoral features. Accordingly, an exploration into this matter is to be conducted. And whether combining intratumoral and peritumoral

features improves predictive efficacy is a question awaiting investigation.

Consequently, we will construct a predictive model. The model will integrate bi-regional features (intratumoral and peritumoral) with two machine learning algorithms (CR and DLR). It is hypothesized that this provides valuable utility for preoperative non-invasive prediction of MVI. Thus facilitating the implementation of clinically personalized treatment. This is the novelty of this study.

The objective of this study is to gain further insight into the potential of combining CR and DLR. This will result in the creation of models that are both high-performing and generalizable across institutions. Furthermore, integrating intratumoral and peritumoral data provides more comprehensive information on tumor heterogeneity. Both have the potential to offer enhanced benefits in the prediction of MVI. Thus our model can be applied to medical research and diagnosis to assist clinical treatment. The flowchart of this multicentre study is shown in Fig. 1.

Materials and methods

This study was conducted in accordance with the ethical standards set forth in the Declaration of Helsinki. It was also approved by the Medical Ethics Committees of the First People's Hospital of Taicang, the Third Affiliated Hospital of Nantong University, and the First Affiliated Hospital of Soochow University. The necessity of written informed consent was waived due to the retrospective study design. The approval numbers are as follows: 2022-ky-203, EK2023025, and 2,024,269, respectively.

Patient criteria

Patients with surgically pathology-confirmed HCC were enrolled at three hospitals from January 2016 to December 2023. Inclusion criteria: (1) HCC was confirmed by postoperative pathology; (2) the presence of MVI could be diagnosed by postoperative pathology; (3) gadolinium-ethoxybenzyl-diethylenetriamine pentaacetic acid (Gd-EOB-DTPA)-enhanced magnetic resonance imaging (MRI) was received within two weeks before surgery; (4) the images were complete and clear; and (5) the data of laboratory tests were complete within two weeks before surgery. Exclusion criteria: (1) prior local treatment of tumor before surgery, including radiofrequency

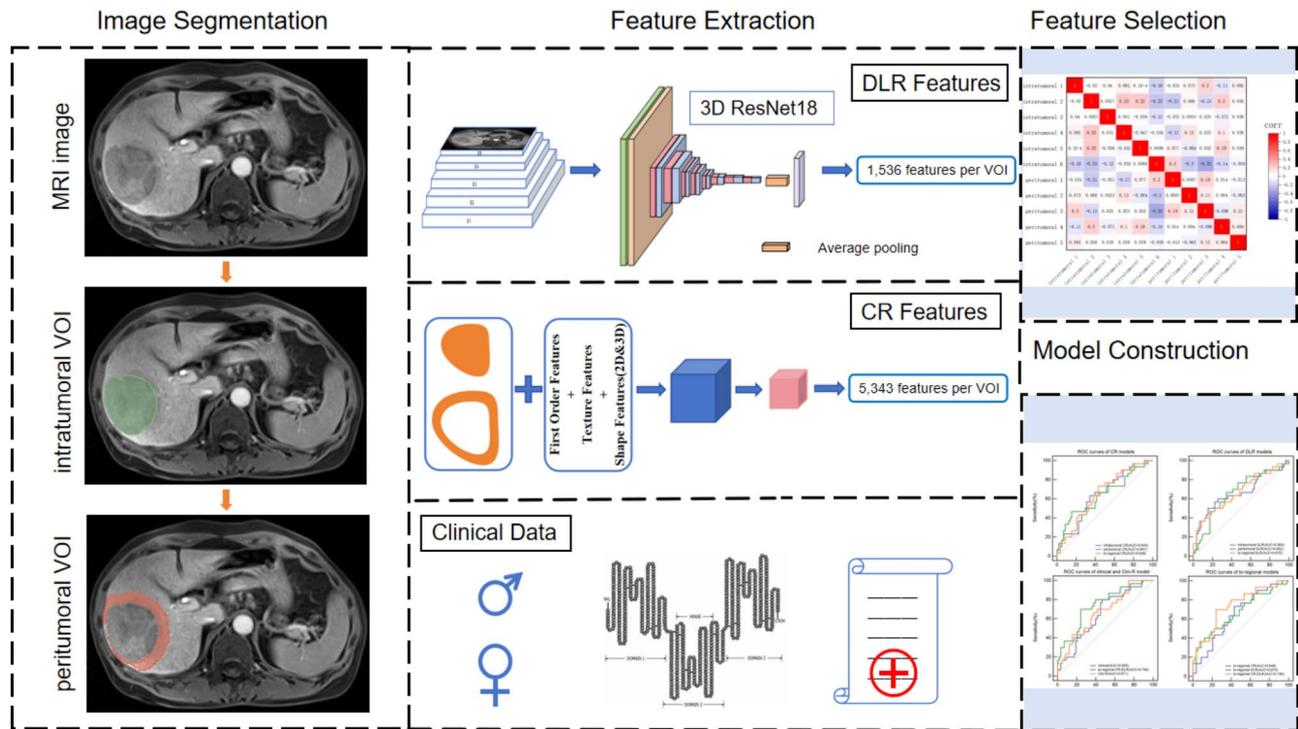


Fig. 1 Flowchart of this multicenter study. It includes image segmentation, clinical data collection, feature extraction, feature selection, and model construction

ablation, radiation therapy, or TACE; (2) invasion of great vessels such as the portal vein, hepatic vein, or inferior vena cava, or the presence of a cancerous embolus was detected on MRI images; (3) extrahepatic metastasis. The flowchart of the patients enrolled in the study is shown in Fig. 2.

Clinical data

Seven clinical data points of patients were recorded, including gender, age, maximum tumor diameter, alpha-fetoprotein (AFP), alanine aminotransferase (ALT), aspartate aminotransferase (AST), and viral hepatitis status.

MRI examination methods

A Siemens Skyra 3.0T, a Philips Medical Systems 3.0T, and a GE Discovery 750 3.0T MR scanner with a 16-channel abdominal phased array coil were used. The patients were fasted and dehydrated for 4–6 h prior to the examination and were positioned in a supine position. The images were captured in three planes: transverse, coronal, and sagittal. The scanning sequence is described in the Supplementary Methods A. The contrast agent Gd-EOB-DTPA (Bayerische Medizintechnik GmbH, Germany), measured at a volume of 0.1 mL/kg, was injected intravenously. Images of the arterial phase (AP), portal venous phase (PP), and hepatobiliary phase

(HBP) were obtained 25–30 s, 55–60 s, and 20 min after the completion of the contrast injection, respectively.

Dataset description

The dataset comprised AP, PP, and HBP images of preoperative Gd-EOB-DTPA-enhanced MRI of HCC patients (Fig. 3). The images we used were 3D images with sizes ranging from 288 × 288 × 90 to 640 × 400 × 60. Images were downloaded from the hospital imaging platforms in DICOM format and transformed into the NII.GZ format for subsequent operations.

A total of 304 HCC patients were included. The training cohort was from the First Affiliated Hospital of Soochow University, including 216 cases (170 males, 46 females, mean age 58.8 years, range 26–87 years), of which 90 MVI+ and 126 MVI- patients. The testing cohort was from the Third Affiliated Hospital of Nantong University and the First People’s Hospital of Taicang City, including 88 cases (53 males, 35 females, mean age 58.3 years, range 34–79 years), of which 30 were MVI+ and 58 were MVI-.

Implementation environment

Our operations were carried out on the computer Lenovo ThinkPad X13, which has the operating system Windows 11. We used a GPU NVIDIA GeForce RTX 4090, 24 GB of RAM and the computing platform CUDA 11.8.

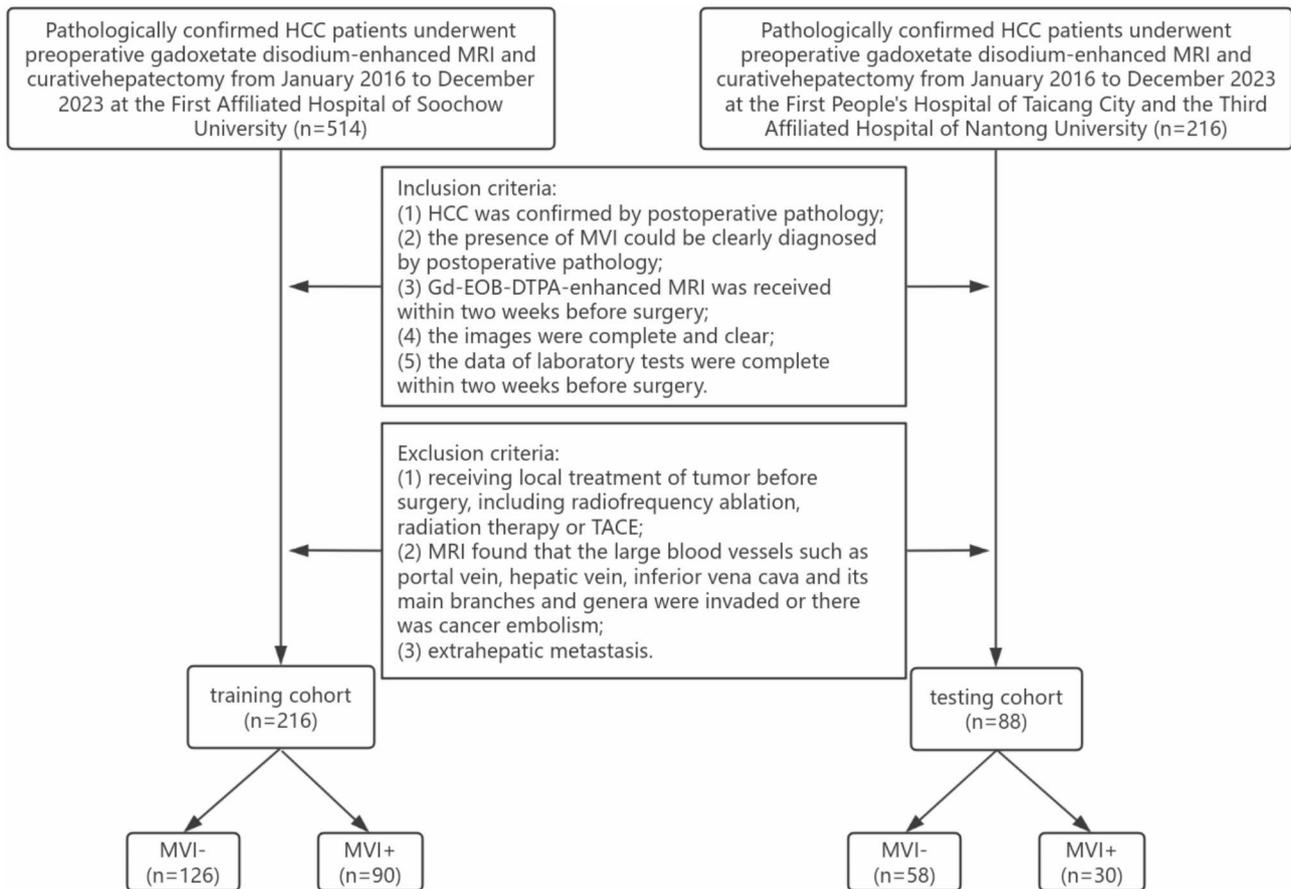


Fig. 2 Flowchart of the patients enrolled in the study

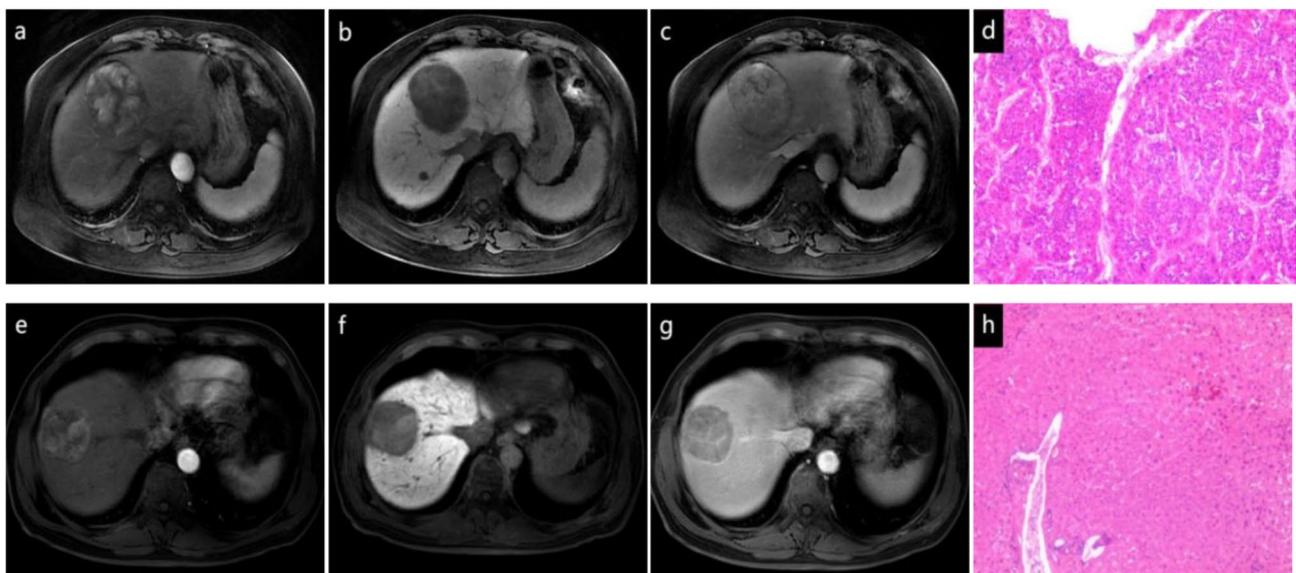


Fig. 3 Representative MRI images and relevant pathological images of two HCC patients. (a–d) show an MVI- case and (e–h) an MVI+ case. (a, e): the AP images show tumor enhancement; (b, f): the PP images show decreased tumor enhancement; (c, g): the HBP images show no enhancement of tumor and enhancement of normal liver parenchyma; (d): the pathological image shows no tumor embolus in the vascular channel (H&E, × 100); (h): the pathological image shows a tumor embolus in the vascular channel (H&E, × 100)

Image segmentation

All operations were done on 3D Slicer 4.10.2 software (<https://www.slicer.org>). Two independent radiologists with 5 and 10 years of experience, respectively, manually delineated the tumor volumes of interest (VOIs) in AP, PP, and HBP images. The intratumoral region was defined as the area within the tumor boundaries labeled by the radiologists. Then, the intratumoral region was expanded outward by 10 mm using the “Hollow” function, thus creating a peritumoral region. In instances where the VOIs extended beyond the boundaries of the liver parenchyma, the external portion was manually erased. The intra-class correlation coefficient (ICC) was employed to evaluate the reproducibility of the delineation of VOIs by two radiologists. VOIs described by radiologists with 10 years of experience were selected for subsequent radiomics analysis.

Extraction of CR features

FeAture Explorer (FAE) 0.5.2 software (<https://github.com/salan668/FAE>) is employed for the processing and extraction of CR features. Computational methods for CR features include the application of wavelet and Laplace of Gaussian (LoG) filters ($\sigma = 1.0$) to either the original or pre-processed images. Feature extraction was carried out on images that were resampled to voxel dimensions of $1 \times 1 \times 1 \text{ mm}^3$, with an intensity bin width of 5 for discretization. CR features were extracted from the intratumoral and peritumoral VOIs, including first-order statistics, shape, and texture features. Texture features included gray level co-occurrence matrix, gray level run length matrix, gray level size zone matrix, gray level dependence matrix, and neighboring gray tone difference matrix.

Extraction of DLR features

Pre-trained (Kinetics dataset-based) 3D ResNet-18 model from the “torchvision” library of the deep learning framework PyTorch v2.1.0 (<https://github.com/pytorch/pytorch>) was used for DLR features extraction. During model training, we used several well-established techniques to minimise the risk of overfitting, including data augmentation and learning rate decay. Data enhancement methods included ScaleIntensityRanged, RandRotate90d, RandFlipd, and so on. In this study, the batch size was set to 32, the initial learning rate to $1e-4$, and the learning rate decayed by 0.1 times in every 5 epochs. CrossEntropyLoss was used as the loss function. We chose the Adam optimizer, a widely used optimizer that could automatically adjust the learning rate. The fully connected layer was first fine-tuned and trained for 50–100 epochs. Once the model training was finished, the trained weights were saved for predictions. Then the fully connected layer of the model was removed, and the

part before the fully connected layer was used as a feature extractor to extract the high-dimensional feature representation from the input image. To ensure that the input image conforms to the model, resampling was performed using linear interpolation to resize the image to a uniform size of $64 \times 64 \times 64$. Finally, the feature extractor of the 3D ResNet-18 model was applied to extract DLR features from each intratumoral and peritumoral region image. As well, the activation function used was ReLU. The architecture of 3D ResNet-18 is given in the Supplementary Methods B. The details of the hyperparameters were tabulated in Supplementary Methods C.

Feature screening and model construction

To enhance the generalization of the models, we used Z-Score to normalize the features. Subsequently, the Pearson correlation coefficient (PCC) values between all features were calculated. Features with PCC values > 0.90 were excluded to prevent the potential for multicollinearity. Recursive feature elimination (RFE) was employed to identify the optimal feature set, while support vector machine (SVM) was utilized to construct the models.

Three types of VOI models were constructed: (i) intratumoral models, (ii) peritumoral models, and (iii) bi-regional models. The bi-region model is a combined intratumoral and peritumoral model. In each VOI model, based on the CR features and DLR features, we constructed CR models and DLR models. The performance of the models was quantified using the area under the receiver operating characteristic (ROC) curve (AUC). The objective was to identify the optimal CR model and DLR model. The feature sets of both models were then aggregated to construct the CR-DLR models. The AUC was further used to select the best CR-DLR model.

The clinical data of the training cohort were subjected to univariate and multivariate logistic regression analysis to identify the independent factors predicting MVI ($P < 0.05$). Subsequently, SVM was applied to construct the clinical model. Moreover, a clinical-radiomics (Clin-R) model was constructed by integrating the clinical features and the features of the best radiomics model.

Statistical analysis

SPSS 26.0 software was used for statistical analysis. Continuous variables were presented as the mean \pm standard deviation or median with interquartile range (IQR). Categorical variables were reported as frequency and proportions. The independent-sample *t* test or Mann-Whitney *U* test was performed to compare the quantitative parameters and the chi-square test to compare the qualitative features. Interobserver reproducibility of feature extraction was assessed using ICC. $\text{ICC} \geq 0.8$ indicated high consistency, 0.5–0.79 middle, and < 0.5 low.

Table 1 Clinical characteristics of patients in the training and testing cohort

Characteristic	Training cohort (n = 226)	Testing cohort (n = 88)	P-value
MVI			0.220
Positive	90	30	
Negative	126	58	
Age (years, mean ± SD)	58.79 ± 10.23	58.33 ± 9.29	0.714
Gender			0.001
Male	170	53	
Female	46	35	
Maximum tumor diameter [cm, median (IQR)]	3.35(2.20, 5.72)	3.05(1.93, 4.20)	0.055
AFP			0.167
<400ng/mL	158	71	
≥400ng/mL	58	17	
ALT			0.820
<40ng/mL	138	55	
≥40ng/mL	78	33	
AST			0.508
<40ng/mL	134	51	
≥40ng/mL	82	37	
Viral hepatitis			0.508
YES	28	9	
NO	188	79	

Table 2 Comparison of clinical characteristics between MVI+ and MVI- group in the training cohort

Characteristic	MVI+ (n = 90)	MVI- (n = 126)	P-value
Age (years, mean ± SD)	57.76 ± 10.67	59.53 ± 9.87	0.209
Gender			0.082
Male	76	94	
Female	14	32	
Maximum tumor diameter [cm, median (IQR)]	5.10(3.00, 7.63)	2.64(1.70, 4.43)	<0.001
AFP			<0.001
<400ng/mL	51	107	
≥400ng/mL	39	19	
ALT			0.666
<40ng/mL	56	82	
≥40ng/mL	34	44	
AST			0.813
<40ng/mL	55	79	
≥40ng/mL	35	47	
Viral hepatitis			0.891
YES	78	110	
NO	12	16	

AUC was applied to quantify the performance of models. Accuracy, AUC, negative predictive value, positive predictive value, sensitivity, and specificity were calculated to assess the performance of the respective models. The Delong test was employed to compare the differences between the different models. All analyses were

deemed statistically significant at P -values of less than 0.05 (two-tailed).

Results

General information

With the exception of gender ($P=0.001$), no significant differences were observed in the clinical characteristics of the training and testing cohort ($P<0.05$) (Table 1).

In the training cohort, 39 MVI+ (43%) and 19 MVI- patients (15%) had higher AFP levels ($P<0.001$). The median maximum diameter was significantly higher in MVI+ HCC (5.10, IQR: 3.00–7.63) than in MVI- HCC (2.64, IQR: 1.70–4.43) ($P<0.001$). In addition, there was no statistically significant difference between MVI+ and MVI- patients in terms of age, gender, ALT, AST, and viral hepatitis status ($P>0.05$) (Table 2).

Consistency assessment

Based on the three-phase combined MRI images, 5,343 CR features and 1,536 DLR features were extracted from each VOI, respectively. Heatmaps showed low correlation coefficients in the selected radiomics features (Fig. 4), making them suitable for establishing the models for predicting MVI.

The ICC range of the two radiologists was 0.854–0.923 for delineating VOIs, indicating excellent reproducibility of the feature extraction.

Performance of the clinical model in predicting MVI in HCC

The results of the multivariate logistic regression analysis indicated that AFP (odds ratio (OR) 0.32; 95% confidence interval (CI): 0.162–0.634) and maximum tumor diameter (OR 1.270; 95% CI: 1.135–1.420) were independent predictors of MVI in the training cohort. These predictors could be used to construct a clinical model. The AUCs of the clinical model in the training and testing cohort were 0.758 (95% CI: 0.694–0.822) and 0.655 (95% CI: 0.537–0.772), respectively.

Performance of radiomics models in predicting MVI in HCC

The radiomics models were classified into CR, DLR, and CR-DLR models by algorithm types, and into intratumoral, peritumoral, and bi-regional models by VOIs.

Of the CR models, the bi-regional model performed best in the training (AUC = 0.868; 95% CI: 81.9–91.7; sensitivity: 81.1%; specificity: 81.0%) and the testing cohort (AUC = 0.648; 95% CI: 53.0–76.7; sensitivity: 46.7%; specificity: 84.5%). The differences between it and the intratumoral model ($P=0.921$) and the peritumoral model ($P=0.931$), respectively, were not statistically significant (all $P>0.05$).

Of the DLR models, the bi-regional model had the best predictive performance. The AUC, sensitivity, and specificity were 0.800 (95% CI: 74.2–85.9), 66.7%, and 79.4%,

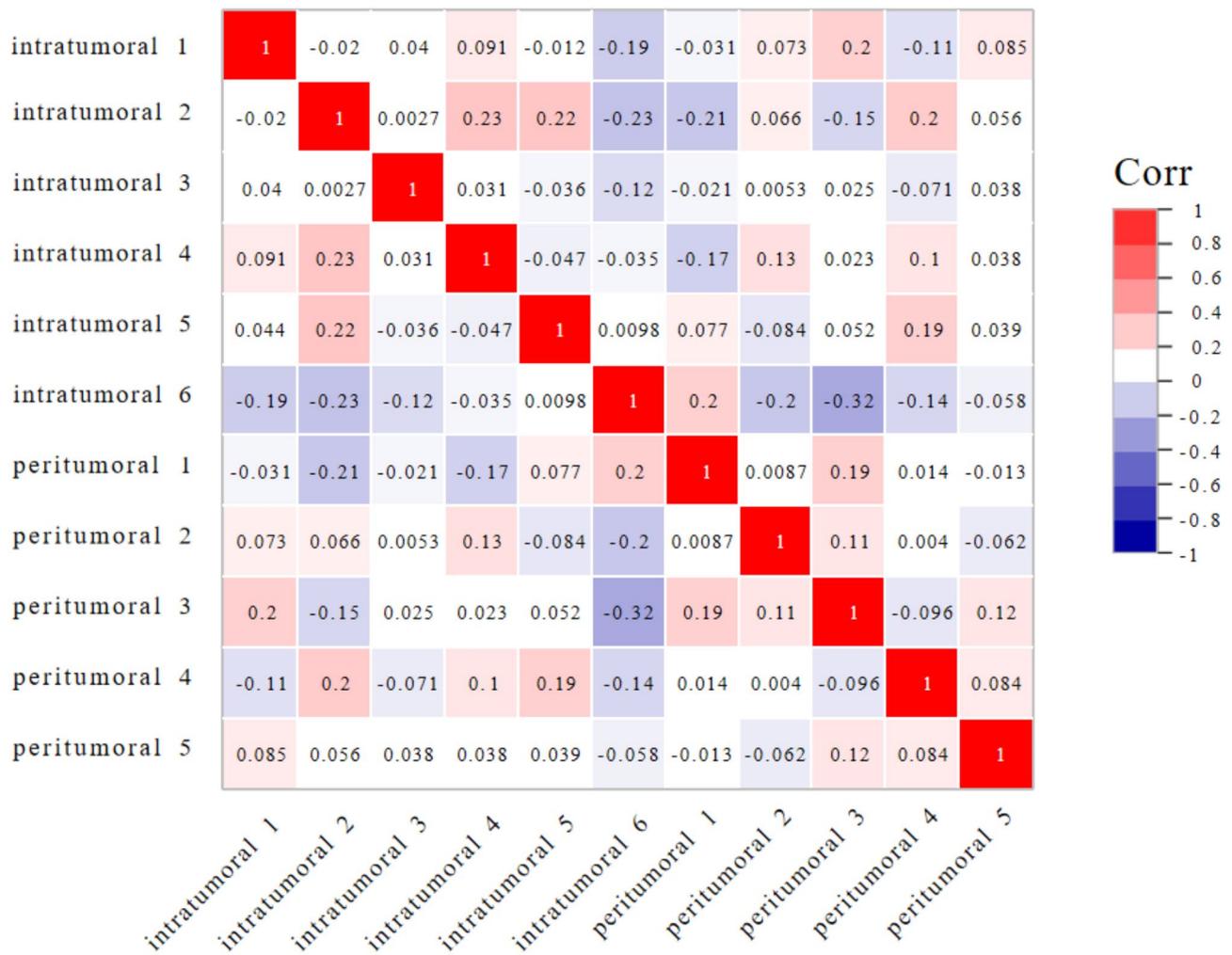


Fig. 4 Heatmap depicting correlation coefficients matrix of 11 selected features in the bi-regional CR model. The larger the value or the darker the color is, the stronger the correlation is

respectively, in the training cohort while 0.670 (95% CI: 54.5–79.4), 66.7%, and 65.5%, respectively, in the testing cohort. The differences between the bi-regional and the intratumoral model were not statistically significant ($P=0.811>0.05$), and the same was the comparison between the bi-regional and the peritumoral model ($P=0.907>0.05$). In addition to this, the intratumoral model is slightly more effective than the peritumoral model. Their AUC, sensitivity, and specificity, respectively, in the training cohort were 0.791 (95% CI: 73.1–85.0), 64.4%, and 80.1%; 0.664 (95% CI: 53.4–79.4), 73.3%, and 57.0%. Meanwhile, in the testing cohort were 0.685 (95% CI: 61.0–75.9), 54.4%, and 81.8%; 0.663 (95% CI: 53.9–78.6), 50.0%, and 84.5%. Their differences were not statistically different ($P=0.985>0.05$). To enhance interpretability, we utilized Class Activation Mapping to visualize the model outputs, aiding in our understanding of the critical regions the models emphasized when predicting MVI. As shown in Fig. 5. The red part that

gathers inward to the blue part is active, indicating that the model pays particular attention to this area.

Of the CR-DLR models, the bi-regional model had the best predictive efficacy in the training (AUC=0.844; 95% CI: 79.2–89.7; sensitivity: 87.8%; specificity: 61.1%) and the testing cohort (AUC=0.740; 95% CI: 62.9–85.1; sensitivity: 80.0%; specificity: 57.0%). The differences between it and the intratumoral model ($P=0.279$) and the peritumoral model ($P=0.062$), respectively, were not statistically significant (all $P>0.05$). In addition, the P -values for the difference between the bi-regional CR-DLR model and the intratumoral, peritumoral, and bi-regional DLR model were 0.135, 0.151, and 0.118, respectively, and none of them were statistically significant (all $P>0.05$).

In general, the bi-regional CR-DLR model demonstrated superior performance, with higher AUC and specificity, compared to all CR models (AUC=0.641–0.648; sensitivity: 46.7–80.0%; specificity: 53.5–84.5%) and DLR models (AUC=0.663–0.670; sensitivity: 50.0–73.3%;

CAM_3D_medical_image

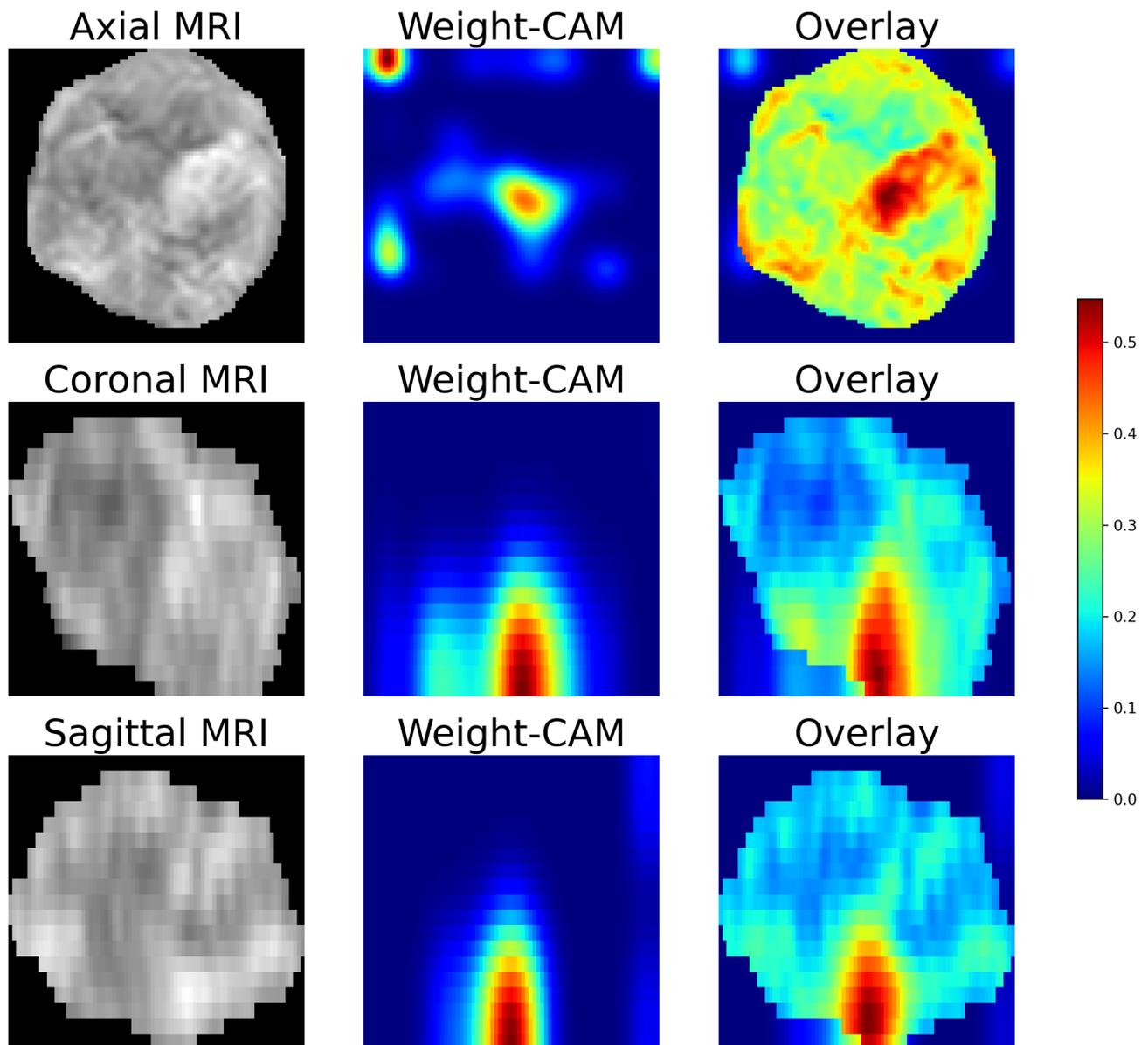


Fig. 5 The class activation maps for DLR models, which highlights important areas of the model predictions

specificity: 57.0–84.5%) in the testing cohort. However, its sensitivity was inferior to that of the majority of models, with the exception of the bi-regional CR model (AUC=0.648; 95% CI: 53.0–76.7; sensitivity: 46.7%; specificity: 84.5%) and the intratumoral CR-DLR model (AUC=0.696; 95% CI: 58.2–81.0; sensitivity: 36.7%; specificity: 91.4%). The feature set of the bi-regional CR-DLR model contained 8 (4 intratumoral, 4 peritumoral) CR features and 6 (3 intratumoral, 3 peritumoral) DLR features. Meanwhile, the features of the DLR model we plotted in a table, as detailed in Supplementary Methods D.

Performance of the Clin-R model in predicting MVI in HCC

The Clin-R model incorporated the 14 features of the bi-regional CR-DLR model with the 2 clinical features of maximum tumor diameter and AFP. The AUC, sensitivity, and specificity were 0.853 (95% CI: 80.1–90.6), 75.6%, and 84.9%, respectively, in the training cohort and 0.672 (95% CI: 55.4–79.0), 66.7%, and 62.1% in the testing cohort.

The AUC (95% CI), accuracy, sensitivity, and specificity of the training and testing cohort of the above 11 HCC models were shown in Table 3. The ROC curves were

Table 3 Model performances in the training and testing cohort

Model		Training cohort				Testing cohort			
		AUC (95% CI)	ACC (%)	SEN (%)	SPE (%)	AUC (95% CI)	ACC (%)	SEN (%)	SPE (%)
CR	intratumoral	0.816 (75.9–87.3)	76.4	72.2	79.4	0.644 (52.4–76.5)	64.8	80.0	53.5
	peritumoral	0.813 (75.5–87.2)	77.3	67.8	84.1	0.641 (51.4–76.8)	71.6	63.3	65.5
	bi-regional	0.868 (81.9–91.7)	81.0	81.1	81.0	0.648 (53.0–76.7)	62.5	46.7	84.5
DLR	intratumoral	0.791 (73.1–85.0)	73.6	64.4	80.1	0.664 (53.4–79.4)	72.7	73.3	57.0
	peritumoral	0.685 (61.0–75.9)	70.4	54.4	81.8	0.663 (53.9–78.6)	65.9	50.0	84.5
	bi-regional	0.800 (74.2–85.9)	74.1	66.7	79.4	0.670 (54.5–79.4)	72.7	66.7	65.5
CR-DLR	intratumoral	0.839 (78.5–89.1)	76.4	77.8	75.4	0.696 (58.2–81.0)	64.8	36.7	91.4
	peritumoral	0.805 (74.7–86.3)	71.8	86.7	61.1	0.657 (53.0–78.4)	72.7	80.0	57.0
	bi-regional	0.844 (79.2–89.7)	76.9	87.8	69.1	0.740 (62.9–85.1)	73.9	50.0	84.5
clinical		0.758 (69.4–82.2)	70.4	64.4	74.6	0.655 (53.7–77.2)	62.5	70.0	75.9
Clin-R		0.853 (80.1–90.6)	81.0	75.6	84.9	0.672 (55.4–79.0)	63.6	66.7	62.1

shown in Fig. 6. All differences between AUCs were not statistically significant (all $P > 0.05$).

Discussion

This study was a multicenter collaborative study. Based on the intratumoral and peritumoral regions of HCC in Gd-EOB-DTPA-enhanced MRI, we developed and validated a series of clinical, Clin-R, and radiomic (CR, DLR, and CR-DLR) models for preoperative prediction of MVI in HCC.

Our results indicated that the bi-regional models demonstrated superior predictive efficacy compared to the intratumoral or peritumoral models. The CR-DLR models had incremental predictive value over the CR or DL models. In contrast, the Clin-R model did not yield any additional predictive value compared to the radiomics models.

Radiomics models

Many studies [16, 17] have found that multi-phase MRI images could reflect a more comprehensive representation of tumor heterogeneity and vascularisation patterns than single-phase images. Consequently, multi-phase MRI images are capable of providing a more comprehensive insight into the aggressiveness of the tumor. Therefore, we chose to combine the multi-phase images of AP, PP, and HBP, to extract features for the study. The best model was the bi-regional CR-DLR model. It contained 9

AP features, 3 PP features, and 2 HBP features. This finding supported the above research point.

Feng [18] demonstrated that peritumoral features (peritumoral enhancement in AP and peritumoral low signal in HBP) could predict MVI in HCC. The peritumoral region could provide considerable information regarding tumor heterogeneity [19]. A peritumoral region of 10 mm is the optimal range recognized by more studies [15, 20], so this region was chosen for this study. Previous peritumoral studies have predominantly utilized CR [21], which was investigated in this study with the innovative use of DLR. The findings revealed that the peritumoral DLR models exhibited comparable predictive efficacy compared to the peritumoral CR models ($P > 0.005$). This indicated that DLR could be a valuable research tool for the extraction of MVI information in the peritumoral region in this study.

The results of this study demonstrated that the bi-regional models exhibited superior predictive efficacy compared to the single-regional models. This was consistent with the findings of Chong [15]. Tumor cells infiltrate outward along the microvasculature from within the tumor to form MVI. The peritumoral region is in close proximity to the tumor, forming a microenvironment critical for the growth and infiltration of tumor cells [22]. The bi-regional features, as a combination of single-regional features, can reflect tumor heterogeneity information from a holistic view. It can be reasonably

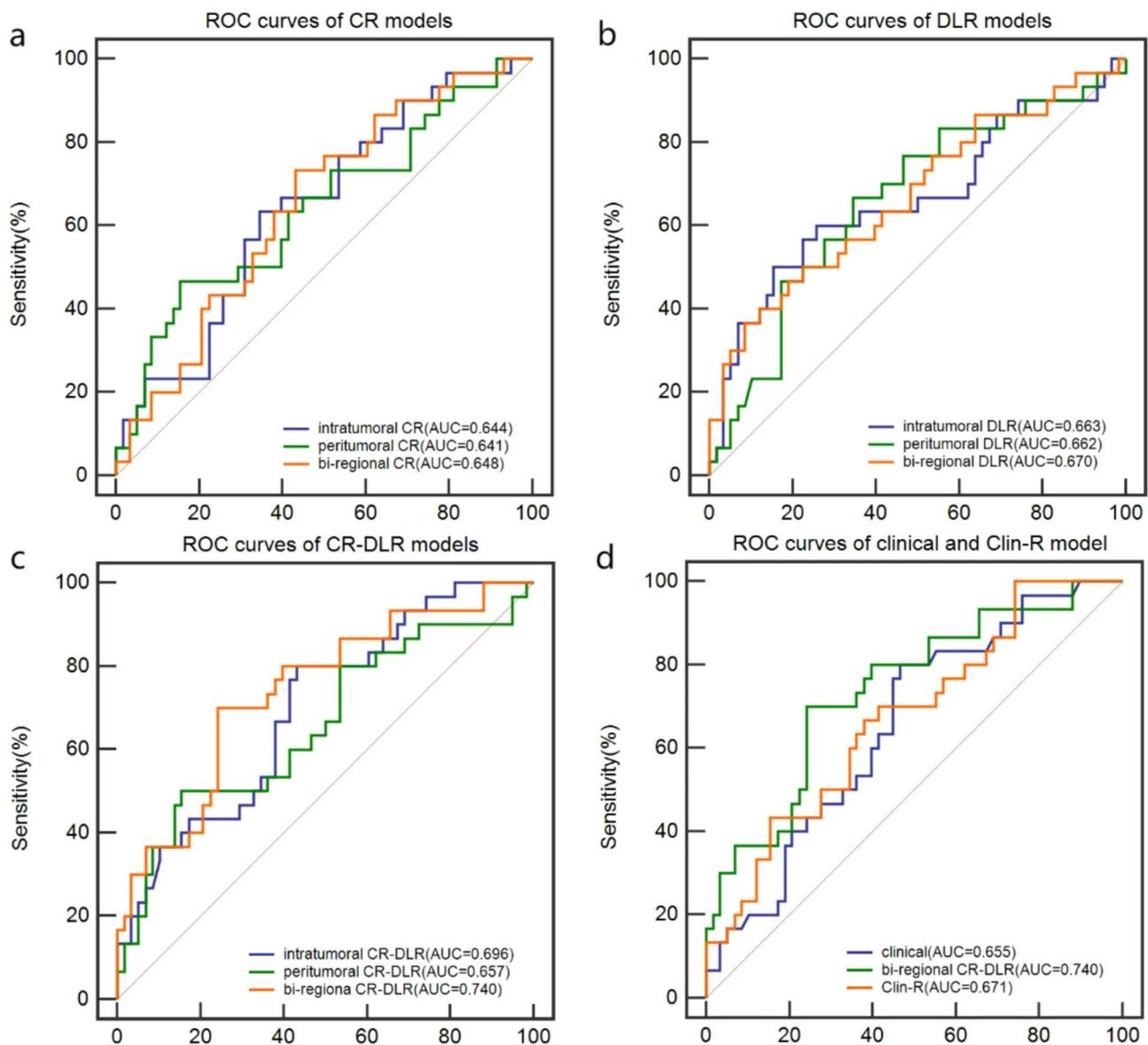


Fig. 6 Performance of prediction models based on the MVI status. (a) ROC curves of CR models; (b) ROC curves of DLR models; (c) ROC curves of CR-DLR models; (d) ROC curves of the clinical and Clin-R model

deduced that exploring them as a whole would result in superior performance.

ResNet is a commonly used deep learning network in numerous studies within the field of medicine. For example, Wang [23] developed the ResNet-based models with multimodal images for the prediction of MVI. It is reasonable to assume that ResNet is a commonly used and effective model. Gao Y [24] proposed a novel vision transformer (ViT)-based deep learning network, referred to as Dual-Style ViT. The model excelled in improving the prediction performance and interpretability of early recurrence of HCC. ViT segments the image into fixed-size patches and uses the Transformer encoder to process these patches. This allows ViT to perform well in

terms of accuracy and feature extraction. However, due to the complexity of the Transformer structure and its heavy reliance on positional information, ViT consumes more computational resources and takes longer to train and reason. In contrast, ResNet has the advantage of using a residual structure. The residual structure is built by connecting the stacked layers. Integration of subsequent input and output variables in each layer provides additional nonlinearity and reduces additional generated weights. This design makes the network deeper while avoiding the problem of vanishing gradients. Furthermore, the network structure is straightforward, facilitating training and implementation. After a comprehensive

review of the available options, we selected ResNet for this study.

CR performs high-throughput quantification of medical images to extract qualitative information about tumors and assess tumor heterogeneity [25]. DLR acquires deep image features through the utilization of computer neural network algorithms. This algorithm has the advantage of stability and reproducibility [26]. However, some important information may be missed due to the radiomics approach relying on hand-crafted features. The diagnosis of MVI is of paramount importance in the development of personalized treatment strategies for HCC. Nevertheless, the high level of diagnostic expertise in MVI is currently limited to a small number of reference centers. This may result in the vast majority of HCC patients being treated in a way that has the negative impact of lagging and mismatch. A promising solution is to employ multiple machine learning fusion models to improve the accuracy of preoperative prediction of MVI. In our study, the bi-regional CR-DLR model demonstrated superior performance compared to the CR and DLR models. The two algorithms collaborated by utilizing different feature mining methods to enhance the validity and accuracy of the prediction results. This will facilitate more effective clinical diagnosis and treatment [27]. It is anticipated that in the future, radiomics models will utilise larger datasets, employ a multimodal framework, and be applied to a broader range of projects.

Clinical and Clin-R models

Seven clinical characteristics were selected because they are most commonly employed in clinical settings and are closely associated with the diagnosis of HCC. Some studies have found that elevated AFP and maximum tumor diameter as independent risk factors for MVI in HCC [28, 29]. Both AST and ALT are biochemical surrogates that indicate hepatocellular necrosis and inflammation. The probability of HCC in patients with chronic hepatitis B virus infection is significantly higher than that of other triggers. Approximately 78% of HCC can be attributed to hepatitis B virus (53%) [30]. In comparison with other studies in the field [10, 12], the clinical characteristics presented in this study appear to be insufficient. Given the extensive duration of the study, the involvement of three hospitals, and the presence of incomplete or absent laboratory test data, we have selected these for analysis, ensuring both data completeness and accuracy. It is anticipated that in the future, the study will be expanded to include more data from more centers. This will facilitate a more comprehensive and in-depth combined study of clinical data and radiomics. The screening of clinical data of diagnostic value allows for a more targeted and effective clinical examination to be conducted.

There was significant gender variability in the training and the testing cohort. Given that the patients were sourced from multiple hospitals, we deemed it appropriate to attribute the observed gender variability to chance.

In this study, AFP and maximum tumor diameter were independent predictors in the clinical model. The predictive efficacy of both the clinical model and the Clin-R model was inferior to that of the bi-regional CR-DLR model. Many previous studies [31, 32] have reached conflicting conclusions regarding the superiority of the Clin-R model over radiomics models. Our findings contradicted this consensus. Song [33] concluded that AFP lacks sensitivity and specificity in the detection of HCC and the assessment of MVI. It was hypothesized that AFP exhibits increased instability and diminished predictive performance in comparison to the highly reproducible radiomics features. Hong [34] concluded that the prevalence of MVI was elevated in tumors with a maximum diameter of >5 cm. In contrast, the proportion of individuals with a measurement of >5 cm was relatively low (29%) in our study. We considered the generalisability and reliability of the MVI information contributed by this feature to be inadequate in this study.

It is also noteworthy that the sensitivity of the bi-regional CR-DLR model in the testing cohort was found to be lower than that of the Clin-R model. The following factors were taken into consideration: (i) Most of the patients in the testing cohort were from the Third Affiliated Hospital of Nantong University, a hospital specializing in liver disease. It had a higher detection rate for early liver cancer and small liver cancer than the hospital in the training cohort. Moreover, the training cohort had fewer MVI+ than MVI-, which has the imbalance of proportion that exists in the real medical situation. So the model built based on the training cohort showed poor generalization in the testing cohort. (ii) The sensitivity of the clinical model was observed to be higher (70%). It can be surmised that the incorporation of clinical data into the Clin-R model has led to an enhancement in sensitivity. (iii) In this study, the AUC was employed as the principal indicator of the model's predictive capacity, with specificity and sensitivity serving as secondary metrics. In the bi-regional CR-DLR model, although the sensitivity was lower, the specificity was higher (73.9%). In order to reduce the probability of misdiagnosing MVI+ as MVI- in clinical practice, greater emphasis was placed on specificity than sensitivity. This approach helps to circumvent the potential adverse outcomes associated with underdiagnosis in clinical settings. However, we have great expectations for the double improvement of sensitivity and specificity. Therefore, how to achieve overall superior performance becomes a proposition waiting to be studied and realized.

Limitations

This retrospective study still has some limitations: (1) the number of cases was limited, and the breadth of case data and multicenter external validation will be further expanded in the future; (2) patient prognostic information was not included in this study; (3) unlike the manual segmentation used in this study, the automatic segmentation technique used in [35, 36] is becoming increasingly popular. This technique has the advantages of improving the accuracy of image segmentation and making full use of massive data training. In the future, we can explore this technique more; (4) as future work, the loss functions in [37] and the augmentation methods in [38] can be integrated into the proposed model to further improve its performance.

Conclusion

In conclusion, we presented a bi-regional CR-DLR model based on Gd-EOB-DTPA-enhanced MRI. It had a good performance in preoperative non-invasive prediction of MVI in HCC. Our model can help surgeons in clinical settings to diagnose and screen for MVI, thereby guiding treatment planning. In the future, multicenter, multi-omics, and large-scale studies are needed. This will allow us to eliminate the above-mentioned limitations and validate our findings to construct more comprehensive diagnostic models.

Abbreviations

AFP	Alpha-fetoprotein
ALT	Alanine aminotransferase
AP	Arterial phase
AST	Aspartate aminotransferase
AUC	Area under the curve
CI	Confidence interval
Clin-R	Clinical-Radiomics
CR	Conventional radiomics
DLR	Deep learning radiomics
FAE	FeAture Explorer
Gd-EOB-DTPA	Gadolinium-ethoxybenzyl-diethylenetriamine pentaacetic acid
HBP	Hepatobiliary phase
HCC	Hepatocellular carcinoma
ICC	Intra-class correlation coefficient
IQR	Interquartile range
LoG	Laplacian of Gaussian
MRI	Magnetic resonance imaging
MVI	Microvascular invasion
OR	Odds ratio
PCC	Pearson correlation coefficient
PP	Portal venous phase
RFE	Recursive feature elimination
ROC	Receiver operating characteristic
SVM	Support vector machine
VOI	Volume of interest
ViT	Vision transformer

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12880-025-01646-9>.

Supplementary Material 1

Acknowledgements

We thank the patients enrolled in this study.

Author contributions

YY and WG were instrumental in the conceptualisation and design of the study. The preparation of materials and the collection of data were conducted by JL, DX, WF, and SD. The data analysis was conducted by KW and ZZ. The initial draft of the manuscript was prepared by ZZ, and all authors provided feedback on previous versions. All authors have read and approved the final manuscript.

Funding

This work was supported by the Science and Technology Program of Suzhou Municipal Health Commission (No: LCZX202126) and the Natural Science Research Project of Jiangsu Provincial Universities (No:22KJB320022).

Data availability

All data generated or analyzed during this study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The retrospective study was approved by the Medical Ethics Committees of the First People's Hospital of Taicang, the Third Affiliated Hospital of Nantong University, and the First Affiliated Hospital of Soochow University, and informed consent for participation was exempted by the ethics committee. The approval numbers are as follows: 2022-ky-203, EK2023025, and 2024269, respectively.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Radiology, The First People's Hospital of Taicang, Affiliated Hospital of Soochow University, Suzhou, Jiangsu 215400, China

²Department of Radiology, The Third Affiliated Hospital of Nantong University, The Third People's Hospital of Nantong, Nantong, Jiangsu 226000, China

³Soochow university, Suzhou, Jiangsu 215000, China

⁴Department of Radiology, The First Affiliated Hospital of Soochow University, Suzhou, Jiangsu 215000, China

Received: 22 September 2024 / Accepted: 18 March 2025

Published online: 31 March 2025

References

1. Brown ZJ, Tsilimigras DI, Ruff SM, et al. Management of hepatocellular carcinoma: A review [J]. *JAMA Surg.* 2023;158(4):410–20. <https://doi.org/10.1001/amasurg.2022.7989>.
2. Govalan R, Lauzon M, Luu M, et al. Comparison of surgical resection and systemic treatment for hepatocellular carcinoma with vascular invasion: National cancer database analysis [J]. *Liver Cancer.* 2021;10(5):407–18. <https://doi.org/10.1159/000515554>.
3. Zhang H, Huo F. Prediction of early recurrence of HCC after hepatectomy by contrast-enhanced ultrasound-based deep learning radiomics [J]. *Front Oncol.* 2022;12:930458. <https://doi.org/10.3389/fonc.2022.930458>.
4. Zhang X, Li J, Shen F, et al. Significance of presence of microvascular invasion in specimens obtained after surgical treatment of hepatocellular carcinoma [J]. *J Gastroenterol Hepatol.* 2018;33(2):347–54. <https://doi.org/10.1111/jgh.13843>.

5. Hwang YJ, Bae JS, Lee Y, et al. Classification of microvascular invasion of hepatocellular carcinoma: correlation with prognosis and magnetic resonance imaging [J]. *Clin Mol Hepatol*. 2023;29(3):733–46. <https://doi.org/10.3350/cmh.2023.0034>.
6. Li SH, Mei J, Cheng Y, et al. Postoperative adjuvant hepatic arterial infusion chemotherapy with FOLFOX in hepatocellular carcinoma with microvascular invasion: A multicenter, phase III, randomized study [J]. *J Clin Oncol*. 2022;41(10):1898–908. <https://doi.org/10.1200/JCO.22.01142>.
7. Court CM, Harlander-Locke MP, Markovic D, et al. Determination of hepatocellular carcinoma grade by needle biopsy is unreliable for liver transplant candidate selection [J]. *Liver Transpl*. 2017;23(9):1123–32. <https://doi.org/10.1002/lt.24811>.
8. Lv K, Cao X, Du P, et al. Radiomics for the detection of microvascular invasion in hepatocellular carcinoma [J]. *World J Gastroenterol*. 2022;28(20):2176–83. <https://doi.org/10.3748/wjg.v28.i20.2176>.
9. Aoude LG, Wong BZY, Bonazzi VF, et al. Radiomics biomarkers correlate with CD8 expression and predict immune signatures in melanoma patients [J]. *Mol Cancer Res*. 2021;19(6):950–6. <https://doi.org/10.1158/1541-7786.MCR-20-1038>.
10. Tian Y, Hua H, Peng Q, et al. Preoperative evaluation of Gd-EOB-DTPA-Enhanced MRI Radiomics-Based nomogram in small solitary hepatocellular carcinoma (≤ 3 cm) with microvascular invasion: A Two-Center study [J]. *J Magn Reson Imaging*. 2022;56(5):1459–72. <https://doi.org/10.1002/jmri.28157>.
11. Schmidhuber J. Deep learning in neural networks: an overview [J]. *Neural Netw*. 2015;61:85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>.
12. He X, Xu Y, Zhou C, et al. Prediction of microvascular invasion and pathological differentiation of hepatocellular carcinoma based on a deep learning model [J]. *Eur J Radiol*. 2024;172:111348. <https://doi.org/10.1016/j.ejrad.2024.111348>.
13. Huynh BN, Groendahl AR, Tomic O, et al. Head and neck cancer treatment outcome prediction: a comparison between machine learning with conventional radiomics features and deep learning radiomics [J]. *Front Med*. 2023;10. <https://doi.org/10.3389/fmed.2023.1217037>.
14. Li N, Wan X, Zhang H, et al. Tumor and peritumor radiomics analysis based on contrast-enhanced CT for predicting early and late recurrence of hepatocellular carcinoma after liver resection [J]. *BMC Cancer*. 2022;22(1):664. <https://doi.org/10.1186/s12885-022-09743-6>.
15. Chong HH, Yang L, Sheng RF, et al. Multi-scale and multi-parametric radiomics of gadoxetate disodium-enhanced MRI predicts microvascular invasion and outcome in patients with solitary hepatocellular carcinoma ≤ 5 cm [J]. *Eur Radiol*. 2021;31(7):4824–38. <https://doi.org/10.1007/s00330-020-07601-2>.
16. Gao L, Xiong M, Chen X, et al. Multi-Region radiomic analysis based on Multi-Sequence MRI can preoperatively predict microvascular invasion in hepatocellular carcinoma [J]. *Front Oncol*. 2022;12:818681. <https://doi.org/10.3389/fonc.2022.818681>.
17. Liu WM, Zhao XY, Gu MT, et al. Radiomics of preoperative Multi-Sequence magnetic resonance imaging can improve the predictive performance of microvascular invasion in hepatocellular carcinoma [J]. *World J Oncol*. 2023;15:58–71. <https://doi.org/10.14740/wjon1731>.
18. Feng ST, Jia Y, Liao B, et al. Preoperative prediction of microvascular invasion in hepatocellular cancer: a radiomics model using Gd-EOB-DTPA-enhanced MRI [J]. *Eur Radiol*. 2019;29(9):4648–59. <https://doi.org/10.1007/s00330-018-5935-8>.
19. Mao N, Shi Y, Lian C, et al. Intratumoral and peritumoral radiomics for preoperative prediction of neoadjuvant chemotherapy effect in breast cancer based on contrast-enhanced spectral mammography [J]. *Eur Radiol*. 2022;32(5):3207–19. <https://doi.org/10.1007/s00330-021-08414-7>.
20. Zhou KQ, Sun YF, Cheng JW, et al. Effect of surgical margin on recurrence based on preoperative Circulating tumor cell status in hepatocellular carcinoma [J]. *eBioMedicine*. 2020;62:103107. <https://doi.org/10.1016/j.ebiom.2020.103107>.
21. Yang Y, Fan W, Gu T, et al. Radiomic features of Multi-ROI and Multi-Phase MRI for the prediction of microvascular invasion in solitary hepatocellular carcinoma [J]. *Front Oncol*. 2021;11:756216. <https://doi.org/10.3389/fonc.2021.756216>.
22. Zhang R, Xu L, Wen X, et al. A nomogram based on bi-regional radiomics features from multimodal magnetic resonance imaging for preoperative prediction of microvascular invasion in hepatocellular carcinoma [J]. *Quant Imaging Med Surg*. 2019;9(9):1503–15. <https://doi.org/10.21037/qims.2019.09.07>.
23. Wang F, Chen Q, Chen Y, et al. A novel multimodal deep learning model for preoperative prediction of microvascular invasion and outcome in hepatocellular carcinoma [J]. *Eur J Surg Oncol*. 2023;49(1):156–64. <https://doi.org/10.1016/j.ejso.2022.08.036>.
24. Gao Y, Yang X, Li H, et al. A knowledge-enhanced interpretable network for early recurrence prediction of hepatocellular carcinoma via multi-phase CT imaging [J]. *Int J Med Informatics*. 2024;189:105509. <https://doi.org/10.1016/j.ijmedinf.2024.105509>.
25. Tortora M, Gemini L, Scaravilli A, et al. Radiomics applications in head and neck tumor imaging: A narrative review [J]. *Cancers*. 2023;15(4):1174. <https://doi.org/10.3390/cancers15041174>.
26. Wang Y, Lombardo E, Avanzo M, et al. Deep learning based time-to-event analysis with PET, CT and joint PET/CT for head and neck cancer prognosis [J]. *Comput Methods Programs Biomed*. 2022;222:106948. <https://doi.org/10.1016/j.cmpb.2022.106948>.
27. Park HJ, Park B, Lee SS. Radiomics and deep learning: hepatic applications [J]. *Korean J Radiol*. 2020;21(4):387–401. <https://doi.org/10.3348/kjr.2019.0752>.
28. Deng Y, Yang D, Tan X, et al. Preoperative evaluation of microvascular invasion in hepatocellular carcinoma with a radiological feature-based nomogram: a bi-centred study [J]. *BMC Med Imaging*. 2024;24(1):29. <https://doi.org/10.1186/s12880-024-01206-7>.
29. Fujita N, Ushijima Y, Ishimatsu K, et al. Multiparametric assessment of microvascular invasion in hepatocellular carcinoma using Gadoteric acid-enhanced MRI [J]. *Abdom Radiol*. 2024;49(5):1467–78. <https://doi.org/10.1007/s00261-023-04179-3>.
30. El-Serag HB. Epidemiology of viral hepatitis and hepatocellular carcinoma [J]. *Gastroenterology*. 2012;142(6):1264–73. e1.
31. Zhang K, Xie SS, Li WC, et al. Prediction of microvascular invasion in HCC by a scoring model combining Gd-EOB-DTPA MRI and biochemical indicators [J]. *Eur Radiol*. 2022;32(6):4186–97. <https://doi.org/10.1007/s00330-021-08502-8>.
32. Zheng R, Zhang X, Liu B, et al. Comparison of non-radiomics imaging features and radiomics models based on contrast-enhanced ultrasound and Gd-EOB-DTPA-enhanced MRI for predicting microvascular invasion in hepatocellular carcinoma within 5 cm [J]. *Eur Radiol*. 2023;33(9):6462–72. <https://doi.org/10.1007/s00330-023-09789-5>.
33. Song PP, Xia JF, Inagaki Y, et al. Controversies regarding and perspectives on clinical utility of biomarkers in hepatocellular carcinoma [J]. *World J Gastroenterol*. 2016;22(1):262–74. <https://doi.org/10.3748/wjg.v22.i1.262>.
34. Hong SB, Choi SH, Kim SY, et al. MRI features for predicting microvascular invasion of hepatocellular carcinoma: a systematic review and meta-analysis [J]. *Liver Cancer*. 2021;10(2):94–106. <https://doi.org/10.1159/000513704>.
35. You C, Yang J, Chapiro J, et al. Unsupervised Wasserstein distance guided domain adaptation for 3D multi-domain liver segmentation [J]. *Springer International Publishing*; 2020. https://doi.org/10.1007/978-3-030-61166-8_17.
36. You C, Zhao R, Staib LH et al. Momentum contrastive voxel-wise representation learning for semi-supervised volumetric medical image segmentation [J]. *Medical image computing and computer-assisted intervention: miccai international conference on medical image computing and computer-assisted intervention*. 2021;13434:639–52. <https://doi.org/10.48550/arXiv.2105.07059>.
37. Goceri E. Polyp segmentation using a hybrid vision transformer and a hybrid loss function [J]. *J Imaging Inf Med*. 2024;37(2):851–63. <https://doi.org/10.1007/s10278-023-00954-2>.
38. Goceri E. Comparison of the impacts of dermoscopy image augmentation methods on skin cancer classification and a new augmentation method with wavelet packets [J]. *Int J Imaging Syst Technol*. 2023;33(5):1727–44. <https://doi.org/10.1002/ima.22890>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.