Development and Preliminary Validation of a Novel Convolutional Neural Network Model for Predicting Treatment Response in Patients with Unresectable Hepatocellular Carcinoma Receiving Hepatic Arterial Infusion Chemotherapy

Bing Quan[1,2] · Jinghuan Li[1,2] · Hailin Mi[4] · Miao Li[1,2] · Wenfeng Liu[1,2] · Fan Yao[1,2] · Rongxin Chen[1,2] · Yan Shan[3] · Pengju Xu[3] · Zhenggang Ren[1,2] · Xin Yin[1,2]

Abstract
The goal of this study was to evaluate the performance of a convolutional neural network (CNN) with preoperative MRI and clinical factors in predicting the treatment response of unresectable hepatocellular carcinoma (HCC) patients receiving hepatic arterial infusion chemotherapy (HAIC). A total of 191 patients with unresectable HCC who underwent HAIC in our hospital between May 2019 and March 2022 were retrospectively recruited. We selected InceptionV4 from three representative CNN models, AlexNet, ResNet, and InceptionV4, according to the cross-entropy loss (CEL). We subsequently developed InceptionV4 to fuse the information from qualified pretreatment MRI data and patient clinical factors. Radiomic information was evaluated based on several constant sequences, including enhanced T1-weighted sequences (with arterial, portal, and delayed phases), T2 FSE sequences, and dual-echo sequences. The performance of InceptionV4 was cross-validated in the training cohort ($n=127$) and internally validated in an independent cohort ($n=64$), with comparisons against single important clinical factors and radiologists in terms of receiver operating characteristic (ROC) curves. Class activation mapping was used to visualize the InceptionV4 model. The InceptionV4 model achieved an AUC of 0.871 (95% confidence interval [CI] 0.761–0.981) in the cross-validation cohort and an AUC of 0.826 (95% CI 0.682–0.970) in the internal validation cohort; these two models performed better than did the other methods (AUC ranges 0.783–0.873 and 0.708–0.806 for cross- and internal validations, respectively; $P < 0.01$). The present InceptionV4 model, which integrates radiomic information and clinical factors, helps predict the treatment response of unresectable HCC patients receiving HAIC treatment.

Keywords Hepatocellular carcinoma · Hepatic arterial infusion chemotherapy · Treatment response · Convolutional neural network

Abbreviations

HCC	Hepatocellular carcinoma
TACE	Transarterial chemoembolization
HAIC	Hepatic arterial infusion of oxaliplatin
CNNs	Convolutional neural networks
BCLC	Barcelona Clinic Liver Cancer
HbsAg	Hepatitis B surface antigen
AFP	Alpha fetoprotein
PIVKA-II	Protein induced by vitamin K absence or antagonist-II
PLT	Platelet
TB	Total bilirubin
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
AKP	Alkaline phosphatase
γ-GT	Gamma-glutamyl transpeptidase
CRP	C-reactive protein
OS	Overall survival
PFS	Progression-free survival
T2WI	T2-weighted imaging
FSE	Fast spin echo
TLs	Target lesions
CR	Complete response
PR	Partial response
SD	Stable disease

Key Points

Treatment response prediction of HCC patients after HAIC can be reliably diagnosed by CNN.

Jinghuan Li and Hailin Mi are co-first author.

Extended author information available on the last page of the article

PD	Progressive disease
ORR	Objective response rate
DCR	Disease control rate
CEL	Cross entropy loss
SVM	Support vector machine
AUC	Area under the ROC curve
SEN	Sensitivity
SPE	Specificity
PPV	Positive predictive value
NPV	Negative predictive value
CAM	Class activation mapping
IQR	Interquartile range
NLR	Neutrophil-lymphocyte ratio
FOV	Field of view
UV	Univariable
MV	Multivariable
CI	Confidence interval
HR	Hazard ratio

Introduction

Hepatocellular carcinoma (HCC), accounting for 75–85% of primary liver cancer cases, is the third leading cause of cancer-related death, with approximately 780,000 deaths annually [1]. A comparative study analyzing more than 8500 patients receiving treatment for HCC showed that less than 10% of the candidates met the criteria for liver resection [2]. For patients with unresectable or advanced HCC who are ineligible for surgery, recommended first-line systemic treatments include the single agents sorafenib, lenvatinib, and donafenib [3–5]. Moreover, tumor immunotherapy has been reported to provide promising outcomes. The combination of atezolizumab and bevacizumab has become the standard first-line systemic therapy according to the IMbrave150 trial [6]. Additionally, the evidence from a recent phase III randomized clinical trial indicated that tislelizumab may represent a potential first-line treatment option for patients with unresectable HCC [7]. However, the effectiveness of molecular targeted therapy and immunotherapy is limited by the low response rate. Additionally, the high cost of immunotherapy may prevent it from being widely used. Transarterial chemoembolization (TACE) is the most widely used treatment option for patients with unresectable HCC [8, 9]. However, TACE is not effective in patients with diffuse HCC, arterioportal/arteriovenous shunts, or major vascular tumor thrombosis [10]. As hepatic arterial infusion chemotherapy (HAIC) is associated with increased local agent concentrations and reduced systemic distribution of anticancer drugs, it has shown stronger treatment efficacy and fewer systemic adverse effects than other systemic therapies [11–15]. Clinical experience has indicated that HAIC is more suitable for patients with arteriovenous/arterioportal shunts and major vascular tumor thrombosis. HAIC was listed in the Japan Society of Hepatology guidelines in 1995 and accepted as an interventional therapy for advanced HCC by the Chinese Society of Clinical Oncology [15, 16]. In addition, according to the Pan-Asian–adapted European Society for Medical Oncology Guidelines, HAIC was recommended as the first-line option for advanced, nonmetastatic HCC with macrovascular invasion [17].

According to the retrospective study of Zhao et al., HAIC was associated with longer OS than sorafenib (14.5 and 7.0 months, $P < 0.001$) in patients with locally advanced HCC [18]. A clinical trial by Shi et al. indicated that FOLFOX-HAIC plus sorafenib had a greater ORR (40.8% vs. 2.46%, $P < 0.001$) than sorafenib alone [19]. On the other hand, Shi et al. showed that FOLFOX-HAIC significantly improved OS compared with TACE in patients with unresectable large HCC [20]. Additionally, compared with sorafenib alone, HAIC plus sorafenib was associated with longer median OS (16.3 vs. 6.5 months, $P < 0.001$) and PFS (9.0 vs. 2.5 months, $P < 0.001$) and a greater ORR (41% vs. 3%, $P < 0.001$) in patients with major portal vein tumor thrombosis [21]. Moreover, HAIC combined with targeted therapy or immunotherapy can be used when patients present with tumor progression after traditional TACE treatment in clinical practice [22]. In many studies, scholars have reported high treatment response rates in patients with unresectable HCC receiving HAIC, and the clinical response significantly affects survival. A comparative study showed that HAIC may improve OS in patients with advanced HCC with major portal vein tumor thrombosis compared to TACE/TAE treatment [23]. Moreover, a meta-analysis indicated that HAIC may have better therapeutic efficacy than TACE in patients with unresectable HCC [24]. To our knowledge, three models have been developed to predict the therapeutic response to HAIC in advanced or unresectable HCC patients [25–27]. Among them, two models used contrast-enhanced computed tomography, and one model used only axial T2-weighted images. However, not all of those studies considered the use of convolutional neural networks (CNNs), which are tools that focus on image processing.

Deep learning has been widely applied in the area of prediction owing to its advantages of being accurate, fast, and reproducible [28]. CNNs have gained extensive attention related to pattern recognition in images as a strategy for deep learning [29, 30]. CNNs can find image details that physicians cannot find and automatically make a quantitative evaluation. The aim of this study was to evaluate the potential of combining CNNs with preoperative MRI and clinical factors for the prediction of treatment response in unresectable HCC patients treated with HAIC.

Patients and Methods

Study Population

This retrospective study was performed in accordance with the institutional ethics committee of our hospital and the Helsinki

Declaration. The ethics review number is KY2022078. Between May 2019 and March 2022, a total of 191 patients with unresectable HCC after HAIC from our hospital were included in this study. HCC was pathologically confirmed by liver biopsy or clinically confirmed by the American Association for the Study of Liver Diseases guidelines. Tumors were considered unresectable either because patients were at an advanced stage or because patients had insufficient remnant liver volume after liver surgery ($< 40\%$ for patients with cirrhosis; $< 30\%$ for patients without cirrhosis). The key inclusion criteria for our study included the following: (1) patients had received at least one HAIC cycle; (2) Eastern Cooperative Oncology Group Performance Status of 0–1; (3) Child−Pugh class A or B; (4) adequate hematologic function—white blood cell counts $\geq 3.0 \times 10^9$/L, neutrophils $\geq 1.5 \times 10^9$/L, platelets $\geq 50 \times 10^9$/L, hemoglobin ≥ 100 g/L; (5) adequate renal function—serum creatinine ≤ 2 mg/dL. The exclusion criteria were as follows: (1) had severe cardiovascular and cerebrovascular diseases, (2) had other malignant tumors in addition to HCC, (3) unavailable qualified pretreatment MRI at our institution, (4) unavailable clinical medical records, and (5) lacked survival data.

Treatment Procedure

For the HAIC procedure, we introduced a 5F RH catheter (Cook, Bloomington, USA) or a 5F MPA catheter (Cordis, USA) through the femoral artery or the radial artery via the Seldinger technique. Then, we performed angiographic surveys of the celiac trunk and superior mesenteric artery. We advanced the catheter or a 2.7F microcatheter (Terumo, Tokyo, Japan) into the hepatic artery with reference to tumor size, location, and arterial supply. The catheter or microcatheter was connected to a chemotherapy infusion pump. After HAIC treatment, the catheter or microcatheter was removed. HAIC was delivered via segmental, lobar, or whole-liver infusion according to the tumor burden. The FOLFOX regimen was used as previously reported [22]. Specifically, at the discretion of the treating physician, oxaliplatin was administered at 85 mg/m^2 for 2 h, leucovorin at 200–400 mg/m^2, and 5-fluorouracil at 2400 mg/m^2 for 46 h (48 h protocol) or 1200 mg/m^2 for 22 h (24 h protocol). HAIC was performed every 3–4 weeks and was repeated unless deterioration of hepatic function, deterioration of clinical conditions, tumor progression, or unacceptable toxicity occurred.

Data Collection

The clinical characteristics of each patient included age, sex, pre-HAIC treatment, number of HAIC cycles, HAIC regimen, tumor number, maximum tumor size, tumor thrombus, extrahepatic metastasis, stage, liver function, hepatitis panel, liver cirrhosis markers, circulating blood cell laboratory values, best tumor response, overall survival (OS), and progression-free survival (PFS). OS was defined as the time interval between the first day of treatment and the date of death or last follow-up. PFS was defined as the time interval between the first day of treatment and the first documented disease progression or death.

MRI Acquisition

Each patient underwent gadoxetic acid–based MRI using a 3.0-Tesla MR scanner (uMR 770; United Imaging Health care, Shanghai, China). The routine plain-scan liver MRI protocol included diffusion-weighted imaging, T2-weighted imaging (T2WI), T1-weighted in-phase imaging, opposed-phase imaging, and contrast-enhanced T1-weighted imaging in the arterial phase (20–30 s), portal venous phase (60–80 s), and delayed phase (180 s) after injection of gadoxetic acid (extracellular) at a dose of 0.1 mmol/kg and at a rate of 2 mL/s, followed by a 20-mL saline flush. The detailed parameters of each MRI sequence are listed in Supplementary Table 1. In this study, enhanced T1-weighted sequences (including arterial, portal, and delayed phases), T2 fast spin echo (FSE) sequences, and dual echo sequences were selected for radiomic analysis.

Follow-Up and Response Evaluation

All enrolled patients were treated and monitored regularly. Tumor response was evaluated every 1–3 HAIC cycles according to the enhanced T1-weighted sequence in the arterial phase by two investigators according to Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 [31]. The tumor response could be divided into four conditions: complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD). CR was defined as the disappearance of all target lesions (TLs) (up to two measurable liver lesions). PR was defined as a $\geq 30\%$ decrease in the sum of the greatest one-dimensional diameters of TLs. PD was defined as a $\geq 20\%$ increase in the sum of the diameters of the TLs. SD was defined as a case that did not qualify for either PR or PD. The objective response rate (ORR) was calculated as the percentage of patients with the best ORR or PR. The disease control rate (DCR) was the percentage of patients with the best ORR, PR, or standard deviation (SD). We analyzed adverse events according to the National Cancer Institute Common Terminology Criteria for adverse events version 4.03. In our study, we considered patients who achieved CR, PR, or SD as having a positive treatment response, while patients with PD were considered to have a poor treatment response. The timeframe for evaluating tumor response was during our median follow-up time.

Convolutional Neural Network Development

For each patient, consecutive images were collected from enhanced T1-weighted sequences (including arterial, portal, and delayed phases), T2 FSE sequences, and dual-echo sequences. In this study, we extracted features by deep learning model automatically instead of segmenting the MRI images manually. A total of 191 patients were randomly divided into a training group ($n = 127$) and a validation group ($n = 64$). Because of the limited amount of data in our training group, data augmentation and image processing were performed to increase the amount of data to 10 times the original size according to random geometric image transformations, including shifting, rotation, flipping, and enlarging. In this way, we could make the model focus on the lesions rather than various sources of noise. All the transformed images were resized to 450×320 pixels to standardize the distance scale. The data augmentation method was proven to help prevent network overfitting and memorization of the exact details of the images. All the above steps were conducted in Python (version 3.7) by using the resize function in OpenCV (https://opencv.org).

In our study, we used CNN models to transform the MRI images of each patient into digital features, and different sequences were input into different channels. The image features were combined with the other 27 clinical features as the overall characteristics of the patients. To select the CNN model with the best performance and eliminate the interference of other factors, we conducted unified experiments on three classification models, namely, AlexNet, ResNet, and InceptionV4. We compared the cross entropy loss (CEL) of the three models and selected InceptionV4 because it has the smallest CEL. InceptionV4 was proposed in 2021 and contains three Inception parts, two Reduction parts, and one Stem part. This approach introduces a residual block to substantially accelerate the training speed and achieves dual improvements in accuracy and training speed. Finally, we combined the image features with clinical features as input to the support vector machine (SVM) classifier. All three models were constructed with torch 1.16.0. InceptionV4 was initialized randomly before training. We used the SGD optimizer to train the model with a batch size of 1 and adjusted the parameters by end-to-end supervised learning. The initial learning rate was set to 0.01, and the learning rate decreased by a factor of 10 when the accuracy of the validation set did not further improve for 10 continuous epochs each time. CEL was selected as the loss function for backpropagation. $\text{CEL} = f(x) = \frac{1}{\text{batch\_size}} \sum_{j=1}^{\text{batch\_size}} \sum_{i=1}^{n} -y_{ji} \log \widehat{y}_{ji} - (1 - y_{ji}) \log(1 - \widehat{y}_{ji})$ (n: the total category of classification; $_{ji}$: the jth sample in batch_size, classified as I; y_{ji}: actual value; $\widehat{y}_{ji}$: predictive value). During the training progress, we used the dropout strategy with a probability of 0.2 on the fully connected layers and the L2 regularization strategy for weight and bias to prevent overfitting. All steps were run in Python version 3.7. We generated the model by using a computer with an NVIDIA GeForce RTX 2080 GPU, a $6 \times$ Xeon E5-2678 v3 CPU, and 64 GB of memory.

Model Training and Validation

The patients in the training cohort ($n = 127$) composed the main cohort, and the corresponding data yielded an internal dataset for model development and cross-validation. The remaining patients in the validation cohort ($n = 64$) served as an independent cohort, and the corresponding data yielded a dataset for internal validation.

To further reduce the risk of model overfitting and make full use of our dataset, we adopted a fivefold cross-validation method to train and validate InceptionV4 in the training cohort. The method randomly divided the total existing data into five equal parts, among which four parts were used to train the whole model as the training set and the remaining part was used for testing as the testing set. In each iteration, we conducted model training on the training set, tested the well-trained model on the testing set, and calculated the accuracy. We repeated this series of steps five times and finally averaged the accuracy of the five folds to obtain the final accuracy. Additionally, the model trained in the training cohort was additionally validated in the validation cohort.

Evaluation Metrics, Comparative Methods, and Visualization Methods

We calculated the area under the ROC curve (AUC), sensitivity (SEN), specificity (SPE), positive predictive value (PPV), and negative predictive value (NPV) as evaluation metrics to analyze the diagnostic capacity of the models. We compared the performance of the InceptionV4 model against that of the following methods: (1) a clinical model that combines multiple significant clinical features (age, number of HAIC cycles, tumor thrombus, extrahepatic metastasis and AST), including tumor thrombus, extrahepatic metastasis, and AST, and (2) two skilled radiologists. Two radiologists evaluated the response to HAIC according to the following features of the HCC lesions: (1) tumor size, (2) the distance between the tumor location and the hepatic hilar, (3) tumor number, (4) tumor margin, (5) portal vein tumor thrombosis, (6) gross growth type, (7) signal intensity on T2WI, (8) signal homogeneity on T2WI, (9) arterial enhancement intensity, (10) arterial enhancement range, (11) enhancement pattern, (12) tumor capsule, (13) hemorrhage, (14) necrosis, (15) fat, and (16) hepatic capsule bulge.

To interpret the CNN predictions, we used class activation mapping (CAM) to analyze the parameters of the model.

The CAM is a tool that helps us visualize CNNs with heatmaps. Using CAM, we could clearly observe which area of the image the network focused on. After completing the training of the model, we used CAM to visualize the images and view the areas related to treatment response in heatmaps. It can be understood as the distribution of contributions to the predicted output. The higher the score, the higher the response and contribution of the corresponding region of the original image to the network, indicating the importance of each position to the classification. CAM was implemented by the package grad-cam (https://github.com/jacobgil/pytorch-grad-cam), and heatmaps were produced by using the package Matplotlib (https://pypi.org/project/matplotlib/).

Statistical Analysis

Given that some of our enrolled patients were treated with HAIC plus targeted therapy and/or immunotherapy within 2 weeks, we assigned patients to a HAIC monotherapy group and a HAIC plus targeted therapy and/or immunotherapy group. Then, we analyzed the clinical characteristics, best tumor responses, and survival differences between the two groups.

Continuous and categorical variables are expressed as medians with interquartile ranges (IQRs) and counts with percentages, respectively. Continuous and categorical variables were compared by using Student's t test, the Mann−Whitney U test, and the χ^2 test/Fisher's exact test, respectively. Variables with $P < 0.05$ in univariable analyses were entered into multivariable analyses. Univariate analyses and multivariate analyses for treatment response were performed with logistic regression models. The quantitative results of the AUC, SEN, SPE, PPV, NPV, and 95% confidence interval (CI) were calculated using the Clopper–Pearson method. Survival curves between groups were generated by Kaplan−Meier analysis and the log-rank test. All the statistical analyses were performed with R 3.5.1.

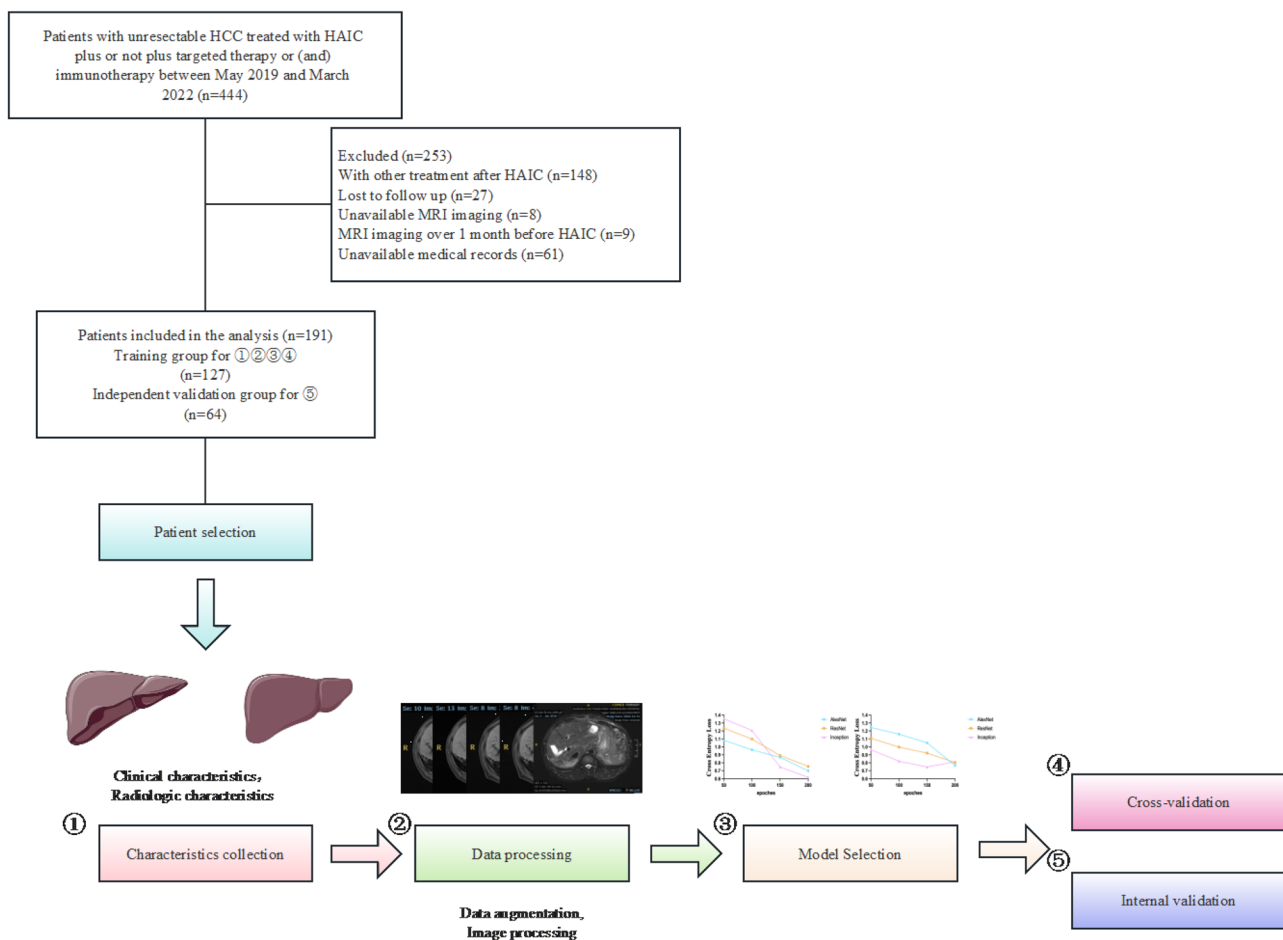The overall flowchart of the outline of our study is shown in Fig. 1.

Fig. 1 Flow chart of the outline of our study

Results

Patient Characteristics

A total of 444 patients with unresectable HCC treated with HAIC were studied, and 148 patients who accepted other treatments after HAIC were excluded. Additionally, 105 patients had incomplete data. Finally, we included 191 patients in the analysis, 127 of whom composed the training cohort and 64 of whom composed the validation cohort. The baseline characteristics of the patients are summarized in Table 1. Among all the enrolled patients, 91 patients (47.6%)

had extrahepatic metastasis, including lymph node metastasis (67 patients), pulmonary metastasis (24 patients), bone metastasis (8 patients), brain metastasis (10 patients), and implantation metastasis (6 patients). The median OS was 4.0 (2.3–7.7) months for patients in the training cohort and 5.0 (2.9–9.7) months for those in the validation cohort. For patients in the training group, the percentages of patients who achieved CR, PR, SD, and PD were 2 (1.6%), 22 (17.3%), 86 (67.7%), and 17 (13.4%), respectively. For patients in the validation cohort, the percentages of patients who achieved CR, PR, SD, and PD were 3 (4.7%), 9 (14.1%), 39 (60.9%), and 13 (20.3%), respectively.

Table 1 Clinical characteristics of the study population

Variables	Number (proportion, %) or median (IQR)		$P*$
	The training group ($n = 127$)	**The validation group ($n = 64$)**	
Age, $\geq 50/ < 50$ (years)	97/30 (76.4/23.6)	51/13 (79.7/20.3)	0.605
Sex, male/female	114/13 (89.8/10.2)	56/8 (87.5/12.5)	0.637
Pre-HALC treatment, positive/negative	98/29 (77.2/22.8)	49/15 (76.6/23.4)	0.926
No. of HAIC cycles	2 (1–3)	2 (1–3)	0.197
HAIC regimen, low dose/high dose	37/90 (29.1/70.9)	17/47 (26.6/73.4)	0.710
Tumor number, multiple/single	115/12 (90.6/9.4)	58/6 (90.6/9.4)	0.987
Maximum tumor size, cm	6.9 (4.4–10.3)	7.3 (4.5–10.7)	0.590
Tumor thrombus, positive/negative	78/49 (61.4/38.6)	41/23 (64.1/35.9)	0.722
Extrahepatic metastasis, positive/negative	63/64 (49.6/50.4)	28/36 (43.8/56.3)	0.444
BCLC stage, B/C	8/119 (14.2/85.8)	7/57 (10.9/89.1)	0.531
Child–Pugh class, A/B	89/38 (70.0/30.0)	42/22 (66.0/34.0)	0.531
HbsAg, positive/negative	101/26 (79.5/20.5)	45/19 (70.3/29.7)	0.157
AFP, ng/mL	166 (7–4081)	332 (11–3343)	0.006
PIVKA-II, mAU/mL	3898 (242–19,474)	2464 (374–12,714)	0.922
Prothrombin time, s	12.9 (12.1–13.8)	12.6 (12.0–13.4)	0.416
PLT, $\times 10^9$/L	134 (92–202)	145 (93–194)	0.559
Neutrophil, $\times 10^9$/L	3.3 (2.3–4.8)	3.4 (2.4–4.8)	0.709
Lymphocyte, $\times 10^9$/L	1.1 (0.8–1.5)	1.1 (0.8–1.4)	0.500
Monocyte, $\times 10^9$/L	0.48 (0.35–0.68)	0.47 (0.34–0.63)	0.431
NLR	3.38 (2.16–4.80)	2.98 (2.20–4.49)	0.693
PLR	122.1 (96.3–166.0)	134.9 (90.4–222.4)	0.458
Total bilirubin, μmol/L	15.8 (10.8–22.8)	16.6 (9.9–26.0)	0.376
ALT, U/L	39 (24–53)	35 (21–45)	0.469
AST, U/L	56 (34–98)	49 (36–75)	0.505
AKP, U/L	155 (113–239)	176 (122–229)	0.464
γ-GT, U/L	138 (76–248)	168 (125–271)	0.854
Albumin level, g/L	38 (35–43)	40 (37–42)	0.636
CRP, mg/L	10.4 (4.8–34.1)	15.2 (3.5–30.6)	0.731
Best tumor response, CR/PR/SD/PD	2/22/86/17 (1.6/17.3/67.7/13.4)	3/9/39/13 (4.7/14.1/60.9/20.3)	0.327
Overall survival, months	4.0 (2.3–7.7)	5.0 (2.8–9.7)	0.590
Progression free survival, months	3.0 (2.0–5.0)	3.3 (2.0–6.0)	0.455

IQR interquartile range, HAIC hepatic arterial infusion of oxaliplatin, BCLC Barcelona Clinic Liver Cancer, HbsAg hepatitis B surface antigen, AFP alpha fetoprotein, PIVKA-II protein induced by vitamin K absence or antagonist-II, PLT platelet, NLR Neutrophil to Lymphocyte ratio, PLR Platelet-to-lymphocyte ratio, ALBI albumin-bilirubin, ALT alanine aminotransferase, AST aspartate aminotransferase, AKP alkaline phosphatase, γ-GT gamma-glutamyl transpeptidase, CRP C-reactive protein, CR complete response, PR partial response, SD stable disease, PD progressive disease

Treatment-Related Adverse Events

Treatment-related adverse events, which occurred in 53.9% of patients, are presented in Supplementary Table 2. The main adverse events were elevated ALT levels/elevated AST levels (34.9%), hyperbilirubinemia (27.9%), and pain (10.1%). Serious adverse events (grades 3–4) were found in 29 (28.2%) of the 191 patients, including 2 thrombocytopenia events, 1 leucopenia event, 2 pain events, 2 upper gastrointestinal bleeding events, 14 elevated ALT/elevated AST events, and 11 hyperbilirubinemia events. Among them, three patients had elevated ALT levels/elevated AST levels as well as hyperbilirubinemia.

Clinical Predictive Factors Related to DCR

As shown in Table 2, in the training cohort, age, number of HAIC cycles, tumor thrombus, extrahepatic metastasis, lymphocyte count, neutrophil-to-lymphocyte ratio (NLR), AST,

Table 2 Univariable and multivariable logistic regression analyses of predicting disease control rate (DCR) in patients after HAIC for unresectable hepatocellular carcinoma in the training group

Variables	OR comparison	UV OR (95% CI)	UV P	MV OR (95% CI)	MV P*
Age, years	≥ 50 vs. < 50	0.281 (0.097–0.812)	**0.019**	0.165 (0.034–0.788)	**0.024**
Sex	Male vs. female	2.143 (0.525–8.744)	0.288		
Pre-HALC treatment	Positive vs. negative	0.670 (0.215–2.088)	0.490		
Number of HAIC cycles	Continuous variable	0.244 (0.103–0.574)	**0.001**	0.134 (0.033–0.538)	**0.005**
HAIC regimen	Low dose vs. high dose	3.500 (0.759–16.147)	0.108		
Tumor number	Multiple vs. single	1.778 (0.215–14.724)	0.594		
Maximum tumor size, cm	continuous variable	0.991 (0.871–1.127)	0.890		
Tumor thrombus	Positive vs. negative	5.595 (1.220–25.660)	**0.027**	11.228 (1.585–79.533)	**0.015**
Extrahepatic metastasis	Positive vs. negative	3.900 (1.196–12.715)	**0.024**	6.651 (1.177–37.588)	**0.032**
HbsAg	Positive vs. negative	1.234 (0.327–4.660)	0.757		
AFP, ng/mL	Continuous variable	1.000 (1.000–1.000)	0.766		
PIVKA-II, mAU/mL	Continuous variable	1.000 (1.000–1.000)	0.832		
Prothrombin time, s	Continuous variable	1.182 (0.829–1.688)	0.356		
PLT, $\times 10^9$/L	Continuous variable	0.995 (0.988–1.003)	0.214		
Neutrophil, $\times 10^9$/L	Continuous variable	0.965 (0.752–1.240)	0.783		
Lymphocyte, $\times 10^9$/L	Continuous variable	0.157 (0.036–0.685)	**0.014**	2.116 (0.449–9.982)	0.344
Monocyte, $\times 10^9$/L	Continuous variable	0.291 (0.029–2.926)	0.295		
NLR	Continuous variable	1.212 (1.019–1.441)	**0.029**	1.151 (0.865–1.532)	0.335
PLR	Continuous variable	1.001 (0.995–1.008)	0.660		
Total bilirubin, μmol/L	Continuous variable	0.992 (0.952–1.035)	0.722		
ALBI	3 vs. 1/2	0.631 (0.218–1.827)	0.396		
ALT, U/L	Continuous variable	1.007 (0.994–1.020)	0.269		
AST, U/L	Continuous variable	1.007 (1.002–1.012)	**0.006**	1.006 (1.000–1.013)	**0.042**
AKP, U/L	Continuous variable	1.003 (1.000–1.006)	**0.038**	1.001 (0.996–1.006)	0.751
γ-GT, U/L	Continuous variable	1.002 (1.000–1.004)	0.038		
Albumin level, g/L	Continuous variable	1.046 (0.987–1.107)	0.126		
CRP, mg/L	Continuous variable	1.005 (0.990–1.020)	0.508		

OR odds ratio, *UV* univariable, *MV* multivariable, *CI* confidence interval, *HAIC* hepatic arterial infusion of oxaliplatin, *BCLC* Barcelona Clinic Liver Cancer, *HbsAg* hepatitis B surface antigen, *AFP* alpha fetoprotein, *PIVKA-II* protein induced by vitamin K absence or antagonist-II, *PLT* platelet, *NLR* neutrophil-to-lymphocyte ratio, *PLR* platelet-to-lymphocyte ratio, *ALBI* albumin-bilirubin, *ALT* alanine aminotransferase, *AST* aspartate aminotransferase, *AKP* alkaline phosphatase, *γ-GT* gamma-glutamyl transpeptidase, *CRP* C-reactive protein

*Those variables found significant at $P < 0.05$ in univariable analyses were entered into multivariable analyses

and AKP were related to the DCR according to univariate logistic regression analysis. Multivariate logistic regression analysis confirmed age, number of HAIC cycles, tumor thrombus, extrahepatic metastasis, and AST as independent prognostic factors.

Construction of the InceptionV4 Model

We combined 27 clinical features (age, sex, pre-HALC treatment, number of HAIC cycles, HAIC regimen, tumor number, maximum tumor size, tumor thrombus, extrahepatic metastasis, HbsAg, AFP, PIVKA-II, prothrombin time, PLT, neutrophil, lymphocyte, monocyte, NLR, PLR, total bilirubin, ALBI, ALT, AST, AKP, γ-GT, albumin level and CRP) and 5 image features obtained from 5 constant sequences (enhanced T1-weighted sequence [including arterial, portal, and delayed phase], T2 FSE sequence and dual-echo sequence) to complete the classification of patient status and constructed three classical CNNs, including AlexNet, ResNet, and InceptionV4. We calculated the CEL of various models to identify the most suitable model. The CEL curves of the three models for the training and validation groups are provided in Fig. 2. We randomly captured the information of one patient from the entire group in each epoch and performed 200 epochs in total. With increasing number of epochs, the CEL of the validation cohort was stable, and the best CEL was 0.7654 according to AlexNet, 0.8032 according to ResNet, and 0.7475 according to InceptionV4. Therefore, we used the InceptionV4 model as the treatment response prediction model for subsequent training and validation. The detailed information about the architecture of InceptionV4 is shown in Fig. 3.
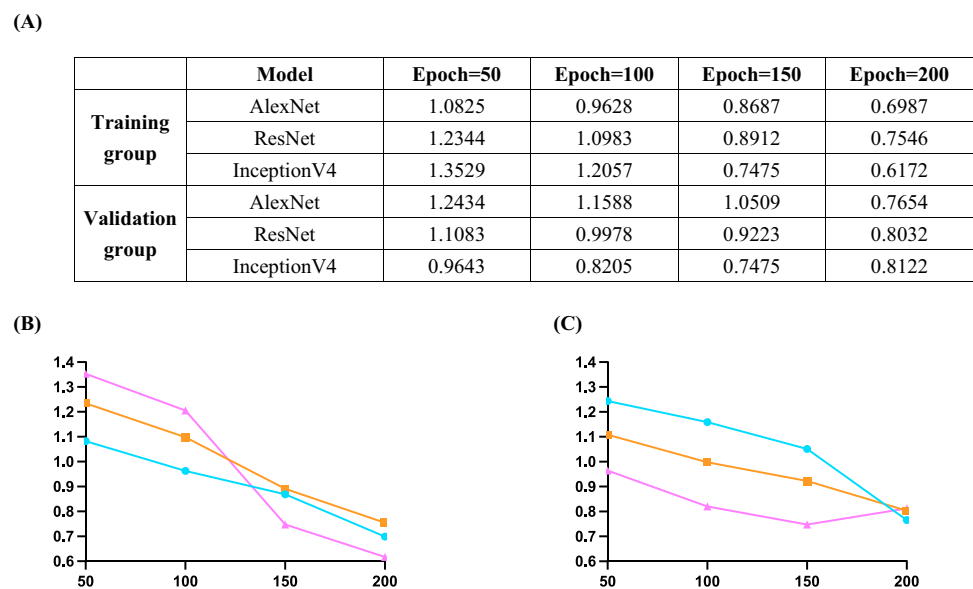
Cross-Validation Results and Internal Validation Results

In the training cohort, the AUCs of the Inception model, the clinical model, and the two radiologists were 0.871, 0.873, 0.807, and 0.783, respectively; in the validation cohort, the AUCs of these models were 0.826, 0.708, 0.806, and 0.748, respectively (Fig. 4B). Although there were no significant differences between the ROCs of the Inception model and the other models according to DeLong's test (inception model vs. clinical model, $Z = 1.39$, $P = 0.16$; inception model vs. R1, $Z = 1.43$, $P = 0.15$; inception model vs. R2, $Z = 0.72$, $P = 0.47$) in the validation cohort, the Inception model had the greatest accuracy (AUC = 0.826, SEN = 0.769, SPE = 0.882, PPV = 0.625, NPV = 0.937) in predicting antitumor responses (Table 3).

Visualization and Interpretability of the InceptionV4 Model

We artificially marked the most likely parts of the lesion in different sequences and planes. For example, as shown in the heatmaps produced by means of the CAM method (Fig. 5), we used red to indicate the most likely location of hepatic hilar tumor and portal vein tumor thrombosis. The purple and green backgrounds in the marked section reflect areas identified by the InceptionV4 model that have greater predictive significance. According to the brightness, we divided the remaining parts into blue and brown, indicating weaker predictive values. This shows that CNNs focused on the most predictive image features regarding treatment response.

Fig. 2 Cross-entropy loss of various models. **A** Cross entropy loss index of different epochs in the training group and the validation group. **B** Cross-entropy loss curves of three CNN models in the training group. **C** Cross-entropy loss curves of three CNN models in the validation group
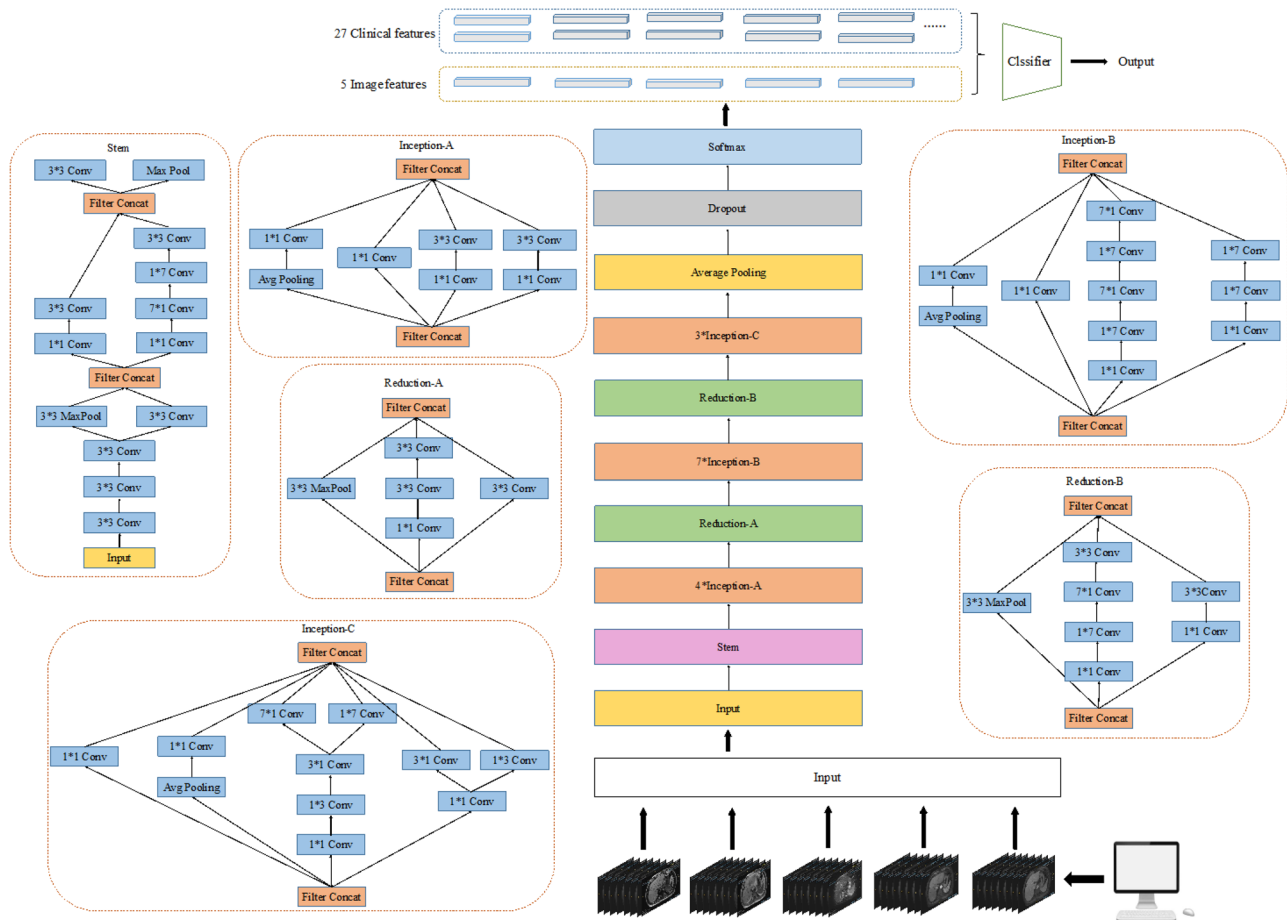
(A)

	Model	Epoch=50	Epoch=100	Epoch=150	Epoch=200
Training group	AlexNet	1.0825	0.9628	0.8687	0.6987
	ResNet	1.2344	1.0983	0.8912	0.7546
	InceptionV4	1.3529	1.2057	0.7475	0.6172
Validation group	AlexNet	1.2434	1.1588	1.0509	0.7654
	ResNet	1.1083	0.9978	0.9223	0.8032
	InceptionV4	0.9643	0.8205	0.7475	0.8122

(B)

(C)

Fig. 3 Architecture of inception

Prognosis Analysis Based on Treatment Procedure, Pre-HAIC Treatment, BCLC Stage, AFP, PIVKA-II, AST, and γ-GT

Among the 191 enrolled patients, 67 patients received HAIC alone, and 124 patients received HAIC plus targeted therapy and/or immunotherapy. The related baseline demographics and clinical characteristics in the two groups are described in Supplementary Table 3. Moreover, there were no significant differences between the two groups. Moreover, we assigned patients to several subgroups based on pre-HAIC treatment status, BCLC stage, AFP, PIVKA-II, AST, and γ-GT. The OS and PFS rates in the different subgroups were determined via Kaplan−Meier analyses with the log-rank test as shown in Supplementary Fig. 1. Patients who had an AST $\leq$ 40 had longer OS than individuals who had an AST > 40 (hazard ratio (HR) = 1.649 [95% CI, 1.017–2.674]; $P = 0.0194$). Patients who had γ-GT $\leq$ 60 had longer OS than individuals who had γ-GT > 60 (HR = 4.156 [95% CI, 1.865–9.262]; $P = 0.0190$). In addition, patients with AFP $\leq$ 400 or PIVKA-II $\leq$ 1000 had a higher ORR than did

the other patients, patients with AST $\leq$ 40 had a higher DCR than did those with AST > 40, and patients with γ-GT $\leq$ 60 had a higher ORR and DCR than did those with γ-GT > 60.

Discussion

In this study, we developed and preliminarily validated a CNN model to predict the response to HAIC in patients with unresectable HCC based on radiologic and clinical factors. The CNN model achieved an outstanding predictive performance compared to the clinical model and our radiologists, suggesting that this CNN model may provide a noninvasive and effective method for HAIC response prediction. To the best of our knowledge, this is the first study to apply the deep learning of CNNs in treatment response prediction for unresectable HCC patients who undergo HAIC.

HAIC is a widely used locoregional therapy for unresectable HCC, especially in Asia, in which chemotherapeutic agents are directly delivered into targeted vessels at increased local concentrations [32]. However, due to the

Fig. 4 ROC curves yielded by different models. **A** The cross-validation results. **B** The internal validation results
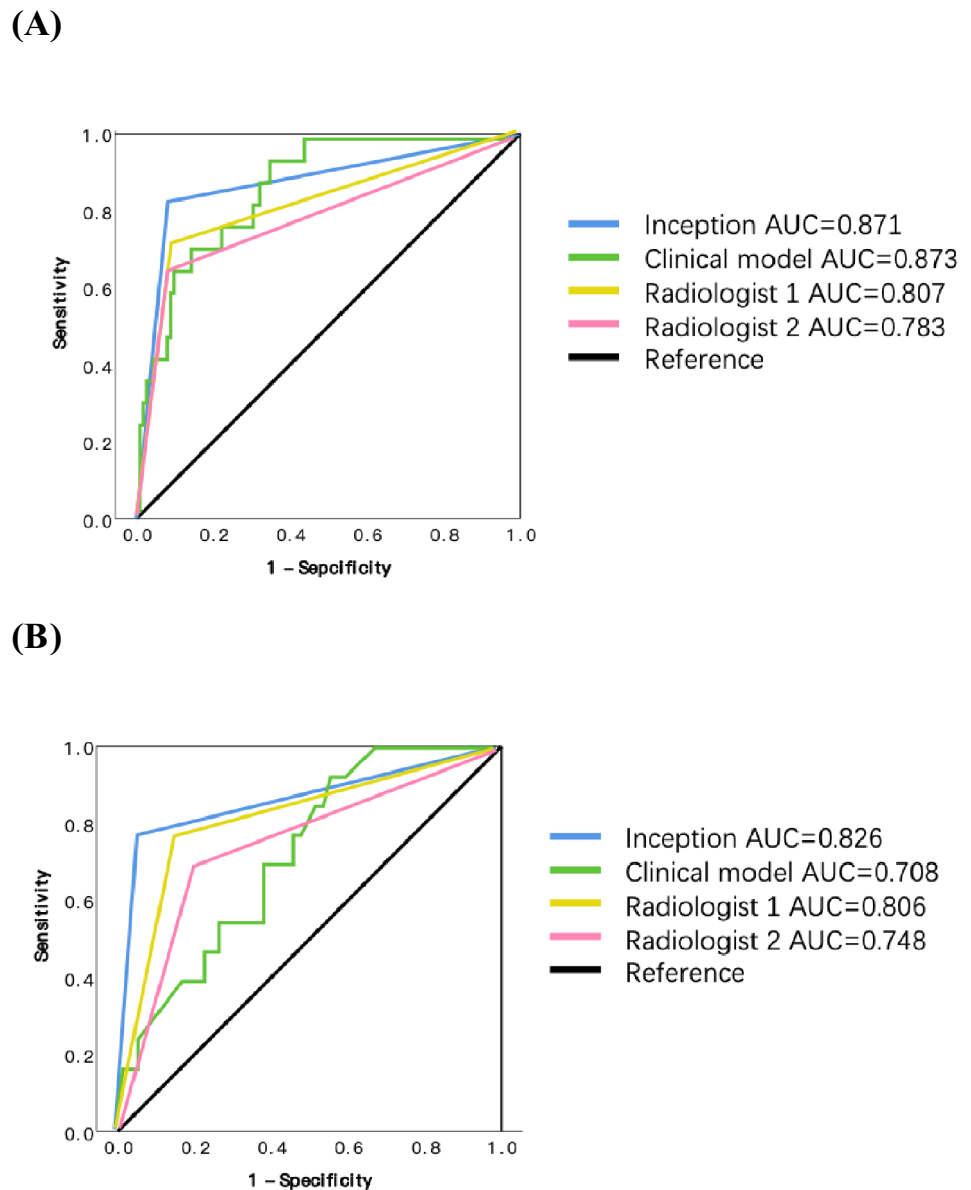
tremendous variations in tumor burden, liver function, and chemosensitivity, the treatment of HCC patients after HAIC is highly heterogeneous [33]. Therefore, a practical prognostic model based on objective measures is urgently needed to facilitate clinical decision-making.

Previous studies have shown that radiomics provides additional value to clinical factors in predicting the treatment response of HCC patients. For example, Park et al. reported that radiologic information obtained from pre-treatment CT was thought to be an important factor for predicting the therapeutic response of patients with HCC after TACE, as this information could prevent unnecessary treatment [34]. Liu et al. suggested that pretreatment MRI-based radiomics could predict the prognosis of HCC patients treated with TACE combined with MWA [35].

Miyaki et al. reported that the combination of imaging assessment and tumor marker data could be useful for predicting the early response to HAIC and the prognosis of patients with advanced HCC [36]. Therefore, our study analyzed radiomic and clinical features to predict treatment response after the HAIC procedure.

In our study, the recruited patients had a median PFS of 3.0 (2.0–5.0) months, a median OS of 4.0 (2.3–7.7) months, an ORR of 18.9%, and a DCR of 86.6% in the training cohort; additionally, they had a median PFS of 3.6 (2.0–7.1) months, a median OS of 5.5 (2.9–10.2) months, an ORR of 18.8%, and a DCR of 79.7% in the validation cohort, values which were significantly lower than those of other previous studies [18, 23, 37]. This inconsistency may be attributed to the poorer baseline characteristics and more advanced stage of the

Table 3 Performance comparison of various methods

Group	Methods	AUC	SEN	SPE	PPV	NPV
The training group	Inception	0.871 (0.761–0.981)	0.824 (0.558–0.953)	0.918 (0.846–0.960)	0.609 (0.388–0.795)	0.971 (0.912–0.993)
	Clinical model	0.873 (0.796–0.949)	0.941 (0.932–0.949)	0.651 (0.644–0.658)	0.288 (0.279–0.297)	0.987 (0.984–0.988)
	Radiologist 1	0.807 (0.675–0.940)	0.706 (0.440–0.886)	0.909 (0.835–0.953)	0.545 (0.327–0.749)	0.952 (0.887–0.982)
	Radiologist 2	0.783 (0.642–0.923)	0.647 (0.386–0.847)	0.918 (0.846–0.960)	0.550 (0.320–0.762)	0.944 (0.877–0.977)
The validation group	Inception	0.826 (0.682–0.970)	0.769 (0.460–0.938)	0.882 (0.754–0.951)	0.625 (0.359–0.837)	0.937 (0.818–0.984)
	Clinical model	0.708 (0.567–0.850)	0.923 (0.778–1.000)	0.431 (0.295–0.567)	0.293 (0.153–0.432)	0.957 (0.873–1.000)
	Radiologist 1	0.806 (0.660–0.952)	0.769 (0.460–0.938)	0.843 (0.709–0.925)	0.556 (0.313–0.776)	0.935 (0.811–0.983)
	Radiologist 2	0.748 (0.588–0.908)	0.692 (0.389–0.896)	0.804 (0.665–0.897)	0.474 (0.252–0.705)	0.911 (0.779–0.971)

AUC area under the receiver operating characteristic curve, *SEN* sensitivity, *SPE* specificity, *PPV* positive predictive value, *NPV* negative predictive value, *AST* aspartate aminotransferase

recruited patients, which included tumor thrombus and extrahepatic metastasis. Moreover, more than half of the patients in our study had already received various therapies prior to HAIC, such as traditional TACE procedures. These therapies may increase the risk of decompensated liver function and damage the tumor-feeding artery, which impairs the delivery and perfusion of arterial infusion chemotherapy. Notably, since SD is also a result of effective treatment in our study, it is more accurate to consider patients who achieved CR, PR, or SD as having a positive treatment response.

Complex models, such as CNNs, can be affected by the problem of overfitting, which results in poor performance for data not included in the training process. Therefore, we divided the whole dataset into a training group and a validation group. Then, we used MRI images based on several constant sequences (enhanced T1-weighted sequences [including arterial, portal, and delayed phases]; T2 FSE sequence; and dual-echo sequence) to construct three representative CNN models, AlexNet, ResNet, and InceptionV4, in the training group. We selected the most appropriate model by CEL. The best-performing model yielded satisfactory predictions in the validation cohort, with a CEL of 0.7475 obtained by InceptionV4 at the 150th epoch. Finally, we combined the image features with clinical features as the total features of each patient in the SVM classifier because SVM has superior performance in classifying gene features and low sample size data [38]. Our model yielded satisfactory performance in the validation cohort, with an AUC of 0.826, an SEN of 0.769, an SPE of 0.882, a PPV of 0.609, and an NPV of 0.971.

Furthermore, we identified variables related to treatment response by logistic regression analysis in the training cohort.

Age, number of HAIC cycles, tumor thrombus, extrahepatic metastasis, and AST (continuous variable) were found to be independent predictors. Our model had a greater AUC (AUC = 0.826) than did the clinical model (including the above five significant clinical features) and two experienced radiologists in the validation cohort. Additionally, it had optimal accuracy (SEN, SPE, PPV, NPV). Therefore, we believe that a combined model could be a more powerful predictor.

Recently, great progress has been made in deep learning, allowing machines to discover intricate structures in large datasets and change their internal parameters [29]. CNN has been applied to establish a direct connection between disease prediction and diagnostic images [39]. Conventional machine learning requires features from images to be extracted before learning, while CNNs allow the image itself to be used during the training process [29]. Therefore, CNNs enable all the information contained in images to be used, although this approach can be limited by the feature parameters chosen in conventional machine learning. However, deep learning models are black boxes, and future studies are needed to compare the efficiency of automated AI and handcrafted feature extraction methods.

Notably, in our study, there were no significant differences between the HAIC group and the HAIC combined with molecular targeted therapy/immunotherapy group regarding the best tumor response, OS or PFS. This is likely due to the increased risk of liver dysfunction caused by aggressive HAIC combination treatment that might compromise the potential survival benefits of initial treatment. According to previous studies, the survival benefit of HAIC combined with sorafenib in patients with unresectable HCC is controversial [19; 40]. Kudo et al. compared sorafenib
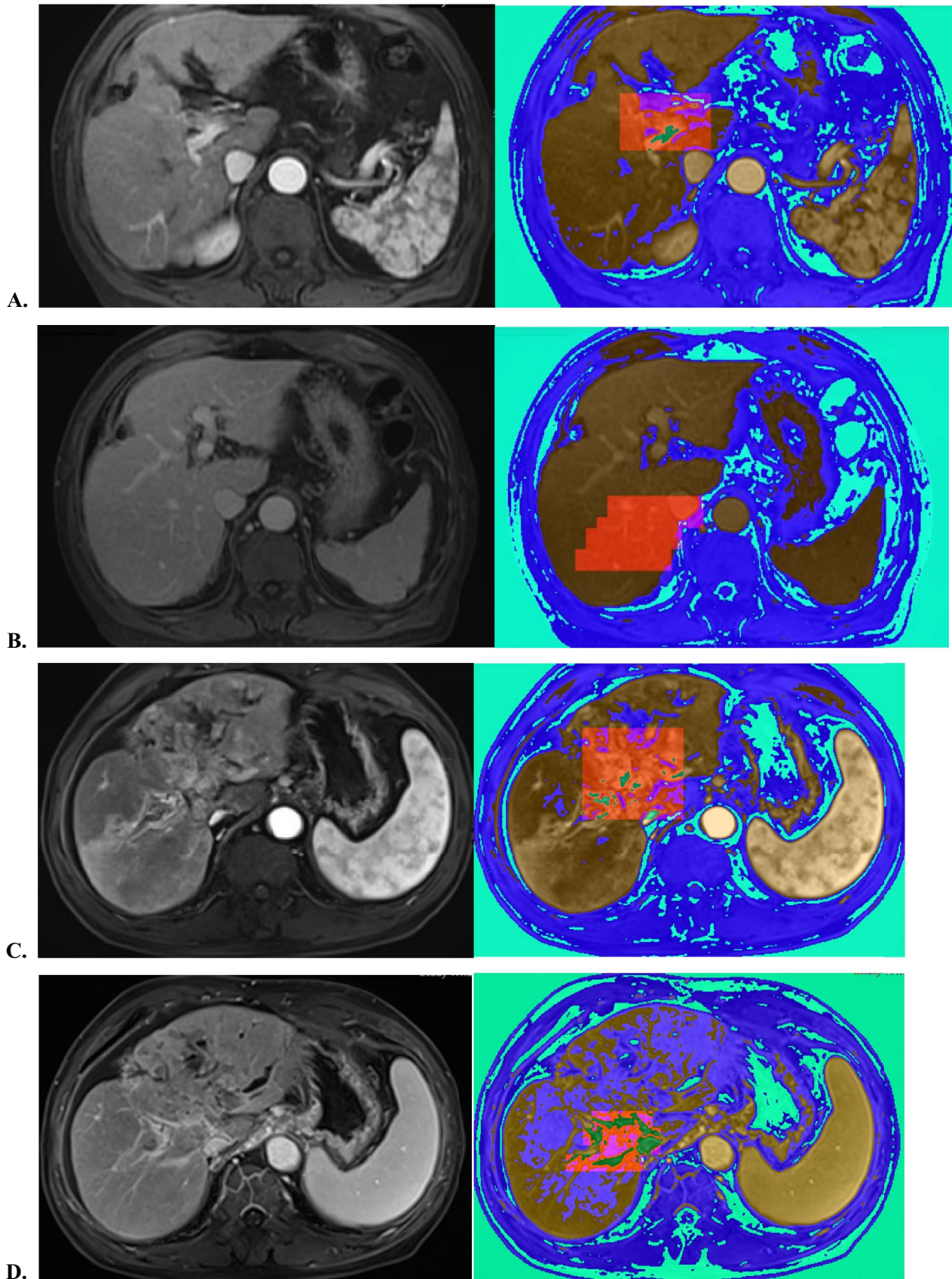
1294

Journal of Imaging Informatics in Medicine (2024) 37:1282–1296

◄**Fig. 5** Visualization of the InceptionV4 model. MRI images and heat maps of two HCC patients treated after HAIC. **A**, **B** Images in a 65-year-old man with best tumor response of CR, time to PD of 8 months. **A** Arterial phase image shows that HCC is not close to hepatic hilar. **B** Delayed phase image shows that the HCC patient is without portal vein tumor thrombosis or hepatic bile duct dilation. **C**, **D** Images in an 81-year-old man with best tumor response of PD, time to PD of 1 month. **C** Arterial phase image shows that HCC is close to hepatic hilar. **D** Delayed phase image shows that the HCC patient is with portal vein tumor thrombosis and hepatic bile duct dilation

with or without low-dose cisplatin and fluorouracil (HAIC) in a randomized phase III trial of advanced HCC patients [40]. The median OS was similar between patients treated with sorafenib plus HAIC and patients treated with sorafenib monotherapy (HR = 1.009 [0.743–1.371], $P = 0.955$). Thus, the optimization of HAIC combination treatment warrants future clinical studies in unresectable HCC patients. Our subgroup analyses suggested that the AST and γ-GT correlated with better OS outcomes. Low AST and γ-GT levels are correlated with better liver functional reserve.

Our study has several limitations. First, this was a retrospective study involving a single center with limited data, and selection bias was unavoidable. However, further external validation studies with larger prospective studies are needed before the findings can be applied to clinical use. Second, most patients enrolled in our study had hepatitis B virus–related HCC, whereas different ethnic backgrounds or common risk factors for HCC, such as alcoholic liver disease and HCV-related HCC, were not specifically evaluated in Western countries. Third, all the images were obtained from the same scanner, and our results require further validation with other MRI vendors to determine the accuracy of the findings. However, this is a pilot study in which deep learning was applied for treatment response prediction in patients with unresectable HCC after HAIC, resulting in a more accurate and reproducible prognostic evaluation than qualitative reasoning.

In conclusion, we prove that a CNN-based model can predict with good performance the treatment response in patients with unresectable HCC who undergo HAIC. The model may be an effective strategy to help physicians identify appropriate candidates for HAIC in unresectable HCC patients. With further validation, our CNN-based model has great potential to serve as a practical decision support tool in clinical applications.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10278-024-01003-2.

Author Contribution Bing Quan, Jinghuan Li and Hailin Mi contributed to this work equally. Bing Quan: writing — original draft (lead). Jinghuan Li: methodology (lead). Hailin Mi: software (lead); validation (lead). Miao Li, Wenfeng Liu, Fan Yao: data curation (equal); Yan Shan, Pengju Xu: investigation (equal); visualization (equal). Zhenggang Ren, Xin Yin: conceptualization (lead); writing — review and editing (lead).

Funding Funding for the study was provided by the National Natural Science Foundation of China (No. 81972889, Dr Yin; No. 81802320, Dr Li).

Availability of Data and Materials The data used in the article are available with reasonable request from the corresponding author.

Declarations

Ethics Approval Institutional Review Board approval was obtained.

Consent to Participate Informed consent was obtained from all individual participants included in the study.

Consent for Publication The authors affirm that human research participants provided informed consent for publication of the information in this study.

Competing Interests The authors declare no competing interests.

References

1. Asrani SK, Devarbhavi H, Eaton J, Kamath PS (2019) Burden of liver diseases in the world. J Hepatol 70:151-171
2. Roayaie S, Jibara G, Tabrizian P et al (2015) The role of hepatic resection in the treatment of hepatocellular cancer. Hepatology 62:440-451
3. Llovet JM, Ricci S, Mazzaferro V et al (2008) Sorafenib in advanced hepatocellular carcinoma. N Engl J Med 359:378-390
4. Kudo M, Finn RS, Qin S et al (2018) Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet 391:1163-1173
5. Qin S, Bi F, Gu S et al (2021) Donafenib Versus Sorafenib in First-Line Treatment of Unresectable or Metastatic Hepatocellular Carcinoma: A Randomized, Open-Label, Parallel-Controlled Phase II-III Trial. J Clin Oncol 39:3002-3011
6. Finn RS, Qin S, Ikeda M et al (2020) Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. N Engl J Med 382:1894-1905
7. Qin S, Kudo M, Meyer T et al (2023) Tislelizumab vs Sorafenib as First-Line Treatment for Unresectable Hepatocellular Carcinoma: A Phase 3 Randomized Clinical Trial. JAMA Oncol. https://doi.org/10.1001/jamaoncol.2023.4003
8. Llovet JM, Real MI, Montaña X et al (2002) Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: a randomised controlled trial. Lancet 359:1734-1739
9. Lo CM, Ngan H, Tso WK et al (2002) Randomized controlled trial of transarterial lipiodol chemoembolization for unresectable hepatocellular carcinoma. Hepatology 35:1164-1171
10. Pinter M, Hucke F, Graziadei I et al (2012) Advanced-stage hepatocellular carcinoma: transarterial chemoembolization versus sorafenib. Radiology 263:590-599
11. Ota H, Nagano H, Sakon M et al (2005) Treatment of hepatocellular carcinoma with major portal vein thrombosis by combined therapy with subcutaneous interferon-alpha and intra-arterial 5-fluorouracil; role of type 1 interferon receptor expression. Br J Cancer 93:557-564
12. Yoshikawa M, Ono N, Yodono H, Ichida T, Nakamura H (2008) Phase II study of hepatic arterial infusion of a fine-powder formulation of cisplatin for advanced hepatocellular carcinoma. Hepatol Res 38:474-483

Springer

13. Monden M, Sakon M, Sakata Y, Ueda Y, Hashimura E (2012) 5-fluorouracil arterial infusion + interferon therapy for highly advanced hepatocellular carcinoma: A multicenter, randomized, phase II study. Hepatol Res 42:150-165

14. Nouso K, Miyahara K, Uchida D et al (2013) Effect of hepatic arterial infusion chemotherapy of 5-fluorouracil and cisplatin for advanced hepatocellular carcinoma in the Nationwide Survey of Primary Liver Cancer in Japan. Br J Cancer 109:1904-1907

15. Ueshima K, Ogasawara S, Ikeda M et al (2020) Hepatic Arterial Infusion Chemotherapy versus Sorafenib in Patients with Advanced Hepatocellular Carcinoma. Liver Cancer 9:583-595

16. Kudo M, Kawamura Y, Hasegawa K et al (2021) Management of Hepatocellular Carcinoma in Japan: JSH Consensus Statements and Recommendations 2021 Update. Liver Cancer 10:181-223

17. Chen LT, Martinelli E, Cheng AL et al (2020) Pan-Asian adapted ESMO Clinical Practice Guidelines for the management of patients with intermediate and advanced/relapsed hepatocellular carcinoma: a TOS-ESMO initiative endorsed by CSCO, ISMPO, JSMO, KSMO, MOS and SSO. Ann Oncol 31:334-351

18. Lyu N, Kong Y, Mu L et al (2018) Hepatic arterial infusion of oxaliplatin plus fluorouracil/leucovorin vs. sorafenib for advanced hepatocellular carcinoma. J Hepatol 69:60-69

19. He M, Li Q, Zou R et al (2019) Sorafenib Plus Hepatic Arterial Infusion of Oxaliplatin, Fluorouracil, and Leucovorin vs Sorafenib Alone for Hepatocellular Carcinoma With Portal Vein Invasion: A Randomized Clinical Trial. JAMA Oncol 5:953-960

20. Li QJ, He MK, Chen HW et al (2022) Hepatic Arterial Infusion of Oxaliplatin, Fluorouracil, and Leucovorin Versus Transarterial Chemoembolization for Large Hepatocellular Carcinoma: A Randomized Phase III Trial. J Clin Oncol 40:150-160

21. Zheng K, Zhu X, Fu S et al (2022) Sorafenib Plus Hepatic Arterial Infusion Chemotherapy versus Sorafenib for Hepatocellular Carcinoma with Major Portal Vein Tumor Thrombosis: A Randomized Trial. Radiology 303:455-464

22. Hsu SJ, Xu X, Chen MP et al (2021) Hepatic Arterial Infusion Chemotherapy with Modified FOLFOX as an Alternative Treatment Option in Advanced Hepatocellular Carcinoma Patients with Failed or Unsuitability for Transarterial Chemoembolization. Acad Radiol 28 Suppl 1:S157-s166

23. Hu J, Bao Q, Cao G et al (2020) Hepatic Arterial Infusion Chemotherapy Using Oxaliplatin Plus 5-Fluorouracil Versus Transarterial Chemoembolization/Embolization for the Treatment of Advanced Hepatocellular Carcinoma with Major Portal Vein Tumor Thrombosis. Cardiovasc Intervent Radiol 43:996-1005

24. Liu J, Zhang J, Wang Y, Shu G, Lou C, Du Z (2022) HAIC versus TACE for patients with unresectable hepatocellular carcinoma: A systematic review and meta-analysis. Medicine (Baltimore) 101:e32390

25. He X, Li K, Wei R et al (2023) A multitask deep learning radiomics model for predicting the macrotrabecular-massive subtype and prognosis of hepatocellular carcinoma after hepatic arterial infusion chemotherapy. Radiol Med. https://doi.org/10.1007/s11547-023-01719-1

26. Xu Z, An C, Shi F et al (2023) Automatic prediction of hepatic arterial infusion chemotherapy response in advanced hepatocellular carcinoma with deep learning radiomic nomogram. Eur Radiol. https://doi.org/10.1007/s00330-023-09953-x

27. Zhao Y, Huang F, Liu S et al (2023) Prediction of therapeutic response of unresectable hepatocellular carcinoma to hepatic arterial infusion chemotherapy based on pretherapeutic MRI radiomics and Albumin-Bilirubin score. J Cancer Res Clin Oncol 149:5181-5192

28. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H (2018) Artificial intelligence in radiology. Nat Rev Cancer 18:500-510

29. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436-444

30. Huang H, Hu X, Zhao Y et al (2018) Modeling Task fMRI Data Via Deep Convolutional Autoencoder. IEEE Trans Med Imaging 37:1551-1561

31. Eisenhauer EA, Therasse P, Bogaerts J et al (2009) New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 45:228-247

32. Bartkowski R, Berger MR, Aguiar JL et al (1986) Experiments on the efficacy and toxicity of locoregional chemotherapy of liver tumors with 5-fluoro-2'-deoxyuridine (FUDR) and 5-fluorouracil (5-FU) in an animal model. J Cancer Res Clin Oncol 111:42-46

33. Obi S, Sato S, Kawai T (2015) Current Status of Hepatic Arterial Infusion Chemotherapy. Liver Cancer 4:188-199

34. Park HJ, Kim JH, Choi SY et al (2017) Prediction of Therapeutic Response of Hepatocellular Carcinoma to Transcatheter Arterial Chemoembolization Based on Pretherapeutic Dynamic CT and Textural Findings. AJR Am J Roentgenol 209:W211-w220

35. Liu J, Pei Y, Zhang Y, Wu Y, Liu F, Gu S (2021) Predicting the prognosis of hepatocellular carcinoma with the treatment of transcatheter arterial chemoembolization combined with microwave ablation using pretreatment MR imaging texture features. Abdom Radiol (NY) 46:3748-3757

36. Miyaki D, Kawaoka T, Aikata H et al (2015) Evaluation of early response to hepatic arterial infusion chemotherapy in patients with advanced hepatocellular carcinoma using the combination of response evaluation criteria in solid tumors and tumor markers. J Gastroenterol Hepatol 30:726-732

37. Lyu N, Kong Y, Pan T et al (2019) Hepatic Arterial Infusion of Oxaliplatin, Fluorouracil, and Leucovorin in Hepatocellular Cancer with Extrahepatic Spread. J Vasc Interv Radiol 30:349-357.e342

38. Golub TR, Slonim DK, Tamayo P et al (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286:531-537

39. Liu B, Zeng Q, Huang J et al (2022) IVIM using convolutional neural networks predicts microvascular invasion in HCC. Eur Radiol 32:7185-7195

40. Kudo M, Ueshima K, Yokosuka O et al (2018) Sorafenib plus low-dose cisplatin and fluorouracil hepatic arterial infusion chemotherapy versus sorafenib alone in patients with advanced hepatocellular carcinoma (SILIUS): a randomised, open label, phase 3 trial. Lancet Gastroenterol Hepatol 3:424-432

Authors and Affiliations

Bing Quan[1,2] · **Jinghuan Li**[1,2] · **Hailin Mi**[4] · **Miao Li**[1,2] · **Wenfeng Liu**[1,2] · **Fan Yao**[1,2] · **Rongxin Chen**[1,2] · **Yan Shan**[3] · **Pengju Xu**[3] · **Zhenggang Ren**[1,2] · **Xin Yin**[1,2]

✉ Xin Yin
yin.xin@zs-hospital.sh.cn

Bing Quan
quanbingfu2021@163.com

Jinghuan Li
li.jinghuan@zs-hospital.sh.cn

Hailin Mi
mihailin@hrbeu.edu.cn

Miao Li
li.miao@zs-hospital.sh.cn

Wenfeng Liu
liuwenfengdoc@163.com

Fan Yao
yaofan1021@163.com

Rongxin Chen
chen.rongxin@zs-hospital.sh.cn

Yan Shan
shan.yan@zs-hospital.sh.cn

Pengju Xu
xu.pengju@zs-hospital.sh.cn

Zhenggang Ren
ren.zhenggang@zs-hospital.sh.cn

1   Liver Cancer Institute, Zhongshan Hospital, Fudan University, Shanghai 200032, China

2   National Clinical Research Center for Interventional Medicine, Shanghai 200032, China

3   Department of Radiology, Zhongshan Hospital, Fudan University, 136 Yi Xue Yuan Road, Shanghai 200032, China

4   Department of Computer Science and Technology, Harbin Engineering University, Harbin, China