Cancer Science

WILEY

Cancer Science

ORIGINAL ARTICLE OPEN ACCESS

A Contrast-Enhanced Ultrasound Cine-Based Deep Learning Model for Predicting the Response of Advanced Hepatocellular Carcinoma to Hepatic Arterial Infusion Chemotherapy Combined With Systemic Therapies

Xu Han[1] | Chuan Peng[1] | Si-Min Ruan[2] | Lingling Li[1] | Minke He[3] | Ming Shi[3] | Bin Huang[4] | Yudi Luo[4] | Jingming Liu[4] | Huiying Wen[4] | Wei Wang[2] | Jianhua Zhou[1] | Minhua Lu[4] | Xin Chen[4] | Ruhai Zou[1] | Zhong Liu[4]

[1]State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Department of Ultrasound, Sun Yat-Sen University Cancer Center, Guangzhou, China | [2]Department of Medical Ultrasonics, Ultrasomics Artificial Intelligence X-Lab, Institute of Diagnostic and Interventional Ultrasound, the First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China | [3]State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Department of Hepatobiliary Oncology, Sun Yat-Sen University Cancer Center, Guangzhou, China | [4]National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen University, Shenzhen, China

Correspondence: Ruhai Zou (zourh@sysucc.org.cn) | Zhong Liu (liuzhong@szu.edu.cn)

Received: 30 November 2024 | **Revised:** 1 April 2025 | **Accepted:** 18 April 2025

Funding: This work was supported by the Basic and Applied Basic Research Foundation of Guangdong Province under Grant 2023A1515110202, and 2025A1515011934; the National Natural Science Foundation of China under Grants 82171957, 82272014, 82327804, 62471315, and 61901282; Science & Technology Projects in Guangzhou under Grant 202102080249; Shenzhen Basic Science Research (Key Program) under Grant JCYJ20220818095612027, and JCYJ20241202124228035; and Medicine-Engineering Interdisplinary Funding of Shenzhen University.

Keywords: combination therapy | contrast-enhanced ultrasound | deep learning | hepatic arterial infusion chemotherapy | hepatocellular carcinoma

ABSTRACT

Recently, a hepatic arterial infusion chemotherapy (HAIC)-associated combination therapeutic regimen, comprising HAIC and systemic therapies (molecular targeted therapy plus immunotherapy), referred to as HAIC combination therapy, has demonstrated promising anticancer effects. Identifying individuals who may potentially benefit from HAIC combination therapy could contribute to improved treatment decision-making for patients with advanced hepatocellular carcinoma (HCC). This dual-center study was a retrospective analysis of prospectively collected data with advanced HCC patients who underwent HAIC combination therapy and pretreatment contrast-enhanced ultrasound (CEUS) evaluations from March 2019 to March 2023. Two deep learning models, AE-3DNet and 3DNet, along with a time-intensity curve-based model, were developed for predicting therapeutic responses from pretreatment CEUS cine images. Diagnostic metrics, including the area under the receiver-operating-characteristic curve (AUC), were calculated to compare the performance of the models. Survival analysis was used to assess the relationship between predicted responses and prognostic outcomes. The model of AE-3DNet was constructed on the top of 3DNet, with innovative incorporation of spatiotemporal attention modules to enhance the capacity for dynamic feature extraction. 326 patients were included, 243 of whom

Abbreviations: 3D Conv, three-dimensional convolution layer; 3D, instance normalization; 3D, Instance-Norm; 3DNet, three-dimensional convolutional neural network; ACC, accuracy; AUC, area under the receiver-operating-characteristic curve; BCLC, Barcelona Clinic Liver Cancer; CBs, convolutional blocks; CEUS, contrast-enhanced ultrasound; CIs, confidence intervals; FC, fully connected; Grad-CAM, gradient-weighted class activation mapping; HAIC, hepatic arterial infusion chemotherapy; HCC, hepatocellular carcinoma; mRECIST, Modified Response Evaluation Criteria in Solid Tumors; OR, objective response; ReLU, rectified linear unit; ROC, receiver-operating-characteristic; SEN, sensitivity; SPE, specificity; STAMs, spatiotemporal attention modules; TACE, transarterial chemoembolization; TIC, time-intensity curve.

Xu Han, Chuan Peng, Si-Min Ruan and Lingling Li contributed equally to this work and are joint first authors.

formed the internal validation cohort, which was utilized for model development and fivefold cross-validation, while the rest formed the external validation cohort. Objective response (OR) or non-objective response (non-OR) were observed in 63% (206/326) and 37% (120/326) of the participants, respectively. Among the three efficacy prediction models assessed, AE-3DNet performed superiorly with AUC values of 0.84 and 0.85 in the internal and external validation cohorts, respectively. AE-3DNet's predicted response survival curves closely resembled actual clinical outcomes. The deep learning model of AE-3DNet developed based on pretreatment CEUS cine performed satisfactorily in predicting the responses of advanced HCC to HAIC combination therapy, which may serve as a promising tool for guiding combined therapy and individualized treatment strategies.
Trial Registration: NCT02973685.

1 | Introduction

Primary liver cancer ranks as the sixth most commonly diagnosed cancer and the third leading cause of cancer-related deaths globally [1]. Hepatocellular carcinoma (HCC) accounts for the majority (75%–85%) of primary liver cancer cases, with prognosis and treatment strategies varying significantly across stratified stages. Early-stage HCC typically has a favorable prognosis, with curative treatment options such as surgical resection, thermal ablation, and liver transplantation available. By contrast, advanced HCC exhibits a poor prognosis, and research continues to identify more effective therapies for improving the prognosis. Recent advancements in systemic therapy [2] and hepatic arterial infusion chemotherapy (HAIC) [3, 4] have spurred investigations into combining these approaches to improve outcomes for advanced HCC. Particularly, a HAIC-associated combination therapeutic regimen, comprising HAIC (with drugs of oxaliplatin, fluorouracil, and leucovorin, FOLFOX) and systemic therapies (molecular targeted therapy plus anti-PD-1 [programmed cell death-1] immunotherapy), referred to as HAIC combination therapy, has demonstrated promising anticancer effects, with objective response rates ranging from 54.1% to 77.1% in patients with advanced HCC [5–10]. With safety confirmation, an increasing number of clinicians are considering HAIC combination therapy as a viable option to enhance patient prognosis [11–13]. However, there are still at least 20% of patients who fail to benefit from the combination therapy or even suffer from heightened toxic side effects. Therefore, it is imperative to develop methods for predicting the efficacy of HAIC combination therapy and prognosis before treatment initiation, such that individuals who may potentially benefit from the combination therapy could be identified for personalized management and ultimately improving the prognosis.

Abnormal vascular structure and blood supply in advanced HCC could significantly impact the effectiveness of anticancer therapies. Contrast-enhanced ultrasound (CEUS) provides high-resolution imaging of the vascular structure and real-time monitoring of tumor blood supply through contrast agent infusion [14], and thus it would be promising to predict the therapeutic efficacy by analyzing the hepatic CEUS cine data before treatment initiation. Currently, the analysis of hepatic CEUS cine data predominantly focuses on parameters such as enhancement patterns, perfusion intensity, time to washout, and degree of washout, which are often derived by naked-eye observation or employing an engineering method based on time-intensity curve (TIC) analysis [15]. However, CEUS cine data contain a wealth of information that may surpass human perception or exceed the capabilities of TIC analysis. A comprehensive exploration of this hidden information could contribute to improving predictions regarding therapeutic efficacy. Recent advances in deep learning technologies have empowered the exploration of more latent crucial details within CEUS cines by deep and supervised learning on the data. In this perspective, deep learning models based on pretreatment CEUS cine images have been established for predicting the response of HCC to various therapeutic regimens, including HAIC [16–21], systemic therapy [22] or transarterial chemoembolization (TACE) [19]. Of note, the clinical application of these predictive models has the potential to facilitate personalized treatments for patients with advanced HCC, ultimately improving prognosis. However, there remains a paucity of studies predicting the efficacy of HAIC combination therapy prior to treatment initiation.

Therefore, this study aims to establish and validate a deep learning model based on pretreatment CEUS cine images to predict the response of advanced HCC to HAIC combination therapy prior to treatment initiation. This approach seeks to identify advanced-HCC patients who are likely to benefit from the combination therapy, facilitating personalized treatment plans and improving clinical outcomes.

2 | Methods

2.1 | Patient Cohorts

The work has been reported in line with the STARD (Standards for the Reporting of Diagnostic accuracy studies) criteria [23]. The patients included in this dual-center retrospective study were derived from clinical trial registration no. NCT02973685 and adhered to the ethical review standards of Sun Yat-sen University Cancer Center (Hospital Center A) and the First Affiliated Hospital of Sun Yat-Sen University (Hospital Center B). Patients who received HAIC combination therapy from March 2019 to March 2023 at either of these two hospital centers were evaluated for eligibility. The inclusion criteria were as follows: unresectable advanced HCC confirmed by contrasted-enhanced CT/MRI and histopathological examination before treatment; Barcelona Clinic Liver Cancer (BCLC) stage B or C; receiving therapeutic treatment for the first time; and undergoing CEUS examination within one month before treatment initiation. The exclusion criteria included the following: diffuse liver cancer without definite lesion boundaries, poor CEUS image quality, or incomplete clinical information. Poor CEUS image quality was defined as significant acoustic attenuation caused by factors such as anomalous anatomical positioning of the lesion, acoustic shadowing from overlying structures, or suboptimal patient cooperation, resulting in inadequate

visualization of the target lesion. Patients from Hospital Center A served as the internal validation cohort for model deviation and cross-validation, while those from the other center served as the external cohort for external validation of the model derived from the internal validation cohort. More details on patient screening are shown in Figure 1. All enrolled patients and the associated data were assessed by the same surgical team in Hospital Center A.

2.2 | Therapeutic Protocols

Patients received FOLFOX-HAIC, targeted therapy (apatinib) plus immunotherapy (camilizumab) as the therapeutic strategy. FOLFOX-HAIC was administered as follows: Firstly, an initial round of HAIC with FOLFOX was conducted by following the femoral artery approach [24]. A 3.5 French catheter was utilized for super-selective catheterization of relevant arteries, and a 2.7 French microcatheter was meticulously positioned into the feeding arteries of the tumor and tumor thrombus, using arteriography to identify the tumor's arterial supply. Following the first HAIC circle, patients were transferred to the ward and confined to bed for 48 h, and then a second round of HAIC was conducted. For this circle, a marked microcatheter was connected to an arterial infusion pump for the administration of chemotherapy agents: oxaliplatin 85 mg/m^2, leucovorin 400 mg/m^2, fluorouracil bolus 400 mg/m^2 on day 1, and fluorouracil infusion 2400 mg/m^2 for 46 h. In addition to the HAIC treatments, all patients received anti-PD-1 immunotherapy (camilizumab, 200 mg, repeated every 21 days from day 4 of the second HAIC cycle) and targeted therapy (apatinib, 250 mg, orally daily, starting on day 8 of the initial HAIC cycle).

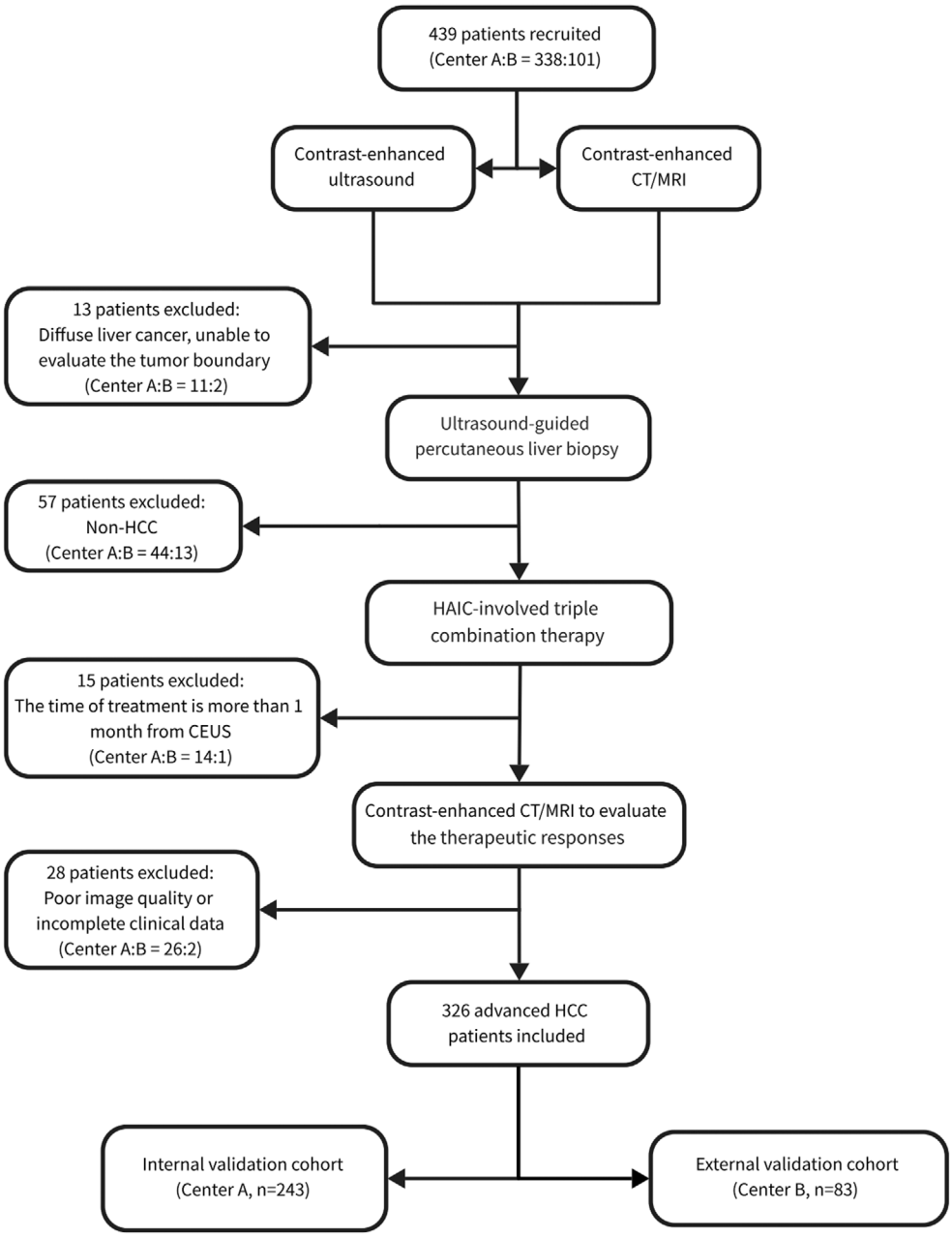

FIGURE 1 | Workflow for patient screening.

2.3 | Patient Follow-Up and Response Assessment

All patients were followed up for at least one year, with the follow-up period ending in March 2024. Tumor responses to the HAIC combination therapy were confirmed by contrast-enhanced CT/MRI at 12–16 weeks after treatment. Two independent radiologists assessed the tumor response according to the guidelines of modified Response Evaluation Criteria in Solid Tumors (mRECIST) [25]. Objective response (OR) was defined as complete response or partial response, while non-objective response (non-OR) was defined as stable disease or progressive disease. In cases of discrepancies between the radiologists' assessments, discussions were conducted to reach a consensus.

2.4 | CEUS Examinations

CEUS examinations were conducted within one month prior to therapy by senior radiologists, each possessing a minimum of 10 years of experience in liver CEUS. The examinations were performed using an ultrasound machine (Siemens Acuson Sequoia) equipped with a 5C1 probe. For the CEUS examination of each patient, a bolus of the second-generation contrast agents (2.0 mL, SonoVue, Bracco Imaging) was injected through the antecubital vein within 1–2 s, followed by a 5 mL saline flush. A timer was started immediately after the injection of contrast agents, and a dynamic digital video with a dual-view display (demonstrated in Figure 2a) of CEUS and B-mode images of the target lesion and the surrounding tissues was continuously observed and recorded for at least 90 s. All video data were stored in Digital Imaging and Communications in Medicine (DICOM) format. In cases of multiple lesions, preference was given to recording the larger and more clearly visualized lesion.

2.5 | Data Preprocessing

The CEUS video data from each patient were firstly decomposed into consecutive frames using the native function of the MicroDicom DICOM viewer 2.8.3. Then, image patches corresponding to the region of interest (ROI, including tumor and peritumoral regions) were extracted from the image sequence. The extraction of ROIs was conducted by a radiologist who was blinded to the histopathological results of the patient

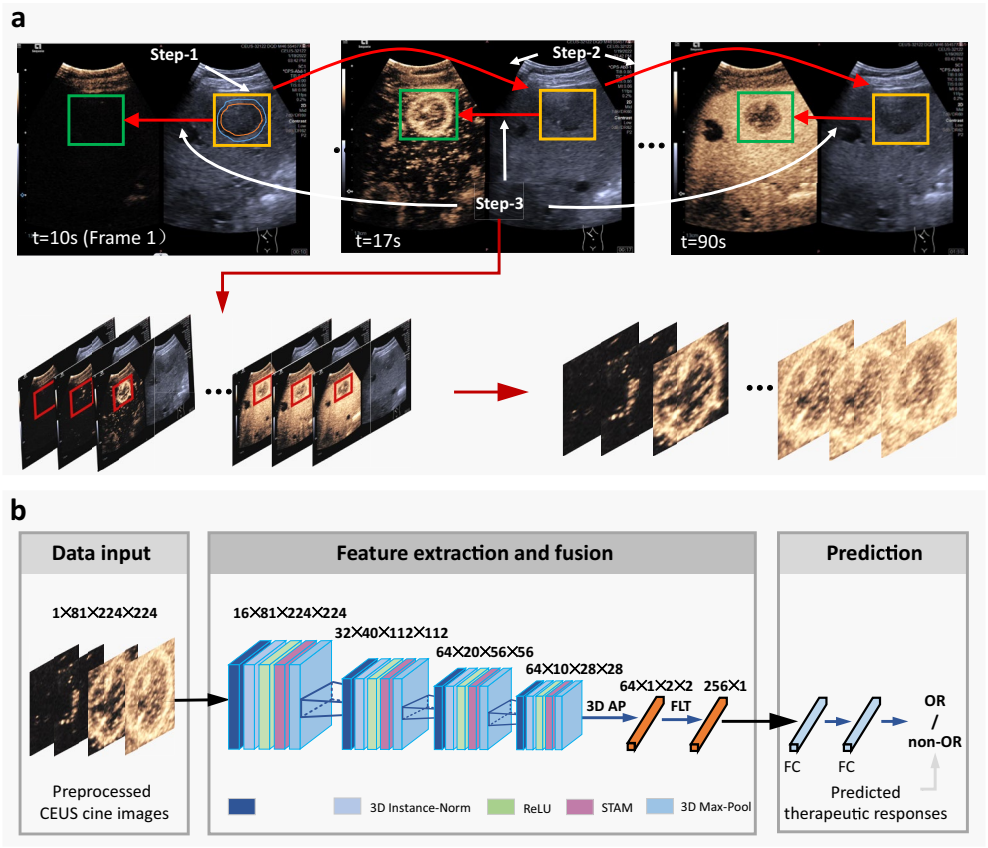

FIGURE 2 | (a) Procedure for the cropping of CEUS image patches. Step-1: for the first frame of the dual-view CEUS data, tumor boundaries were delineated by the radiologist on the B-mode image side and a bounding box denoting the ROI (tumor and peritumoral regions) was generated by the software according to the delineation; Step-2: the ROI of the first frame on the B-mode image side was tracked by using an object-tracking method, by which bounding boxes for the subsequent B-mode images were automatically generated by the software. Step-3: the bounding boxes on the B-mode image side were mapped onto the CEUS image side, and then CEUS images were cropped by the mapped bounding boxes. (b) Architecture of AE-3DNet. The data input to the model were a sequence of 81 frames sampled from the cropped CEUS image patches. AE-3DNet = attention-enhanced three-dimensional deep convolutional neural network, AP = average pooling, Conv = convolution, FC = fully-connected, FLT = flattening, InstanceNorm = instance normalization, Max-Pool = Max pooling, non-OR = non-objective response, OR = objective response, ReLU = rectified linear unit, STAM = spatiotemporal attention module.

using a semi-automatic segmentation tool developed in-house under the MATLAB platform. The detailed procedure for CEUS image cropping is illustrated in Figure 2a and section 1 of Data S1. Due to the limited data input capacity of the deep learning model, the CEUS image patches extracted from each video clip were evenly sampled (one frame per second for 10–90 s), resulting in a sequence of 81 frames for subsequent model development.

2.6 | Model Developments

An attention-enhanced three-dimensional deep convolutional neural network, AE-3DNet, was devised for predicting therapeutic responses from the preprocessed CEUS data. AE-3DNet was constructed on top of a three-dimensional convolutional neural network (3DNet) [26], with the innovative incorporation of spatiotemporal attention modules (STAMs) to enhance the capacity for dynamic feature extraction. As shown in Figures 2b and S1, the data input to AE-3DNet was a sequence of CEUS image patches obtained via data preprocessing. Four convolutional blocks (CBs) were sequentially stacked to extract deep features from the input data, in which each CB was composed of a three-dimensional convolution layer (3D Conv) followed by the other layers, including a 3D instance normalization (3D Instance-Norm) layer, a rectified linear unit (ReLU), and a 3D max pooling (3D Max-Pool) layer, with an STAM lying between the ReLU and 3D Max-Pool layers. At the back end of the model, the deep features were processed by two fully connected (FC) layers to produce the prediction results, that is, whether the tumor would achieve an OR or non-OR after therapeutic treatment. Additionally, we developed other models, including the conventional three-dimensional convolutional neural network (3DNet) without the attention module and a traditional model based on the time-intensity curve (TIC), to compare against AE-3Dnet. Detailed information can be found in Data S1, Tables S1 and S2.

2.7 | Model Evaluation and Interpretation

Patients from Hospital Center A served as the internal validation cohort, on which fivefold cross-validation (FFCV) was performed to develop and internally evaluate the models, while patients from Hospital Center B served as the external validation cohort in order to assess the generalizability of the models derived from the internal validation cohort. The predictive performances of various models were evaluated in terms of accuracy (ACC), sensitivity (SEN), specificity (SPE), receiver-operating-characteristic (ROC) curve, and the area under the ROC curve (AUC). To interpret the inferring behavior of the AE-3DNet model, we employed gradient-weighted class activation mapping (Grad-CAM) [27] to convert the model's outputs into pseudocoloured images (attention or heatmaps) for qualitative analysis. Besides, the association between AE-3DNet scores and patient survivals was evaluated in the external validation cohort using the Kaplan–Meier survival curve, with comparisons against the other models. For this survival analysis, three different models were applied for prediction, the probability of efficacy evaluation as OR was generated, the cutoff value was obtained by maximizing the Youden Index [28] of the ROC curve to classify patients as OR or non-OR, and the survival curve was obtained according to the predicted grouping and compared with the real survival curve.

2.8 | Statistical Analysis

Continuous variables are expressed as the median (interquartile ranges, IQRs) and were compared by Student's *t*-test. Categorical variables are expressed as numbers (percentages) and were compared by the chi-squared test. The Clopper–Pearson method [29] was used to calculate 95% confidence intervals (CIs) for ACC, SEN, SPE, and AUC. Differences between the ROC curves and the corresponding AUC values were measured using the DeLong test [30]. The log-rank test was used to evaluate the significance of the differences between Kaplan–Meier survival curves. All statistical tests were two-sided, and a *p*-value less than 0.05 indicates statistical significance. Data analysis was performed using MedCalc (version 18.2) and MATLAB (version R2023a) software.

3 | Results

3.1 | Patient Characteristics

From March 1, 2019, to March 30, 2023, a total of 439 patients were initially enrolled in our study, of whom 326 patients met the screening criteria and were included for analysis. Among these included patients (median age, 51 years; IQR, 43–60 years; 308 male), 243 individuals (median age, 52 years; IQR, 45–60 years; OR, 159; non-OR, 84) were from Hospital Center A and formed the internal validation cohort, and 83 subjects (median age, 50 years; IQR, 43–59 years; OR, 47; non-OR, 36) were from Hospital Center B and constituted the external validation cohort. The baseline characteristics of the included patients are summarized in Table 1.

3.2 | Performance in the Internal Validation Cohort

By running FFCV ten times on the internal validation cohort, the mean ACC, SEN, SPE, and AUC of AE-3DNet were 0.77 (95% CI: 0.75, 0.79; standard deviation [SD], 0.06), 0.78 (95% CI: 0.75, 0.81; SD, 0.103), 0.76 (95% CI: 0.74, 0.78; SD, 0.07), and 0.84 (95% CI: 0.83, 0.86; SD, 0.06), respectively, which were higher than the corresponding values obtained by the conventional 3DNet (ACC, 0.74 [95% CI: 0.72–0.76; SD, 0.07]; SEN, 0.75 [95% CI: 0.72–0.78; SD, 0.09]; SPE, 0.74 [95% CI: 0.71–0.76; SD, 0.10]; AUC, 0.81 [95% CI: 0.79–0.83; SD, 0.06]), and by the TIC model (ACC, 0.67 [95% CI: 0.65–0.69; SD, 0.08]; SEN, 0.69 [95% CI: 0.66–0.72; SD, 0.12]; SPE, 0.66 [95% CI: 0.64–0.69; SD, 0.12]; AUC, 0.75 [95% CI: 0.73–0.77; SD, 0.09]). These results are summarized in Table 2. Averaged ROC curves of the three models and box-and-whisker plots of the AUC values corresponding to multiple runs of the FFCV are shown in Figures 3a and 4, respectively. AE-3DNet exhibited the highest performance metrics and ROC curves, followed by 3DNet and the TIC model. DeLong test confirmed significant differences between the ROC curves of AE-3DNet and those of the other models ($p_s < 0.001$), while

TABLE 1 | Patient baseline characteristics.

Characteristics	Internal validation cohort ($n = 243$)	External validation cohort ($n = 83$)	p
Age (years)[a]	52 (45–60)	50 (43–59)	0.97
Sex			
Male	227 (93)	81 (98)	0.15
Female	16 (7)	2 (2)	
Tumor size on US (cm)[a]	10.5 (7.3–13.5)	8.7 (5.4–11.6)	0.04
AFP, ng/mL			
< 20	47 (19)	22 (27)	0.25
20–200	43 (18)	17 (20)	
≥ 200	153 (63)	44 (53)	
Differentiation type			
Well	28 (11)	15 (18)	0.31
Medium	101 (42)	31 (37)	
Poor	114 (47)	37 (45)	
HBV infection			
Positive	227 (93)	65 (78)	< 0.001
Negative	16 (7)	18 (22)	
Liver cirrhosis			
Positive	178 (73)	40 (48)	< 0.001
Negative	65 (27)	43 (52)	
BCLC Stage			
Stage B	90 (37)	31 (37)	0.96
Stage C	153 (63)	52 (63)	
Child-Pugh Score			
Child-Pugh A	177 (73)	70 (84)	0.03
Child-Pugh B	66 (27)	13 (16)	
mRECIST			
PR	138 (57)	46 (56)	0.05
CR	21 (9)	1 (1)	
PD	18 (7)	11 (13)	
SD	66 (27)	25 (30)	
Tumor response			
OR	159 (65)	47 (57)	0.15
Non-OR	84 (35)	36 (43)	

Note: Except where indicated, data are numbers of participants, with percentages in parentheses. Student's *t* test was used to compare continuous variables. The chi-squared test was used to compare categorical variables.
Abbreviations: AFP = alpha-fetoprotein, BCLC = Barcelona Clinic Liver Cancer, CR = complete response, HBV = hepatitis B virus, mRECIST = Modified Response Evaluation Criteria in Solid Tumors, non-OR = non-objective response, OR = objective response, PD = progressive disease, PR = partial response, SD = stable disease.
[a]Data are medians; data in parentheses are interquartile ranges (IQRs).

TABLE 2 | Performances of various models on internal and external validation cohorts.

Cohorts	Models	ACC	SEN	SPE	AUC
Internal validation	TIC	0.67 [0.65–0.69] (±0.08)	0.69 [0.66–0.72] (±0.12)	0.66 [0.64–0.69] (±0.12)	0.75 [0.73–0.77] (±0.09)
	3DNet	0.74 [0.72–0.76] (±0.07)	0.75 [0.72–0.78] (±0.09)	0.74 [0.71–0.76] (±0.10)	0.81 [0.79–0.83] (±0.06)
	AE-3DNet	0.77 [0.75–0.79] (±0.06)	0.78 [0.75–0.81] (±0.10)	0.76 [0.74–0.78] (±0.07)	0.84 [0.83–0.86] (±0.06)
External validation	TIC	0.69 [0.58–0.78]	0.69 [0.52–0.84]	0.68 [0.53–0.81]	0.70 [0.59–0.79]
	3DNet	0.75 [0.64–0.84]	0.72 [0.55–0.86]	0.77 [0.62–0.88]	0.79 [0.68–0.87]
	AE-3DNet	0.81 [0.71–0.89]	0.81 [0.64–0.92]	0.81 [0.67–0.91]	0.85 [0.76–0.92]

Note: Data in brackets and parentheses denote respectively the 95% confidence intervals and standard deviations.
Abbreviations: 3DNet = three-dimensional deep convolutional neural network, ACC = accuracy, AE-3DNet = attention enhanced three-dimensional deep convolutional neural network, AUC = area under the receiver-operating-characteristic curve, SEN = sensitivity, SPE = specificity, TIC = time intensity curve-based model.

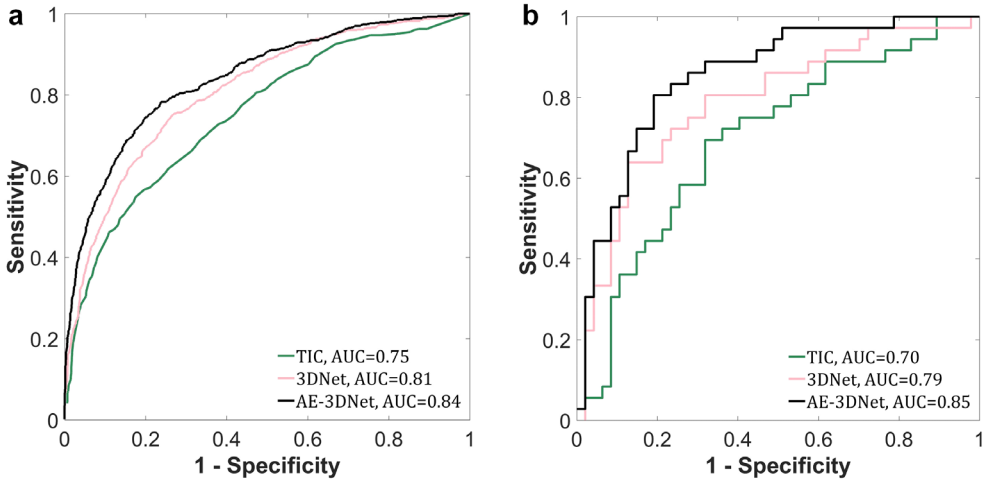

FIGURE 3 | ROC curves of various models including TIC, 3DNet, and AE-3DNet. (a) ROC curves on the internal validation cohort; (b) ROC curves on the external validation cohort. 3DNet = conventional three-dimensional deep convolutional neural network, AE-3DNet = attention-enhanced three-dimensional deep convolutional neural network, TIC = time intensity curve-based model.

paired Student's *t* test indicated that the differences among the AUC values obtained by the various models were significant ($p_s < 0.05$).

3.3 | Performance in the External Validation Cohort

To assess the models' generalizability, we evaluated the performance of various models on the external validation cohort. The ACC, SEN, SPE, and AUC achieved by AE-3DNet were 0.81 (95% CI: 0.71–0.89), 0.81 (95% CI: 0.64–0.92), 0.81 (95% CI: 0.67–0.91), and 0.85 (95% CI: 0.76–0.92), respectively, which were again higher than the values obtained by 3DNet (ACC, 0.75 [95% CI: 0.64–0.84]; SEN, 0.72 [95% CI: 0.55–0.86]; SPE, 0.77 [95% CI: 0.62–0.88]; AUC, 0.79 [95% CI: 0.68–0.87]), and by the TIC model (ACC, 0.69 [95% CI: 0.58–0.78]; SEN, 0.69 [95% CI: 0.52–0.84]; SPE, 0.68 [95% CI: 0.53–0.81]; AUC, 0.70 [95% CI: 0.59–0.79]). The corresponding results are summarized in the bottom rows of Table 2. ROC curves yielded by various models in the external validation cohort are shown in Figure 3b, where it can be observed that AE-3DNet yielded the best quantitative results and ROC curve, followed by 3DNet and the TIC model, the results of which are consistent with those obtained in the internal validation. We also performed the DeLong test, which confirmed statistically significant differences between the ROC curves of AE-3DNet and those of the other models ($p = 0.041$ and $p = 0.038$ for 3DNet and TIC, respectively).

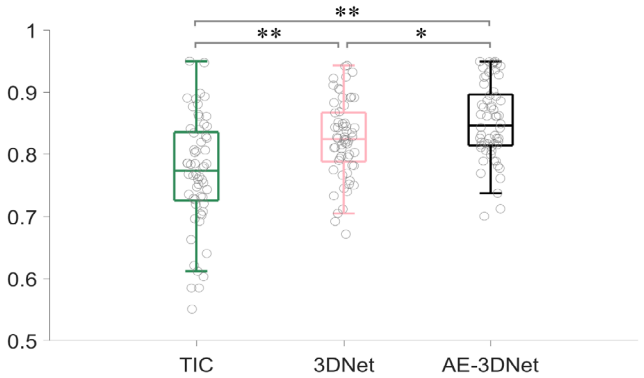

FIGURE 4 | Box-and-whisker plots of the AUC values obtained by ten independent runs of the FFCV. The lines through the boxes represent the median values. The upper and lower sides of the boxes denotes the first (Q1) and third (Q3) quartiles of the data. The length of the box denotes the interquartile ranges (IQRs) within which the middle 50% of the data are located. The upper and lower whiskers represent data outside the middle 50% of the data. The uppermost horizontal line attached to the upper whisker denotes the maximum value (Q1 + 1.5 × IQR), while the lowest horizontal line attached to the lower whisker denotes the minimum value (Q1 − 1.5 × IQR). Circle points outside the box and whisker denote outliers. $^{**}p < 0.01$; $^{*}p < 0.05$. 3DNet = three-dimensional deep convolutional neural network, AE-3DNet = attention-enhanced three-dimensional deep convolutional neural network, FFCV = fivefold cross-validation, TIC = time intensity curve-based model.

3.4 | Model Visualization

Although AE-3DNet demonstrated excellent performance, clinicians may be more concerned with understanding how the model infers its prediction outcomes. To interpret the model's behavior, we used Grad-CAM to convert the feature maps (generated by the AE-3DNet model as it inferred the tumor response from the CEUS data) into pseudocolored maps, that is, attention maps or heatmaps. By reading the attention maps, we summarized several rules that might be used by AE-3DNet for decision-making, most of which were related to the feeding artery and feeding speed of the tumor, as reflected by the CEUS data. First, the presence of apparent feeding arteries in the early stage of the arterial phase was indicative of a higher likelihood of achieving an OR, as illustrated in Figure 5a. Second, a relatively high feeding speed, defined as the time from contrast agent arrival at the tumor to full enhancement, was associated with a higher probability of achieving an OR. This relationship is exemplified in Figure 5b, where the model assigns greater attention to early-stage CEUS frames for patients who achieved an OR, reflecting the higher feeding speed, compared to late-stage frames (vice versa for patients who achieved non-OR).

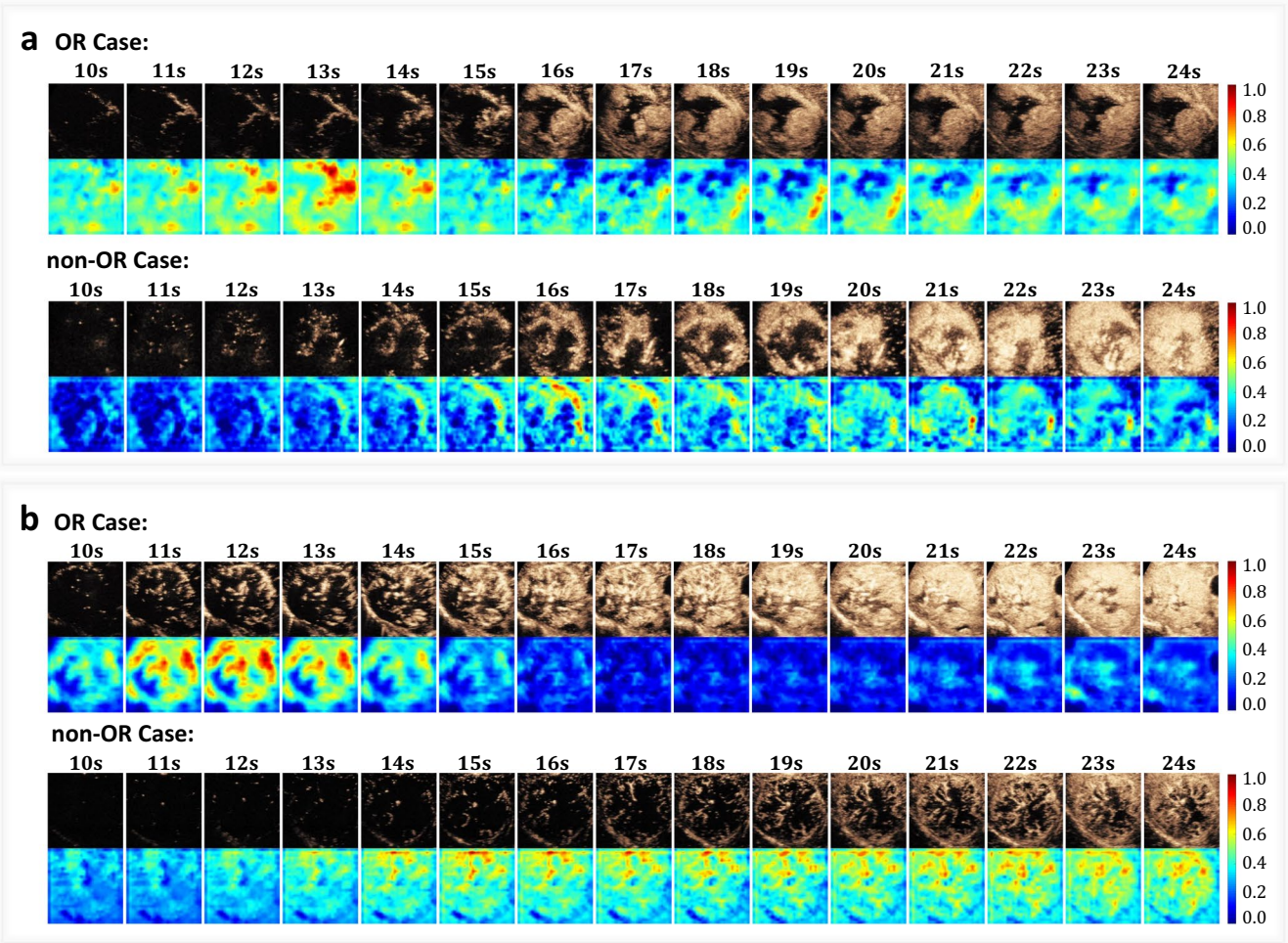

FIGURE 5 | (a, b) Feature map visualization of AE-3DNet for cases whose responses were correctly predicted by the model. Red and warm colors indicate stronger contributions to the predictive outcome, whereas blue and cold colors indicate weaker contributions. AE-3DNet = attention-enhanced three-dimensional deep convolutional neural network, non-OR = non-objective response, OR = objective response.

3.5 | Association Between Therapeutic Responses and Patient Survivals

This study was censored on March 30, 2024. In the internal validation cohort, patients achieving an OR had a mean overall survival (OS) of 19.7 months (95% CI: 17.92–21.52), whereas those achieving a non-OR had a mean OS of 11.6 months (95% CI: 9.94–11.34) ($p < 0.001$). In the external validation cohort, the mean OS was 14.8 months (95% CI: 13.39–16.19) for patients who achieved an OR and 10.3 months (95% CI: 8.40–12.20) for patients who achieved a non-OR ($p < 0.05$). The efficacy of HAIC combination therapy was significantly correlated with the long-term prognosis of advanced HCC (Figures 6a and S2). The prognosis was further stratified based on predicted therapeutic responses from three deep learning models. The predicted therapeutic responses of AE-3DNet could significantly distinguish patients with poor outcomes in the external validation cohort ($p < 0.05$, Figure 6b), whereas those predicted by 3DNet and TIC failed to distinguish the patients with poor outcomes ($p_s > 0.05$, Figure 6c and Figure 6d). Survival curves generated by AE-3DNet (Figure 6b) closely mirrored actual clinical prognosis (Figure 6a).

4 | Discussion

This dual-center retrospective study successfully built an attention-enhanced deep learning model, AE-3DNet, to predict the response of advanced HCC to HAIC combination therapy using pretreatment CEUS cine images. The model's performance was assessed on an internal validation cohort (243 patients, 65% OR) and an external validation cohort (83 patients, 57% OR), with comparisons against 3DNet and TIC. The deep learning model of AE-3DNet demonstrated superior performance in predicting the efficacy of HAIC-associated triple combination therapy for advanced HCC, achieving an AUC value of 0.84 (95% CI: 0.83–0.86) in the internal validation cohort and an AUC value of 0.85 (95% CI: 0.76–0.92) in the external validation cohort. Given the promising performance, AE-3DNet may serve as a valuable tool to identify patients who could potentially benefit from HAIC-associated triple combination therapy, thereby contributing to improved patient management.

HAIC combination therapy for advanced HCC has been widely studied, particularly in Chinese institutions, due to China's high HCC burden and the urgent need for innovative

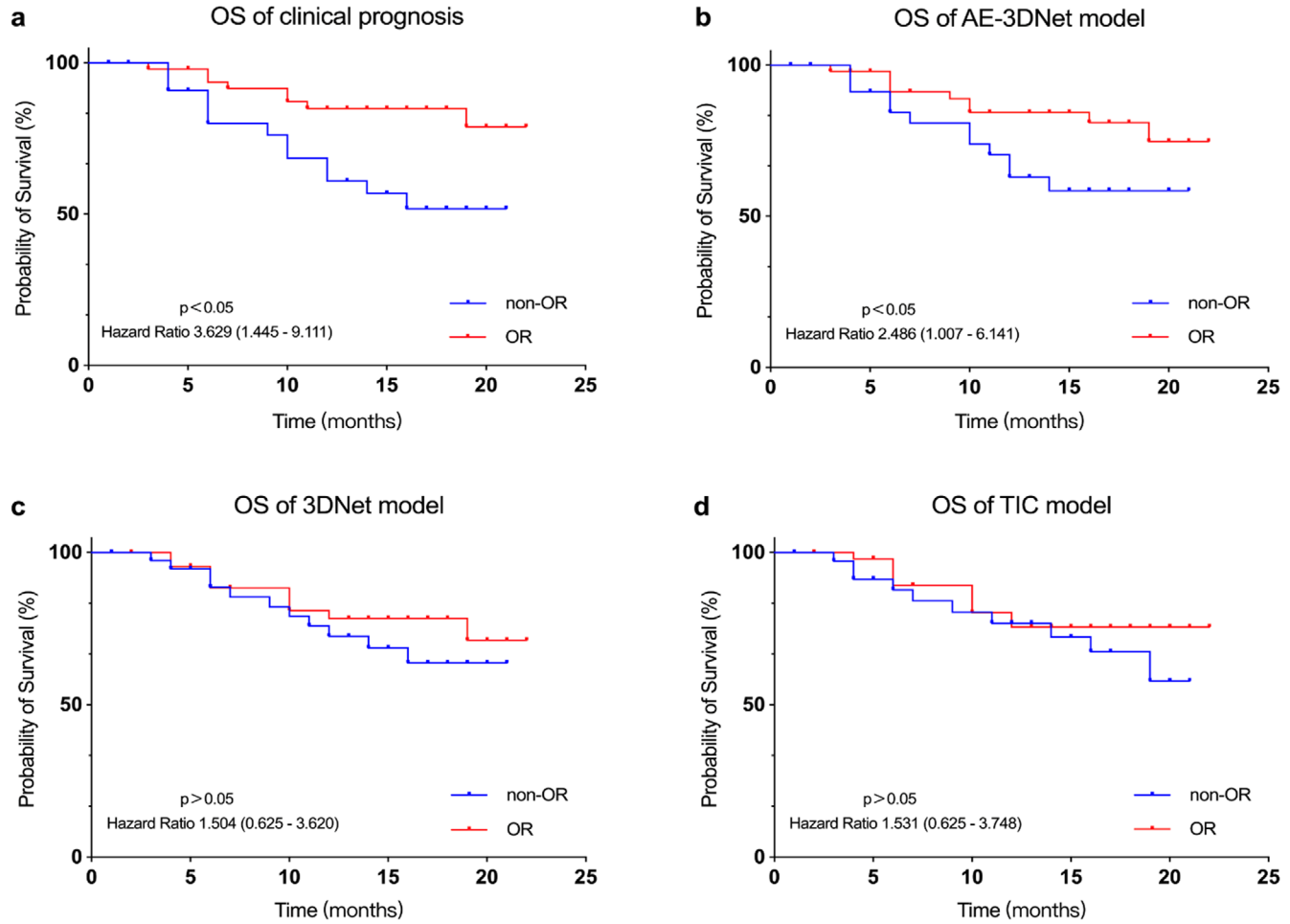

FIGURE 6 | Survival curves of clinical prognosis and predicted response of the three models in the external validation cohort ($n = 83$). (a) OS curves of clinical prognosis; (b) predicted OS curves of AE-3DNet model; (c) predicted OS curves of 3DNet model; (d) predicted OS curves of TIC model. Comparisons between curves were performed with the log-rank test. The hazard ratios were computed using the Mantel–Haenszel method. 3DNet = conventional three-dimensional deep convolutional neural network, AE-3DNet = attention-enhanced three-dimensional deep convolutional neural network, OS = overall survival, TIC = time intensity curve-based model.

treatments. Our center uses real-world data to develop predictive models, helping identify responders and reduce futile interventions. The combination of HAIC, molecular targeted agents (MTAs), and immune checkpoint inhibitors (ICIs) targets tumor vascular structure, angiogenesis, and the immune microenvironment, respectively. While their interactions are not fully understood, their individual efficacy and complementary mechanisms—such as HAIC releasing tumor antigens to enhance ICI response, or MTAs normalizing vasculature to improve drug delivery—support clinical use. However, complexities like excessive vascular pruning or HAIC-induced inflammation may either synergize or counteract effects. Despite these uncertainties, the urgent need for effective therapies and promising preliminary results have driven clinical adoption. Further trials are needed to clarify interactions and optimize sequencing, dosing, and patient selection for maximal benefit, balancing innovation with evidence generation to improve outcomes.

The successful development of AE-3DNet represents a significant advancement in utilizing machine learning to tackle clinical challenges, aligning well with other relevant studies[32]. Here, we demonstrated the superiority of AE-3DNet over the traditional TIC method and the advanced 3DNet in the context of response prediction of advanced HCC. While the TIC method mainly focuses on temporal feature extraction from the time-intensity curve and overlooks spatial patterns in CEUS cine images, the 3DNet and AE-3DNet models effectively capture both spatial and temporal features from the CEUS cine data. This ability likely explains why the 3DNet and its attention-enhanced version exhibited superior prediction performance compared to the TIC method in both internal and external cohorts. Furthermore, our results on the external validation cohort revealed that AE-3DNet displayed enhanced generalizability in contrast to 3DNet. This finding suggests that the attention-enhanced model's extraction of CEUS spatiotemporal features was more robust than that of the model without attention enhancement, underscoring the significance of attention modules in optimizing the effectiveness of deep learning for CEUS data processing.

While AE-3DNet exhibits impressive performance, clinicians are particularly interested in comprehending the rationale behind its prediction outcomes. To enhance the interpretability

of the AE-3DNet model, we conducted visualizations and comparisons of its feature attention maps for populations categorized as OR and non-OR. Upon this analysis, we identified distinct CEUS patterns that may signify different responsive populations, such as the presence of conspicuous feeding arteries early in the arterial phase and elevated feeding speed among the OR patients. These delineated patterns could be invaluable in clinical settings where AI models are not readily available. Understanding and summarizing these patterns could enrich the clinicians' capacity to predict the response of advanced HCC to HAIC combination therapy by analyzing pretreatment CEUS cine images. In the current study, tumors with well-defined feeding arteries in the early arterial phase showed higher OR rates, suggesting that the efficacy of HAIC was closely associated with adequate drug delivery via dominant tumor-feeding vessels. The model's focus on early arterial phase frames demonstrates its ability to detect functional vascular patency, a critical factor for efficient FOLFOX penetration. A shorter tumor enhancement time likely reflects favorable hemodynamic conditions, such as efficient blood flow and reduced vascular resistance, potentially leading to more effective drug delivery during HAIC. Combining feeding artery morphology and hemodynamic profiling has the potential to refine patient selection for therapy.

Although our model was primarily designed to predict the response to HAIC combination therapy, it has the potential for transferability (through transfer learning [28]) to address various predictive tasks. These tasks may include predicting the efficacy of TACE, HAIC, systemic therapies [29], and TACE [30], provided task-specific CEUS data and data labels are available. Of note, with the incorporation of spatiotemporal attention modules (STAMs), the transferred model may exceed existing models, such as the TACE efficacy prediction model [19], in fulfilling specific tasks. While further assessments are necessary to validate our assumptions, even without exploring alternative predictive tasks, AE-3DNet can complement existing models to enhance therapeutic strategy recommendations. For instance, in cases where AE-3DNet predicts unsatisfactory therapeutic efficacy of HAIC-based triple therapy for advanced HCC patients, other models tailored to different therapies could be consulted to determine the most appropriate treatment regimen. By employing multiple predictive models, significant strides can be made in advancing personalized treatments and ultimately improving patient outcomes.

In the current study, we investigated the effectiveness of the deep learning model in predicting OS. Our results revealed a significant association between the efficacy of HAIC combination therapy and the long-term prognosis of patients with advanced HCC. This finding is in accordance with previous relevant studies [30]. Assessment of OS predicted by the three models indicates that the AE-3DNet model exhibits the highest level of concordance with real-world outcomes, suggesting that our model not only demonstrates predictive efficacy but also accurately reflects patients' OS through its efficacy predictions. Inconsistent OS outcomes were observed between the internal and the external validation cohort, which were related to the different follow-up times. The internal cohort was followed up for a longer period of time, so the OS was longer than that of the external cohort.

There are several limitations of our study. Firstly, as HAIC combination therapy has not been extensively used in clinical practice, particularly in our external validation cohort, the sample size was relatively small. Additional research involving multicenter studies with larger sample sizes is warranted to further validate our results. Secondly, the proposed model may be influenced by the specific characteristics of the patient cohorts, highlighting the importance of model adjustment through transfer learning when applying it to different regions.

In conclusion, the deep learning model AE-3Dnet, developed based on pretreatment CEUS cine images, performed satisfactorily in predicting the efficacy of HAIC combination therapy and prognosis in patients with advanced HCC. Therefore, it may be considered a promising tool to identify individuals who could potentially benefit from the triple combination therapy.

Author Contributions

Xu Han: conceptualization, data curation, writing – original draft, writing – review and editing. **Chuan Peng:** formal analysis, investigation, resources. **Si-Min Ruan:** investigation, validation. **Lingling Li:** conceptualization, formal analysis, supervision, writing – review and editing. **Minke He:** resources. **Ming Shi:** resources. **Bin Huang:** methodology, software. **Yudi Luo:** methodology, software. **Jingming Liu:** methodology, software. **Huiying Wen:** methodology, software. **Wei Wang:** resources. **Jianhua Zhou:** resources. **Minhua Lu:** methodology, software. **Xin Chen:** conceptualization, funding acquisition, resources. **Ruhai Zou:** conceptualization, data curation, formal analysis, funding acquisition, project administration, supervision, writing – review and editing. **Zhong Liu:** funding acquisition, methodology, project administration, visualization, writing – original draft, writing – review and editing.

Acknowledgments

The authors have nothing to report.

Ethics Statement

Approval of the research protocol by an Institutional Review Board. This retrospective analysis utilized data from a prospective clinical registry (Ethical Approval no. B2016-042-01). Written informed consent had been obtained from all participants prior to CEUS examinations, which included explicit permission for the utilization of their data in research upon anonymization.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. H. Sung, J. Ferlay, R. L. Siegel, et al., "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians* 71, no. 3 (2021): 209–249, https://doi.org/10.3322/caac.21660.

2. M. Lewin, L. Fartoux, A. Vignaud, L. Arrive, Y. Menu, and O. Rosmorduc, "The Diffusion-Weighted Imaging Perfusion Fraction F Is a Potential Marker of Sorafenib Treatment in Advanced Hepatocellular Carcinoma: A Pilot Study," *European Radiology* 21, no. 2 (2011): 281–290.

3. R. Bartkowski, M. R. Berger, J. L. Aguiar, et al., "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," *Journal of Cancer Research and Clinical Oncology* 111, no. 1 (1986): 42–46.

4. K. Ueshima, S. Ogasawara, M. Ikeda, et al., "Hepatic Arterial Infusion Chemotherapy Versus Sorafenib in Patients With Advanced Hepatocellular Carcinoma," *Liver Cancer* 9, no. 5 (2020): 583–595.

5. Z. Lai, M. He, X. Bu, et al., "Lenvatinib, Toripalimab Plus Hepatic Arterial Infusion Chemotherapy in Patients With High-Risk Advanced Hepatocellular Carcinoma: A Biomolecular Exploratory, Phase II Trial," *European Journal of Cancer* 174 (2022): 68–77.

6. T. Q. Zhang, Z. J. Geng, M. X. Zuo, et al., "Camrelizumab (a PD-1 Inhibitor) Plus Apatinib (An VEGFR-2 Inhibitor) and Hepatic Artery Infusion Chemotherapy for Hepatocellular Carcinoma in Barcelona Clinic Liver Cancer Stage C (TRIPLET): A Phase II Study," *Signal Transduction and Targeted Therapy* 8, no. 1 (2023): 413.

7. S. Chen, B. Xu, Z. Wu, et al., "Pembrolizumab Plus Lenvatinib With or Without Hepatic Arterial Infusion Chemotherapy in Selected Populations of Patients With Treatment-Naive Unresectable Hepatocellular Carcinoma Exhibiting PD-L1 Staining: A Multicenter Retrospective Study," *BMC Cancer* 21, no. 1 (2021): 1126.

8. M. K. He, R. B. Liang, Y. Zhao, et al., "Lenvatinib, Toripalimab, Plus Hepatic Arterial Infusion Chemotherapy Versus Lenvatinib Alone for Advanced Hepatocellular Carcinoma," *Therapeutic Advances in Medical Oncology* 13 (2021): 17588359211002720.

9. Y. Xin, F. Cao, H. Yang, et al., "Efficacy and Safety of Atezolizumab Plus Bevacizumab Combined With Hepatic Arterial Infusion Chemotherapy for Advanced Hepatocellular Carcinoma," *Frontiers in Immunology* 13 (2022): 929141.

10. W. Zhang, K. Zhang, C. Liu, et al., "Hepatic Arterial Infusion Chemotherapy Combined With Anti-PD-1/PD-L1 Immunotherapy and Molecularly Targeted Agents for Advanced Hepatocellular Carcinoma: A Real World Study," *Frontiers in Immunology* 14 (2023): 1127349.

11. M. Ikeda, S. Shimizu, T. Sato, et al., "Sorafenib Plus Hepatic Arterial Infusion Chemotherapy With Cisplatin Versus Sorafenib for Advanced Hepatocellular Carcinoma: Randomized Phase II Trial," *Annals of Oncology* 27, no. 11 (2016): 2090–2096.

12. M. K. He, R. H. Zou, Q. J. Li, et al., "Phase II Study of Sorafenib Combined With Concurrent Hepatic Arterial Infusion of Oxaliplatin, 5-Fluorouracil and Leucovorin for Unresectable Hepatocellular Carcinoma With Major Portal Vein Thrombosis," *Cardiovascular and Interventional Radiology* 41, no. 5 (2018): 734–743.

13. S. Q. Ma, B. R. Cao, H. Zhang, et al., "The Lack of Raf-1 Kinase Feedback Regulation Enhances Antiapoptosis in Cancer Cells," *Oncogene* 36, no. 14 (2017): 2014–2022.

14. E. W. Pepin, S. M. Nordeck, and D. T. Fetzer, "Nontraditional Uses of US Contrast Agents in Abdominal Imaging and Intervention," *Radiographics* 42, no. 6 (2022): 1724–1741.

15. S. M. Badiu, E. C. Gheorghe, C. Nicolau, and A. Săftoiu, "Quantitative Time Intensity Curve Analysis of Contrast-Enhanced Ultrasound (CEUS) Examinations for the Assessment of Focal Liver Lesions," *Medical Ultrasonography* 26, no. 1 (2023): 63–71.

16. I. Oezdemir, C. E. Wessner, C. Shaw, J. R. Eisenbrey, and K. Hoyt, "Tumor Vascular Networks Depicted in Contrast-Enhanced Ultrasound Images as a Predictor for Transarterial Chemoembolization Treatment Response," *Ultrasound in Medicine & Biology* 46, no. 9 (2020): 2276–2286.

17. Z. Xu, C. An, F. Shi, et al., "Automatic Prediction of Hepatic Arterial Infusion Chemotherapy Response in Advanced Hepatocellular Carcinoma With Deep Learning Radiomic Nomogram," *European Radiology* 33, no. 12 (2023): 9038–9051.

18. J. Bo, H. Peng, Z. LianHua, F. Xiang, and L. YuKun, "Intraarterial Contrast-Enhanced Ultrasound to Predict the Short-Term Tumour Response of Hepatocellular Carcinoma to Transarterial Chemoembolization With Lipiodol," *BMC Cancer* 21, no. 1 (2021): 1171.

19. D. Liu, F. Liu, X. Xie, et al., "Accurate Prediction of Responses to Transarterial Chemoembolization for Patients With Hepatocellular Carcinoma by Using Artificial Intelligence in Contrast-Enhanced Ultrasound," *European Radiology* 30, no. 4 (2020): 2365–2376.

20. Y. Zhao, F. Huang, S. Liu, et al., "Prediction of Therapeutic Response of Unresectable Hepatocellular Carcinoma to Hepatic Arterial Infusion Chemotherapy Based on Pretherapeutic MRI Radiomics and Albumin–Bilirubin Score," *Journal of Cancer Research and Clinical Oncology* 149, no. 8 (2023): 5181–5192.

21. Y. Sun, H. Bai, W. Xia, et al., "Predicting the Outcome of Transcatheter Arterial Embolization Therapy for Unresectable Hepatocellular Carcinoma Based on Radiomics of Preoperative Multiparameter MRI," *Journal of Magnetic Resonance Imaging* 52, no. 4 (2020): 1083–1090.

22. C. Sun, Q. Wang, L. Hou, R. Zhang, Y. Chen, and L. Niu, "A Contrast-Enhanced Ultrasound-Based Nomogram for the Prediction of Therapeutic Efficiency of Anti-PD-1 Plus Anti-VEGF Agents in Advanced Hepatocellular Carcinoma Patients," *Frontiers in Immunology* 14 (2023): 1229560.

23. P. M. Bossuyt, J. B. Reitsma, D. E. Bruns, et al., "STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies," *BMJ* 351 (2015): h5527.

24. M. Zhao, Z. Guo, Y. H. Zou, et al., "Arterial Chemotherapy for Hepatocellular Carcinoma in China: Consensus Recommendations," *Hepatology International* 18, no. 1 (2023): 4–31.

25. J. M. Llovet and R. Lencioni, "mRECIST for HCC: Performance and Novel Refinements," *Journal of Hepatology* 72, no. 2 (2020): 288–306.

26. S. Ji, M. Yang, and K. Yu, "3D Convolutional Neural Networks for Human Action Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35, no. 1 (2013): 221–231.

27. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization," *International Journal of Computer Vision* 128, no. 2 (2020): 336–360.

28. H.-C. Shin, H. R. Roth, M. Gao, et al., "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," *IEEE Transactions on Medical Imaging* 35, no. 5 (2016): 1285–1298.

29. M. Reig, A. Forner, J. Rimola, et al., "BCLC Strategy for Prognosis Prediction and Treatment Recommendation: The 2022 Update," *Journal of Hepatology* 76, no. 3 (2022): 681–693.

30. Y. Long, J. Huang, J. Liao, et al., "Safety and Survival Outcomes of Liver Resection Following Triple Combination Conversion Therapy for Initially Unresectable Hepatocellular Carcinoma," *Cancers* 15, no. 24 (2023): 5878.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.