

A Transformer-Based Deep Learning Model for predicting Early Recurrence in Hepatocellular Carcinoma After Hepatectomy Using Intravoxel Incoherent Motion Images

Hongxiang Li^{1,*}, Zehong Qiu^{2,*}, Jing Zhang^{1,*}, Yang Chen², Baoer Liu¹, Zeyu Zheng¹, Xiang Qin¹, Chenggong Yan¹, Wu Zhou², Yikai Xu¹

¹Department of Medical Imaging Center, Nanfang Hospital, Southern Medical University, Guangzhou, People's Republic of China; ²School of Medical Information Engineering, Guangzhou University of Chinese Medicine, Guangzhou, People's Republic of China

*These authors contributed equally to this work

Correspondence: Yikai Xu, Department of Medical Imaging Center, Nanfang Hospital, Southern Medical University, Guangzhou, People's Republic of China, Email yikai.xu@163.com; Wu Zhou, School of Medical Information Engineering, Guangzhou University of Chinese Medicine, Guangzhou, People's Republic of China, Email zhouwu@gzucm.edu.cn

Background: This study aimed to develop and validate a transformer framework-based deep learning (DL) network using intravoxel incoherent motion (IVIM) diffusion-weighted imaging (DWI) to predict early recurrence in hepatocellular carcinoma (HCC).

Materials and Methods: This retrospective study included 122 patients with HCC who underwent magnetic resonance imaging examination, including an IVIM-DWI sequence with nine b-values, before resection. These were divided into training (n=85) and test (n=37) sets. A vision transformer (ViT) framework-based DL was developed to predict early recurrence in HCC. Deep features were extracted from nine b-value DWI images and IVIM parametric maps and fused to construct the fused DL (ViT-fDL) prediction model. A clinical model was constructed using multivariate logistic regression analysis. A combined model was constructed using deep features from the ViT-fDL model and clinical independent features. The performances of the models were evaluated by discrimination, calibration, and clinical applicability.

Results: Among 122 patients (108 males, 14 females; mean age, 51.0 ± 11.9 years), 49 (40.1%) experienced early recurrence. The respective areas under the curve for the training and test sets were 0.755 (95% Confidence interval (CI), 0.650–0.842) and 0.764 (95% CI, 0.596–0.887) using the clinical model, 0.968 (95% CI, 0.905–0.994) and 0.815 (95% CI, 0.653–0.923) using the ViT-fDL model, and 0.991 (95% CI, 0.940–1.000) and 0.821 (95% CI, 0.660–0.927) using the combined model.

Conclusion: The ViT-fDL model based on IVIM can be useful for preoperative prediction early recurrence in HCC. The combined model was a more effective and precise prediction tool than other models, promising to guide individualized postoperative monitoring.

Keywords: hepatocellular carcinoma, early recurrence, deep learning, vision transformer

Introduction

Hepatocellular carcinoma (HCC) is the most common primary liver cancer and a major cause of cancer-related death worldwide.¹ Resection is considered the first-line treatment for HCC, particularly for patients with early-stage HCC and well-preserved liver function.² However, the prognosis of HCC remains poor, and its recurrence rates after surgery reach 80%.³ HCC recurrence is generally classified as early (within two years) or late (>2 years) recurrence.⁴ Early recurrence of HCC is related to intrahepatic metastases and is usually associated with a worse prognosis.^{5,6} Therefore, identifying patients at a high risk of recurrence after surgery might allow clinicians to provide appropriate surveillance to detect HCC

recurrence at its earliest stage. It could also help identify potential candidates for clinical trials of adjuvant therapy, which could improve individualized management and gain prognostic benefits.^{4,5,7}

Some specific factors were associated with early recurrence of HCC, including large tumor size, multiple tumors, lower true diffusion coefficient (D) values, lower apparent diffusion coefficient (ADC) values, microvascular invasion (MVI), and poor histologic differentiation.^{3,4,8–10} Deep learning (DL) can automatically extract and learn deep features from original images to comprehensively quantify tumor heterogeneity. Numerous studies reporting the application of DL have demonstrated remarkable performance in enhancing the diagnosis and predicting the prognosis for HCC, offering promising clinical advantages.^{11–13} Recently, DL, which uses a vision transformer (ViT) framework, has emerged as a promising technique for many medical image processing tasks, replacing convolutional neural networks (CNNs) due to its superior performance and robustness.^{14–16} By incorporating the self-attention and multi-attention mechanisms, the ViT framework gains a deep understanding of the weight of importance of each input data point and effectively captures intricate spatial relationships, patterns across different levels, and locations of multiple sequences within the input data that are usually neglected in CNNs.^{17–19} Moreover, some previous studies about 2D or 2.5D tumor analysis resulted in less effective capture of the morphological and spatial information of the entire tumor.^{20,21} Given the high heterogeneity of HCC, we used 3D analysis based on whole tumor delineation that can comprehensively evaluate the information about the entire tumor.

However, most DL studies for HCC prognosis were based on routine magnetic resonance imaging (MRI) sequences,^{6,8,22} while advanced techniques such as intravoxel incoherent motion (IVIM) have seldom been employed. Notably, a previous study extracted the deep features of multiple b-value and IVIM diffusion-weighted imaging (DWI) images using traditional deep CNNs to predict MVI in HCC.²³ The IVIM technology is derived from DWI with multiple b-value images without using contrast agents, reflecting tissue diffusion and microcapillary perfusion separately. Traditional quantitative IVIM parameters have showed limited diagnostic efficacy in evaluating HCC prognosis,⁹ often neglected high-dimension information hidden behind the IVIM images, and this high-dimension information can be further effectively captured through ViT-DL network. Therefore, this study aimed to develop and validate a ViT-DL network based on IVIM-DWI for accurately predicting early recurrence of HCC.

Materials and Methods

Study Population

This retrospective study was approved by the Ethics Committee of Nanfang Hospital of Southern Medical University (approval number: NFEC-202305-Y10). The study was conducted in compliance with the Declaration of Helsinki. Informed consent was waived by the Ethics Committee due to the retrospective design of anonymized patient data. Between December 2015 and July 2022, 350 consecutive patients with suspected or confirmed malignant hepatic lesions underwent preoperative gadoxetate disodium-enhanced liver MRI examination with a nine b-values DWI sequence at our institution. The inclusion criteria were as follows: (1) patients with postoperative pathological confirmation of HCC; (2) preoperative gadoxetate disodium-enhanced abdominal MRI, including a nine b-values DWI sequence, performed within one month before the operation; (3) tumor without macrovascular invasion on preoperative MRI; (4) no history of preoperative treatment for HCC, such as radiofrequency ablation, transarterial chemoembolization, or other targeted therapies. The exclusion criteria were as follows: (1) signs of extrahepatic metastases on imaging (n=6); (2) history of other malignant tumors (n=1); (3) poor image quality on MRI scans (n=13); (4) incomplete clinical, radiological, pathological, or follow-up data (n=95); (5) small HCCs less than 1 cm in diameter or HCC showing iso- or hypointensity on DWI (n=3). Finally, 122 patients were included in this study. These patients were randomly divided into training and test sets at a 7:3 ratio. The patient flowchart is shown in [Figure 1](#).

Clinical Characteristics and Follow-Up

The clinical information included age, sex, alpha-fetoprotein (AFP) level, alanine transaminase (ALT), aspartate aminotransferase (AST), total bilirubin, direct bilirubin, indirect bilirubin, albumin, platelet count, Child-Pugh grade, Barcelona Clinic Liver Cancer Stage, tumor size and origin of liver disease.

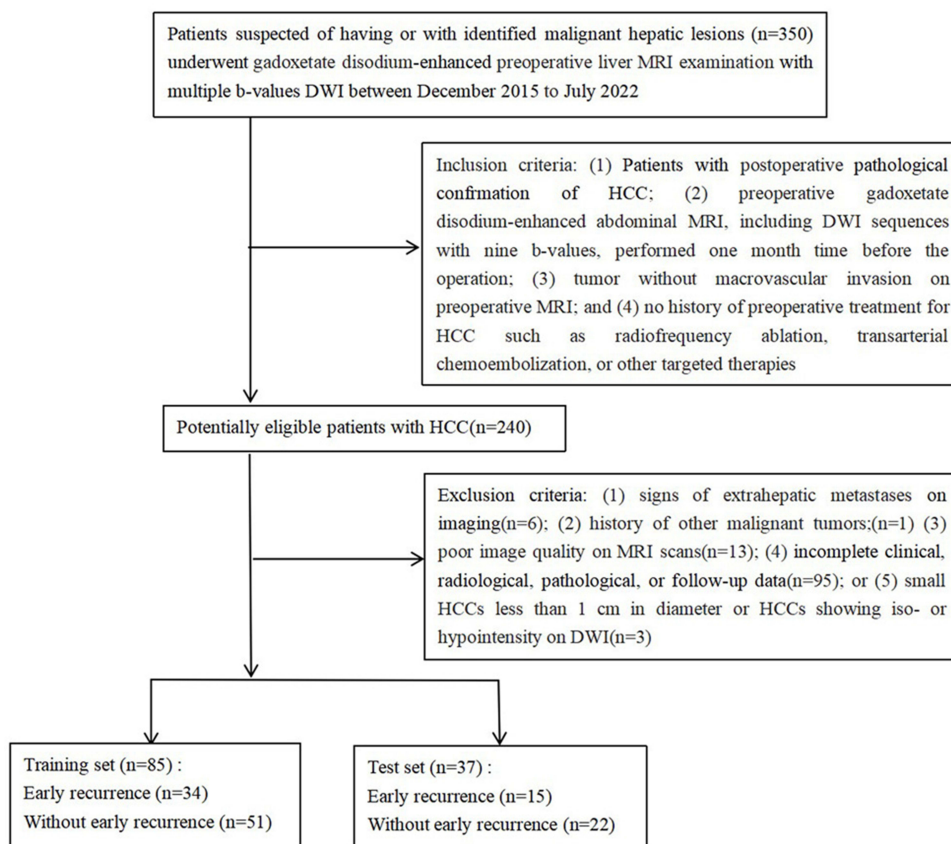


Figure 1 The patient flowchart.

All patients were followed up for at least two years after surgical resection. Patients were screened for tumor recurrence through serum AFP level, ultrasonography, and/or contrast-enhanced CT or MRI examination every three months in the first year and every 3–6 months after that. The censored follow-up date was August 1, 2024.

Early recurrence was defined as one or more of the following events occurring within two years after hepatectomy: (a) presence of new hepatic lesions with typical imaging findings of HCC; (b) atypical imaging findings with biopsy or re-postoperative histopathology confirming HCC diagnosis, or if postoperative transarterial chemoembolization indicated tumor staining; (c) extrahepatic metastases confirmed by typical imaging features or histopathology. No early recurrence was considered if recurrence occurred after more than two years or if no recurrence occurred during follow-up.

MRI

The MRI scans were obtained using two 3.0 T MRI systems (Achieva and Ingenia, Philips Healthcare, the Netherlands). The IVIM-DWI sequences with nine b-values ($b = 0, 10, 20, 40, 80, 200, 400, 600, \text{ and } 1000 \text{ s/mm}^2$) were performed using a respiratory trigger with a single-shot echo-planar sequence. MRI sequences and scan parameters are listed in the MRI protocols in the [Supplementary Materials 1](#) and [Supplement Table S1](#).

Post-Processing IVIM Maps and Parameter Analysis

The nine b-value DWI images were transferred to a workstation (version AW4.6, GE Healthcare) with Functool software to generate post-processing IVIM parameter maps (including ADC, D, D* and f maps).

All nine b-value images were used as input data. IVIM data were analyzed using in-house software based on 64-bit MATLAB 2014b (Math Works, Natick, MA), automatically generating the quantitative IVIM parameters, including ADC, D, D* and f values. Method for quantitative IVIM parameters analysis were provided in [Supplementary Materials 2](#).

Image Pre-Processing and Segmentation

The patient's nine b-value images and IVIM parametric maps were utilized for further analysis. Preprocessing method of original images about nine b-value images and IVIM parametric maps were recorded in [Supplementary Materials 3](#).

Subsequently, above mentioned two radiologists with six and ten years of experience manually delineated the VOIs in the nine b-value DWI images to encompass the entire tumor using ITK-SNAP software. Meanwhile, the delineation of VOIs on DWI ($b=1000\text{s/mm}^2$) automatically transferred to four IVIM parametric maps to gain tumor VOIs on the corresponding four parametric maps. The VOIs were confirmed or corrected by the above mentioned senior radiologist with more than 15 years of work experience. Then, the tumor areas were extracted from the original images using Python. And the detailed information about tumor areas processing were provided in [Supplementary Materials 4](#). The processes of image preprocessing and tumor segmentation were shown in [Figure 2](#).

Histopathological Analysis

The pathological reports of all included patients were retrospectively reviewed. Pathology data included the background liver parenchymal tissue, Edmondson-Steiner grade, and the presence of MVI. MVI was defined as tumor tissue within a vascular space lined by endothelial cells and visible only by microscopy. All pathology examinations were performed

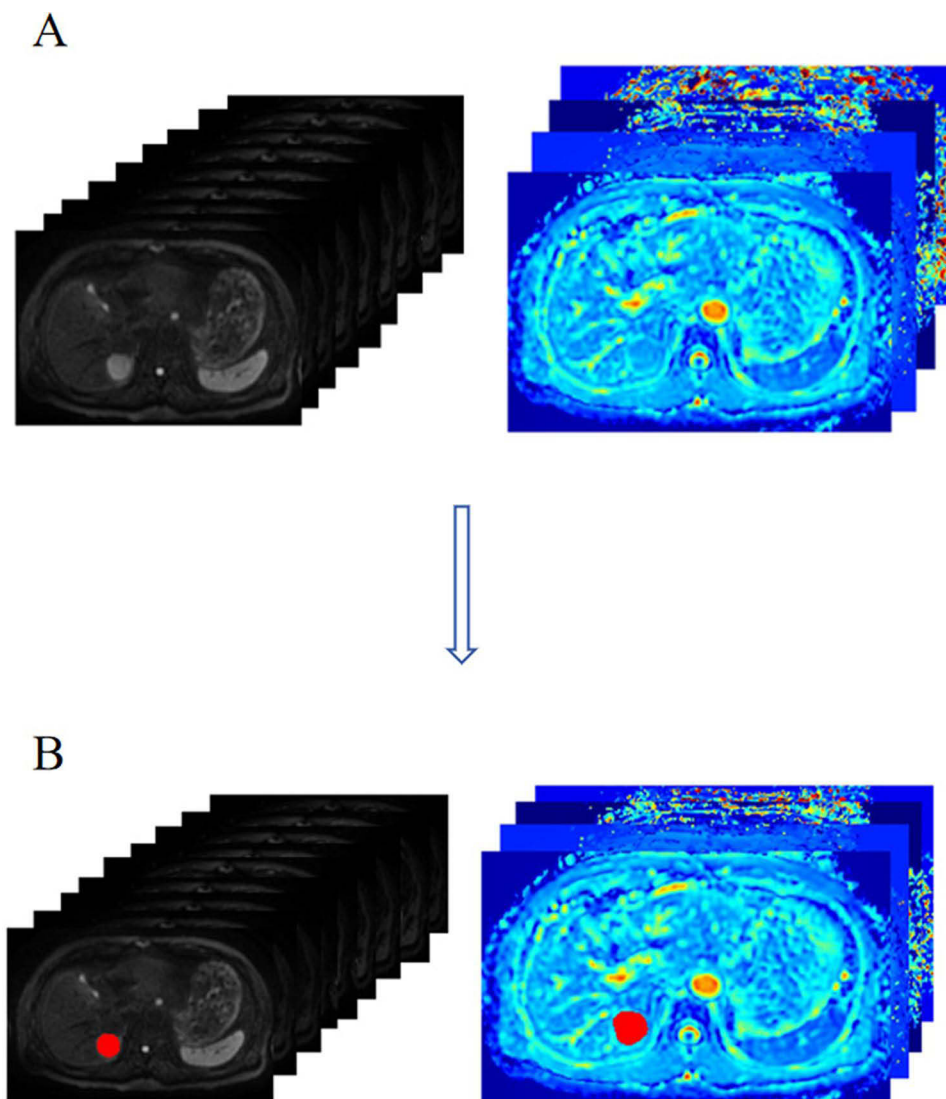


Figure 2 Images preprocessing (A) and tumor segmentation (B).

by two pathologists with more than ten years of experience in liver pathology who were blinded to the radiological and clinical findings.

Vision Transformer DL Network Construction

We used the ViT-DL model to predict early recurrence of HCC based on the IVIM-DWI images. The input dimension was the $64 \times 64 \times 64 \times 9$ IVIM-DWI images generated by channel stacking, which consisted of nine b-value images from the axial view VOIs of the tumor.

We also evaluated the ViT-DL model based on IVIM parametric maps. The input images were normalized tumor areas extracted from four IVIM parametric maps. The deep features used in our combined model were obtained from the last transformer block. The model had 100 training iterations, a batch size of 25, and an initial learning rate of $6e-6$. We chose AdamW optimizer with L2 regularization, set to $3e-4$ in server graphics card [NVIDIA GeForce RTX 3090 (24 GB) GPU]. Details related to the ViT-DL model development are described in [Figure 3](#) and [Supplementary Materials 5](#).

Feature Extraction and DL Model Development

We used the ViT-DL model to extract two feature sets, ViT-bDL and ViT-mDL, from the tumors in nine b-value images and IVIM parametric maps, respectively. A fusion feature set (ViT-fDL) was generated by performing concat and layer normalization on the ViT-bDL and ViT-mDL feature sets. We developed three DL models: (1) using only the ViT-bDL features (ViT-bDL model); (2) using only the ViT-mDL features (ViT-mDL model); (3) using the ViT-fDL features (ViT-fDL model). To prevent model overfitting, we first performed feature dimension reduction using Pearson's coefficient analysis with a threshold of 0.8 and then analyzed the screened features by ANOVA to obtain the most relevant deep features. We also used the five-fold cross-validation method in our DL models in the training set to enhance model robustness. The ViT-fDL model was based on the ten most relevant deep features, and its output probability was used to predict early recurrence, which represented the likelihood of patients with HCC exhibiting early recurrence after surgery.

Construction of the Clinical and Combined Models

Statistically significant clinical predictors were identified by multivariate logistic regression analysis and the calculated odds ratios (ORs) using the training set. Subsequently, a clinical model was developed as a baseline for comparison. A combined model integrating the deep features from the ViT-fDL model and independent clinical predictors was then constructed to further enhance the model's performance in predicting early recurrence of HCC. Details of the clinical, ViT-fDL, and combined model development and training processes are described in [Figure 3](#).

Model Performance Assessment and Interpretability

The diagnostic performance of the models in predicting early recurrence of HCC was assessed by the area under the curve (AUC) of the receiver operating characteristic curve (ROC) with 95% confidence interval (CI). The corresponding sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were also calculated. Model fitting was assessed by using calibration curves using 1000 resampling bootstraps and the Hosmer–Lemeshow test. Decision curve analysis was implemented to estimate the clinical utility of the models by quantifying their net benefits under various threshold probabilities. Gradient-weighted Class Activation Mapping (Grad-CAM) were used to visualize the location and distribution of decision feature information captured by the DL model in predicting early recurrence of HCC.

Statistical Analysis

Statistical analysis was performed using IBM SPSS Statistics for Windows, Version 20.0 (IBM Corporation, Armonk, NY, USA), MedCalc, Version 18.2 (MedCalc Software Ltd, Mariakerke, Belgium), and R (version 4.4.1, <http://www.r-project.org>). Continuous parameters were evaluated using an independent sample *t*-test (normally distributed) or Mann–Whitney *U*-test (non-normally distributed). Categorical variables were compared by the χ^2 or Fisher's exact test. Inter-observer and inter-scanner variability in the IVIM parameters were analyzed by calculating the interclass correlation coefficient (ICC), ICC value of >0.75 was considered to represent good agreement. Univariate predictors with $P < 0.05$ were used in the multivariate logistic regression analysis to identify independent predictors for early recurrence of HCC

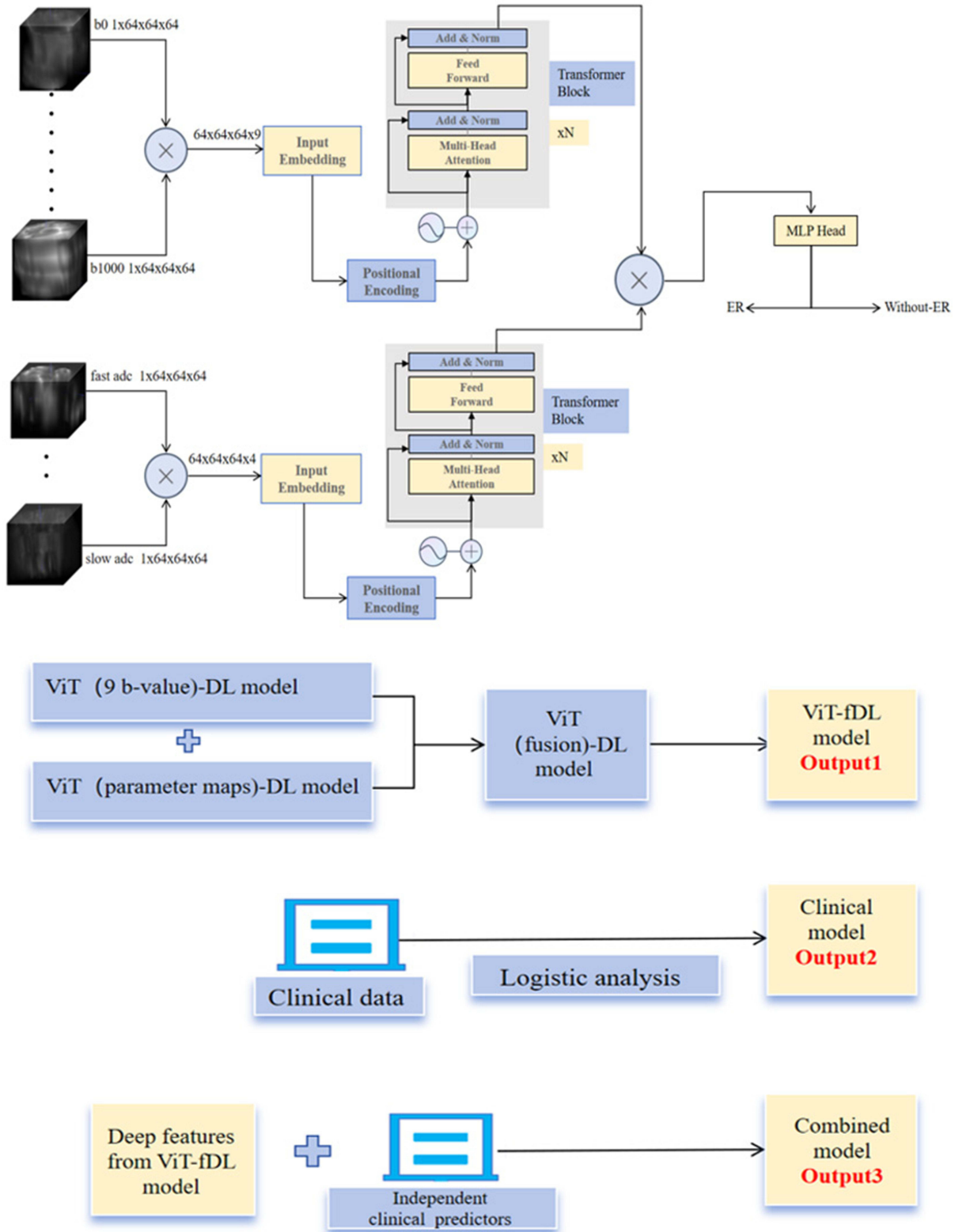


Figure 3 Model development.

Abbreviations: ViT, Vision transformer; DL, deep learning; ER, early recurrence.

to be used in the clinical model. The least absolute shrinkage and selection operator-logistic regression algorithm was applied to identify the most valuable features for predicting early recurrence of HCC in our combined model using the 5-fold cross-validation penalty procedure. Two-sided P values < 0.05 were considered statistically significant.

Results

Baseline Patient Characteristics

The study included 122 patients, 85 in the training set and 37 in the test set. The cohort included 49 (40.1%) patients with early recurrence, with similar rates in the training (40.0%, 34/85) and test (40.5%, 15/37) sets ($P = 0.971$). By the end of follow-up, the remaining 73 of the 122 patients censored (without observed early recurrence), and the proportion of censored cases in total population was 59.84%.

Baseline patient clinical characteristics in the training and test sets were similar (Table 1), and baseline clinical characteristics were not significantly different between the training and test sets (all $P > 0.05$). In the training set, tumor size, AST, MVI, and D were significantly different between patients with and without early recurrence (all $P < 0.05$; Table 2), so these were incorporated into the multivariate logistic regression model. Multivariable logistic regression showed that tumor size (OR=1.870; $P = 0.003$) and D value (OR=0.037; $P = 0.011$) were independent predictors for early recurrence of HCC. The baseline clinical characteristics of patients with and without early recurrence of HCC in the test set are shown in Table 3.

Table 1 Baseline Clinical Characteristic of Training and Test Sets

Variables	Training Set (n = 85)	Test Set (n =37)	P value
Age (years)	50.94±12.40	51.16±10.88	0.925 ^a
Sex (male vs female) (n)	73/12	35/2	0.281 ^c
Tumor size, (cm)	2.30±1.65	2.20±1.20	0.738 ^b
AFP level (<20 ng/mL VS ≥20ng/mL) (n)	48/37	20/17	0.805 ^c
ALT (IU/L)	26.00±17.00	27.00±24.00	0.224 ^b
AST (IU/L)	23.00±9.50	26.00±13.00	0.225 ^b
Total bilirubin (umol/L)	12.20±7.50	13.40±7.60	0.156 ^b
Direct bilirubin (umol/L)	4.90±3.40	5.00±3.30	0.222 ^b
Indirect bilirubin (umol/L)	7.00±5.05	8.20±3.90	0.093 ^b
Albumin (g/L)	40.50±5.00	41.70±6.00	0.332 ^b
Platelet count ($\times 10^9/L$)	161.45±69.58	166.51±55.28	0.696 ^a
Edmondson-Steiner grade ((G1/G2 vs G3/G4))(n)	67/18	26/11	0.308 ^c
Microvascular invasion (positive vs.negative) (n)	34/51	12/25	0.428 ^c
Barcelona Clinic Liver Cancer Stage (0 vs A/B) (n)	27/58	9/28	0.408 ^c
Child-Pugh (A/B) (n)	81/4	36/1	0.987 ^c
Origin of liver disease (HBV vs Other) (n)	83/2	36/1	1.000 ^c
Background liver (cirrhosis or chronic hepatitis vs non-cirrhotic chronic hepatitis) (n)	82/3	36/1	1.000 ^c
D (10^{-3} mm ² /s)	0.88±0.21	0.90±0.20	0.386 ^b
D* (10^{-3} mm ² /s)	54.00±35.00	51.00±46.00	0.863 ^b
f	0.15±0.10	0.14±0.13	0.802 ^b
ADC (10^{-3} mm ² /s)	1.40±0.50	1.40±0.60	0.700 ^b

Notes: Data are presented as the mean±standard deviation or median±interquartile range; Numbers, (n); ^aComparisons were performed by independent samples t -test; ^bComparisons were performed by Mann–Whitney U -test; ^cComparisons were performed by χ^2 or Fisher's exact test; D*represent abbreviation word of the pseudo-diffusion coefficient of IVIM.

Abbreviations: AFP, alpha-fetoprotein; ALT, alanine aminotransferase; AST, aspartate aminotransferase.

Table 2 Comparison Baseline Clinical Characteristics Between Early Recurrence and without Early Recurrence HCCs in Training Set

Variables	Early Recurrence (n = 34)	Without Early Recurrence (n =51)	P value
Age (years)	53.00±11.00	48.00±21.00	0.353 ^b
Sex (male vs female) (n)	28/6	45/6	0.656 ^c
Tumor size, (cm)	3.20±2.07	2.00±1.20	0.003^b
AFP level (<20 ng/mL VS ≥20ng/mL) (n)	17/17	31/20	0.326 ^c
ALT (IU/L)	29.00±19.50	25.00±17.00	0.205 ^b
AST (IU/L)	24.00±11.25	21.00±8.00	0.002^b
Total bilirubin (umol/L)	12.75±7.20	12.10±8.30	0.939 ^b
Direct bilirubin (umol/L)	4.85±3.20	4.90±3.60	0.462 ^b
Indirect bilirubin (umol/L)	6.80±4.35	7.00±5.70	0.809 ^b
Albumin (g/L)	40.10±6.10	41.10±4.30	0.127 ^b
Platelet count (×10 ⁹ /L)	169.00±113.50	170.00±85.00	0.932 ^b
Edmondson-Steiner grade ((G1/G2 vs G3/G4))	26/8	41/10	0.665 ^c
Microvascular invasion (positive vs.negative) (n)	18/16	16/35	0.047^c
Barcelona Clinic Liver Cancer Stage (0 vs A/B) (n)	7/27	20/31	0.071 ^c
Child-Pugh (A/B) (n)	31/3	50/1	0.347 ^c
Origin of liver disease (HBV vs Other) (n)	33/1	50/1	1.000 ^c
Background liver (cirrhosis or chronic hepatitis vs non-cirrhotic chronic hepatitis) (n)	33/1	49/2	1.000 ^c
D (10 ⁻³ mm ² /s)	0.81±0.22	0.92±0.20	0.011^b
D* (10 ⁻³ mm ² /s)	52.38±24.99	58.51±31.69	0.346 ^a
f	0.16±0.10	0.15±0.10	0.781 ^b
ADC (10 ⁻³ mm ² /s)	1.30±0.62	1.50±0.50	0.258 ^b

Notes: Data are presented as the mean±standard deviation or median±interquartile range; Numbers, (n); ^aComparisons were performed by independent samples t-test; ^bComparisons were performed by Mann-Whitney U-test; ^cComparisons were performed by χ^2 or Fisher's exact test; D*represent abbreviation word of the pseudo-diffusion coefficient of IVIM; Significant results are in bold.

Abbreviations: AFP, alpha-fetoprotein; ALT, alanine aminotransferase; AST, aspartate aminotransferase.

Table 3 Comparison Baseline Clinical Characteristics Between Early Recurrence and without Early Recurrence HCCs in Test Set

Variables	Early Recurrence (n = 15)	Without Early Recurrence (n =22)	P value
Age (years)	54.47±12.31	48.91±9.43	0.129 ^a
Sex (male vs female) (n)	15/0	20/2	0.505 ^c
Tumor size, (cm)	2.80±2.50	2.15±0.92	0.015^b
AFP level (<20 ng/mL VS ≥20ng/mL) (n)	5/10	15/7	0.050 ^c
ALT (IU/L)	25.00±21.00	27.00±27.50	0.610 ^b
AST (IU/L)	25.00±19.00	26.50±12.50	0.467 ^b
Total bilirubin (umol/L)	14.55±6.16	14.39±5.27	0.932 ^a
Direct bilirubin (umol/L)	4.90±4.10	5.00±2.45	0.665 ^b
Indirect bilirubin (umol/L)	8.61±4.18	8.91±3.87	0.824 ^a
Albumin (g/L)	42.04±6.67	41.00±4.44	0.572 ^a
Platelet count (×10 ⁹ /L)	178.80±47.09	158.13±59.83	0.270 ^a
Edmondson-Steiner grade ((G1/G2 vs G3/G4)) (n)	10/5	16/6	0.728 ^c
Microvascular invasion (positive vs.negative) (n)	8/7	4/18	0.036^c
Barcelona Clinic Liver Cancer Stage (BCLC 0 vs A/B) (n)	2/13	7/15	0.262 ^c
Child-Pugh (A/B) (n)	14/1	22/0	0.405 ^c
Origin of liver disease (HBV vs Other) (n)	15/0	21/1	1.000 ^c

(Continued)

Table 3 (Continued).

Variables	Early Recurrence (n = 15)	Without Early Recurrence (n =22)	P value
Background liver (cirrhosis or chronic hepatitis vs non-cirrhotic chronic hepatitis) (n)	15/0	21/1	1.000 ^c
D (10 ⁻³ mm ² /s)	0.78±0.20	0.94±0.20	0.038 ^b
D* (10 ⁻³ mm ² /s)	48.60±24.02	64.82±36.14	0.137 ^a
f	0.18±0.11	0.15±0.08	0.341 ^a
ADC (10 ⁻³ mm ² /s)	1.30±0.70	1.55±0.60	0.555 ^b

Notes: Data are presented as the mean±standard deviation or median± interquartile range; Numbers, (n); ^aComparisons were performed by independent samples t-test; ^bComparisons were performed by Mann–Whitney U-test; ^cComparisons were performed by χ^2 or Fisher's exact test; D* represent abbreviation word of the pseudo-diffusion coefficient of IVIM; Significant results are in bold.

Abbreviations: AFP, alpha-fetoprotein; ALT, alanine aminotransferase; AST, aspartate aminotransferase.

Interobserver Agreement in the Training Set

The ICC assessment indicated good-to-excellent agreement between radiologists in evaluating the IVIM parameters (ICC=0.800–0.884). Agreement between the two MRI scanners in evaluating the IVIM parameters was good-to-excellent (ICC=0.996, [Supplementary Table S2](#)).

Performance of the ViT-DL Models

The respective AUCs in the training and test sets were 0.957 (95% CI, 0.889–0.989) and 0.747 (95% CI, 0.577–0.875) in the ViT-bDL model, 0.751 (95% CI, 0.646–0.839) and 0.594 (95% CI, 0.420–0.752) in the ViT-mDL model. The ViT-fDL model which combine deep learning features of 9 b-value images and IVIM parametric maps improved performance in both training and test sets, with AUC of 0.968 (95% CI 0.905–0.994) in the training set and 0.815 (95% CI 0.653–0.923) in the test set.

Performance of the Clinical and Combined Model

The AUCs in the training and test sets in the clinical model were 0.755 (95% CI, 0.650–0.842) and 0.764 (95% CI, 0.596–0.887), respectively. We selected the best-performing ViT-fDL model from three deep learning models to integrate its ten most relevant deep features with the independent clinical predictors to generate the combined model, which achieved AUCs of 0.991 (95% CI, 0.940–1.000) and 0.821 (95% CI, 0.660–0.927) in the training and test sets, respectively ([Figure 4](#) and [Table 4](#)).

Decision curve analysis revealed that the combined model performed good clinical utility in both the training and test sets in predicting early recurrence of HCC ([Figure 4](#)). Calibration curves for the combined model in both the training and test sets are presented in [Supplementary Figure S1](#), which showed predicted and actual probabilities of HCC with early recurrence. And both P values of the Hosmer-Lemeshow test in the training (P=0.932) and test sets (P=0.125) were greater than 0.05, which also showed that fitting of combined model was acceptable, it means no significant difference between the predicted probability and the actual probability.

Model Interpretability

Grad-CAM provided valuable information for predicting early recurrence of HCC. It facilitated visualizing the pixel weight distribution using various colors. The visualization results of the ViT-DL model are shown in [Figure 5](#) to better understand the differences in the IVIM images between patients with and without early recurrence of HCC. Notably, Grad-CAM indicated that the DL network fixated most of its attention on the multiple and large tumor marginal regions in early recurrence of HCC, while Grad-CAM showed that few regions were activated within the HCC tumor of without early recurrence.

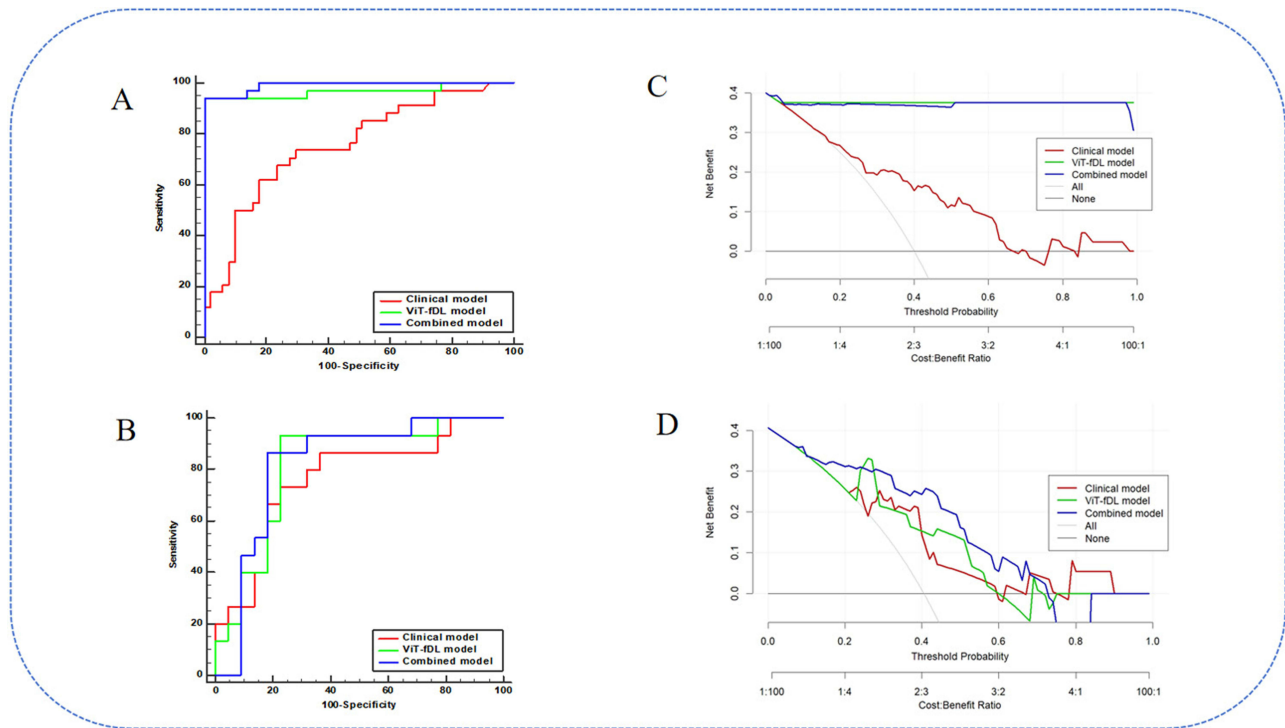


Figure 4 Model evaluation. The ROCs for the clinical model, ViT-fDL model, and combined model of the training set (A) and test set (B). The DCAs for the clinical model, ViT-fDL model, and combined model of the training set (C) and test set (D).

Radiomics quality score (RQS) provided in [Supplementary material 6](#), and the codes for ViT-DL model analysis are available in a GitHub repository (<https://github.com/alan-qi-gzum/CL-VIT>).

Discussion

This study established a ViT-fDL model for predicting early recurrence of HCC by extracting deep features from nine b-value images and IVIM parametric maps. The ViT-fDL model demonstrated better predictive performance for early

Table 4 Diagnostic Performance of the Clinical Model, ViT-fDL Model, and Combined Model

	Training Set				
	AUC (95% CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Clinical model	0.755 (0.650–0.842)	73.5	70.6	62.5	80.0
ViT-fDL model	0.968 (0.905–0.994)	94.1	100	100	96.2
Combinedmodel	0.991 (0.940–1.000)	94.1	100	100	96.2
	Test set				
	AUC (95% CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Clinical model	0.764 (0.596–0.887)	73.3	77.3	68.8	81.0
ViT-fDL model	0.815 (0.653–0.923)	93.3	77.3	73.7	94.4
Combined model	0.821 (0.660–0.927)	86.7	81.8	76.5	90.0

Abbreviations: AUC, area under the curve; PPV, positive prediction value; NPV, negative prediction value; CI, confidence interval.

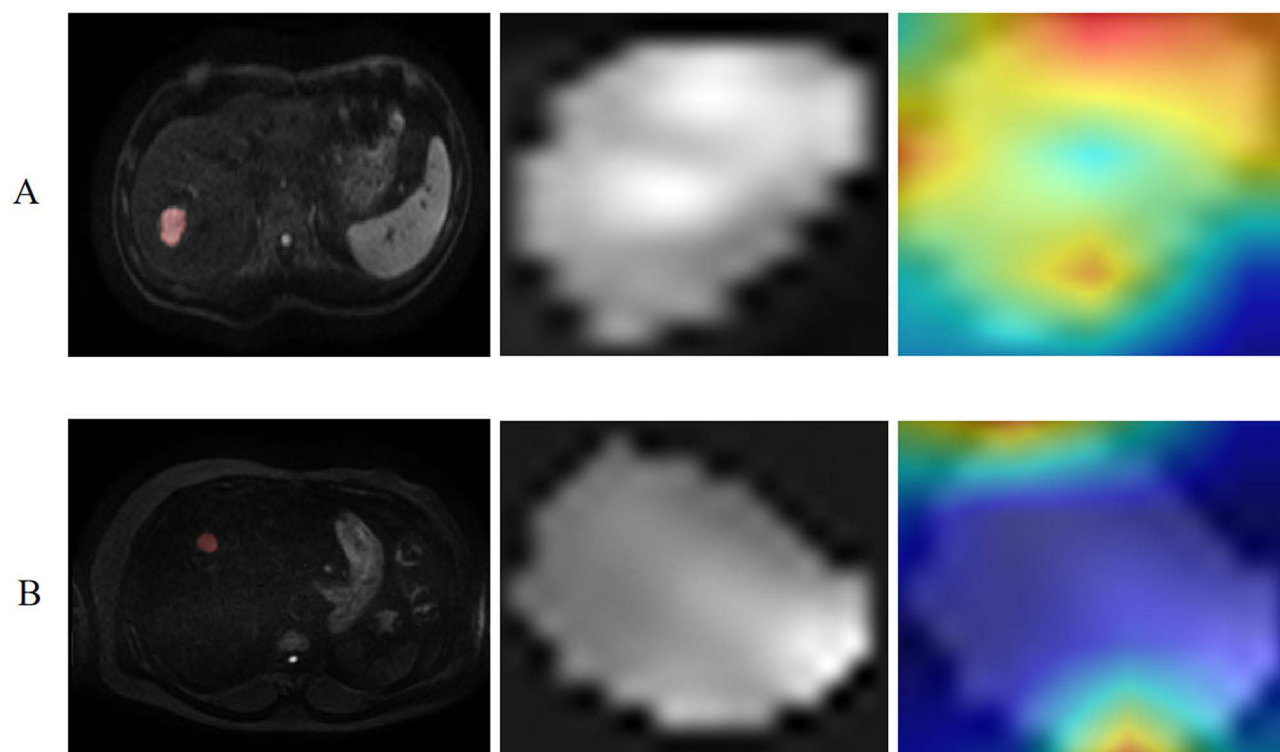


Figure 5 Visualization of gradient-weighted class activation mapping (Grad-CAM). Representative images and illustration of visualization results for early recurrence case (**A**) and without early recurrence case (**B**). From left to right in each row: origin DWI $b=1000s/mm^2$ image; the cropped tumor region for input; Grad-CAM; Grad-CAM overlaid on the cropped tumor region, where red indicates highest contribution to the classification, followed by yellow, while green and blue regions indicate lower contribution. (**A**) Grad-CAM showed that large and multiple marginal regions activated within the tumor of early recurrence case, indicating that tumor heterogeneity and margin invasion contribute to prediction of early recurrence in HCC; (**B**) Grad-CAM emphasized few regions activated within the tumor of without early recurrence case.

recurrence of HCC than the ViT-bDL, ViT-mDL, and clinical models. This ViT-fDL model could become a noninvasive and effective approach for predicting early recurrence of HCC.

By integrating deep features from the ViT-fDL model with clinical features, we developed and validated a combined model that accurately identified HCC patients at high risk of early recurrence, achieving the best predictive performance with AUCs of 0.991 and 0.821 in the training and test sets, respectively.

Notably, considering IVIM sequences did not require contrast agents, our combined model demonstrated efficacy in predicting early recurrence of HCC, similar to previous studies based on contrast-enhanced imaging.^{8,22} Therefore, the combined model is of great clinical significance for patients with HCC who cannot undergo examinations with contrast agents.

This study evaluated the performance of quantitative IVIM parameters for predicting early recurrence of HCC. Previous studies showed that lower ADC and D values were useful for predicting early recurrence of HCC. However, we found that only the D value was an independent predictor for early recurrence in the clinical model. This difference could be because previous studies derived the ADC from a single ROI on a maximum tumor cross-section.^{3,9} In contrast, we used VOIs, which included necrotic portions, yielding different results. We believe that whole tumor analysis, as done in our study, captures the high heterogeneity of HCC. Relying solely on maximum tumor cross-section might not offer comprehensive information about the entire tumor. Our results showed the tumor size was another independent predictor for early recurrence of HCC, similar to previous studies.^{5,21} We built a clinical model based on the D value and tumor size, resulting in AUCs of 0.755 and 0.764 in the training and test sets, respectively. However, high-dimension features of these data could not be extracted by traditional statistical methods for further analysis, limiting the value of clinical model.

Previous research predominantly relied on 2D analysis of the maximum HCC tumor cross-section to predict MVI and survival outcomes.^{8,21,23} Considering the high heterogeneity of HCC tumors, relying solely on maximum tumor cross-

section might not offer comprehensive information about the entire tumor.¹⁷ Moreover, some studies have confirmed that the whole-tumor volume was associated with HCC prognosis.^{24–26} Our ViT-bDL model processed and integrated diverse data through a multi-attention mechanism, efficiently mining high-dimensional image features to comprehensively quantify tumor information. Furthermore, our ViT-bDL model superimposed the nine b-value images in the channel dimension, facilitating a deep understanding of the content of each b-value image and capturing potential patterns and complex spatial relationships between multiple b-values.^{17,18,27} Consequently, the ViT-bDL model provided accurate and comprehensive insights for predicting early recurrence of HCC.

We also evaluated the performance of IVIM parametric maps, using the ViT-mDL model, in predicting early recurrence of HCC. Our results showed that the ViT-bDL model performed better than the ViT-mDL model. This difference adequately reflects the advantages of multiple b-value images over IVIM parametric maps when using ViT deep features for predicting early recurrence of HCC. This difference is due to the fact that the IVIM parametric map calculation is usually prone to errors and is calculation intensive due to the pixel-by-pixel fitting of the bi-exponential model.^{23,28} These drawbacks could be responsible for the relatively poor predictive efficiency of the IVIM parametric maps for early recurrence of HCC. To overcome this limitation and improve the predictive efficiency, we constructed a fusion ViT-fDL model that combined the ViT-bDL and ViT-mDL models.

Grad-CAM is a visual interpretation tool that helps clinicians explain how the ViT-DL model extracted deep features from regions inside the tumor that are most influential for the model's predictions. Grad-CAM showed that most attention was paid to extracting deep features from the tumor's margins for predicting early recurrence of HCC. Tumors in patients with early recurrence of HCC showed more heterogeneity than those in patients without early recurrence of HCC. Consistent with Zhao et al,²² our results suggest that intra-tumoral heterogeneity and invasive margins contribute to predicting early recurrence of HCC.

The limitations of this study must be acknowledged. First, single center with a relatively small population especially test set and potential selection bias may affect the generalization of our model, The IVIM sequence might not be available at some institutions. Furthermore, the setting of the multiple b values in the IVIM sequence is uncertain. However, our combined model yielded a predictive accuracy for HCC prognosis similar to that of DL research using contrast-enhanced MRI. Therefore, expanding the database and performing prospective multi-center studies using IVIM sequences should be performed in future in order to ascertain the generalizability of the model. Second, the calibration curves were not very ideal due to the relatively small sample size. However, the mean absolute error values of the calibration curves in both the training and test sets were less than 0.1, which indicated that the overall fitting effect of our combined model were acceptable, In general, the overall fitting effect of our combined model through MAE analysis and Hosmer–Lemeshow test (all $P > 0.05$) is reasonable. Thus, there is room for improvement in the calibration curve through large-scale and multi-center studies in future. Third, we used early stopping and cross-validation techniques to mitigate the risk of overfitting arising from our small sample. Finally, our study aimed to explore whether a transformer framework-based DL network based on IVIM images could predict early recurrence of HCC preoperatively. Future research should integrate other conventional scanning sequences, such as T2WI. The fusion of multi-sequence information could help to explore effective biomarkers for predicting early recurrence of HCC.

Conclusion

A combined model that incorporated deep features from the ViT-fDL model and independent clinical predictors achieved excellent predictive performance for predicting early recurrence of HCC. Thus, the combined model is proposed as a useful tool for predicting early recurrence of HCC and holds promise for guiding individualized therapies. Our research currently represents preliminary findings, and future work will require expanding the sample size and external validation for further investigation.

Abbreviations

ADC, Apparent diffusion coefficient; AFP, Alpha-Fetoprotein; ALT, alanine transaminase; AST, aspartate aminotransferase; AUC, Area under the curve; CI, Confidence interval; CNN, Convolutional neural networks; D, True diffusion coefficient; DL, Deep learning; DWI, Diffusion-weighted imaging; Grad-CAM, Gradient-weighted class activation

mapping; HCC, Hepatocellular carcinoma; ICC, Interclass correlation coefficient; IVIM, Intravoxel incoherent motion; MRI, Magnetic resonance imaging; MVI, Microvascular invasion; NPV, negative predictive value; OR, Odds ratio; PPV, positive predictive value; ROC, Receiver operating characteristic curve; ViT, Vision transformer; ViT-bDL, Vision transformer-b-value deep learning; ViT-fDL, Vision transformer-fused deep learning; ViT-mDL, Vision transformer-intravoxel incoherent motion parametric map deep learning; VOI, Volumes of interest.

Data Sharing Statement

The datasets used and analyzed during the current study are available from the corresponding author (Yikai Xu) upon reasonable request.

Ethics Approval and Informed Consent

This retrospective study was conducted in accordance with the Helsinki Declaration. This retrospective study was approved by the Ethics Committee of Nanfang Hospital of Southern Medical University (approval number: NFEC-202305-Y10). Informed consent was waived by the Ethics Committee due to the retrospective design of anonymized patient data. We confirmed that the data was anonymized or maintained with confidentiality.

Acknowledgments

We sincerely thank Phiips' scientist Jun Peng.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was funded by the National Natural Science Foundation of China (82271939).

Disclosure

The authors declare that they have no competing interests in this work.

References

1. Vogel A, Meyer T, Sapisochin G, Salem R, Saborowski A. Hepatocellular carcinoma. *Lancet*. 2022;400(10360):1345–1362. doi:10.1016/S0140-6736(22)01200-4
2. Reig M, Forner A, Rimola J, et al. BCLC strategy for prognosis prediction and treatment recommendation: the 2022 update. *J Hepatol*. 2022;76(3):681–693. doi:10.1016/j.jhep.2021.11.018
3. Lu Y, Wang H, Li C, et al. Preoperative and postoperative mri-based models versus clinical staging systems for predicting early recurrence in hepatocellular carcinoma. *Eur J Surg Oncol*. 2024;50(9):108476. doi:10.1016/j.ejso.2024.108476
4. Chan A, Zhong J, Berhane S, et al. Development of pre and post-operative models to predict early recurrence of hepatocellular carcinoma after surgical resection. *J Hepatol*. 2018;69(6):1284–1293. doi:10.1016/j.jhep.2018.08.027
5. Xing H, Zhang WG, Cescon M, et al. Defining and predicting early recurrence after liver resection of hepatocellular carcinoma: a multi-institutional study. *HPB (Oxford)*. 2020;22(5):677–689. doi:10.1016/j.hpb.2019.09.006
6. Gao W, Wang W, Song D, et al. A predictive model integrating deep and radiomics features based on gadobenate dimeglumine-enhanced mri for postoperative early recurrence of hepatocellular carcinoma. *Radiol Med*. 2022;127(3):259–271. doi:10.1007/s11547-021-01445-6
7. Zhong JH, Li LQ. Postoperative adjuvant transarterial chemoembolization for participants with hepatocellular carcinoma: a meta-analysis. *Hepatol Res*. 2010;40(10):943–953. doi:10.1111/j.1872-034X.2010.00710.x
8. Yan M, Zhang X, Zhang B, et al. Deep learning nomogram based on gd-eob-dtpa mri for predicting early recurrence in hepatocellular carcinoma after hepatectomy. *Eur Radiol*. 2023;33(7):4949–4961. doi:10.1007/s00330-023-09419-0
9. Zhang Y, Kuang S, Shan Q, et al. Can IVIM help predict hcc recurrence after hepatectomy? *Eur Radiol*. 2019;29(11):5791–5803. doi:10.1007/s00330-019-06180-1
10. Chen J, Sun W, Wang W, et al. Diffusion-based virtual mr elastography for predicting recurrence of solitary hepatocellular carcinoma after hepatectomy. *Cancer Imaging*. 2024;24(1):106. doi:10.1186/s40644-024-00759-8

11. Zhang Y, Lv X, Qiu J, et al. Deep learning with 3d convolutional neural network for noninvasive prediction of microvascular invasion in hepatocellular carcinoma. *J Magn Reson Imaging*. 2021;54(1):134–143. doi:10.1002/jmri.27538
12. Sun BY, Gu PY, Guan RY, et al. Deep-learning-based analysis of preoperative MRI predicts microvascular invasion and outcome in hepatocellular carcinoma. *World J Surg Oncol*. 2022;20(1):189. doi:10.1186/s12957-022-02645-8
13. Song D, Wang Y, Wang W, et al. Using deep learning to predict microvascular invasion in hepatocellular carcinoma based on dynamic contrast-enhanced mri combined with clinical parameters. *J Cancer Res Clin Oncol*. 2021;147(12):3757–3767. doi:10.1007/s00432-021-03617-3
14. Wang W, Wang Y, Song D, et al. A transformer-based microvascular invasion classifier enhances prognostic stratification in hcc following radiofrequency ablation. *Liver Int*. 2024;44(4):894–906. doi:10.1111/liv.15846
15. Jiang X, Zhao H, Saldanha OL, et al. An mri deep learning model predicts outcome in rectal cancer. *Radiology*. 2023;307(5):e222223. doi:10.1148/radiol.222223
16. Shamshad F, Khan S, Zamir SW, et al. Transformers in medical imaging: a survey. *Med Image Anal*. 2023;88:102802. doi:10.1016/j.media.2023.102802
17. Zheng Y, Qiu B, Liu S, et al. A transformer-based deep learning model for early prediction of lymph node metastasis in locally advanced gastric cancer after neoadjuvant chemotherapy using pretreatment ct images. *EClinicalMedicine*. 2024;75:102805. doi:10.1016/j.eclinm.2024.102805
18. Qu H, Zhang S, Guo M, et al. Deep learning model for predicting proliferative hepatocellular carcinoma using dynamic contrast-enhanced mri: implications for early recurrence prediction following radical resection. *Acad Radiol*. 2024;31(11):4445–4455. doi:10.1016/j.acra.2024.04.028
19. Xu Y, Zhou C, He X, et al. Deep learning-assisted li-rads grading and distinguishing hepatocellular carcinoma (hcc) from non-hcc based on multiphase ct: a two-center study. *Eur Radiol*. 2023;33(12):8879–8888. doi:10.1007/s00330-023-09857-w
20. Zhang YB, Chang ZQ, Bu Y, Lei P, Yang W, Zhang W. Construction of a 2.5d deep learning model for predicting early postoperative recurrence of hepatocellular carcinoma using multi-view and multi-phase ct images. *J Hepatocell Carcinoma*. 2024;11:2223–2239. doi:10.2147/JHC.S493478
21. Mu T, Zheng X, Song D, et al. Deep learning based on multiparametric mri predicts early recurrence in hepatocellular carcinoma patients with solitary tumors ≤ 5 cm. *Eur J Radiol Open*. 2024;13:100610. doi:10.1016/j.ejro.2024.100610
22. Zhao Y, Wang S, Wang Y, et al. Deep learning radiomics based on contrast enhanced mri for preoperatively predicting early recurrence in hepatocellular carcinoma after curative resection. *Front Oncol*. 2024;14:1446386. doi:10.3389/fonc.2024.1446386
23. Liu B, Zeng Q, Huang J, et al. IVIM using convolutional neural networks predicts microvascular invasion in HCC. *Eur Radiol*. 2022;32(10):7185–7195. doi:10.1007/s00330-022-08927-9
24. Wei H, Zheng T, Zhang X, et al. Deep learning-based 3d quantitative total tumor burden predicts early recurrence of bclc a and b hcc after resection. *Eur Radiol*. 2025;35(1):127–139. doi:10.1007/s00330-024-10941-y
25. Wang J, Chen Z, Wang L, et al. A new model based inflammatory index and tumor burden score (tbs) to predict the recurrence of hepatocellular carcinoma (hcc) after liver resection. *Sci Rep*. 2022;12(1):8670. doi:10.1038/s41598-022-12518-5
26. Jeon SK, Lee DH, Park J, et al. Tumor volume measured using mr volumetry as a predictor of prognosis after surgical resection of single hepatocellular carcinoma. *Eur J Radiol*. 2021;144:109962. doi:10.1016/j.ejrad.2021.109962
27. Sato M, Moriyama M, Fukumoto T, et al. Development of a transformer model for predicting the prognosis of patients with hepatocellular carcinoma after radiofrequency ablation. *Hepatol Int*. 2024;18(1):131–137. doi:10.1007/s12072-023-10585-y
28. Scalco E, Rizzo G, Mastropietro A. The quantification of intravoxel incoherent motion - mri maps cannot preserve texture information: an evaluation based on simulated and in-vivo images. *Comput Biol Med*. 2023;154:106495. doi:10.1016/j.combiomed.2022.106495

Journal of Hepatocellular Carcinoma

Publish your work in this journal

The Journal of Hepatocellular Carcinoma is an international, peer-reviewed, open access journal that offers a platform for the dissemination and study of clinical, translational and basic research findings in this rapidly developing field. Development in areas including, but not limited to, epidemiology, vaccination, hepatitis therapy, pathology and molecular tumor classification and prognostication are all considered for publication. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-hepatocellular-carcinoma-journal>

Dovepress
Taylor & Francis Group