

RESEARCH

Open Access



# Malignancy diagnosis of liver lesion in contrast enhanced ultrasound using an end-to-end method based on deep learning

Hongyu Zhou<sup>1,2,3,4†</sup>, Jianmin Ding<sup>1,2,3,4†</sup>, Yan Zhou<sup>1,2,3,4</sup>, Yandong Wang<sup>1,2,3,4</sup>, Lei Zhao<sup>1,2,3,4</sup>, Cho-Chiang Shih<sup>5</sup>, Jingping Xu<sup>5</sup>, Jianan Wang<sup>5</sup>, Ling Tong<sup>5</sup>, Zhouye Chen<sup>5</sup>, Qizhong Lin<sup>5</sup> and Xiang Jing<sup>1,2,3,4\*</sup>

## Abstract

**Background** Contrast-enhanced ultrasound (CEUS) is considered as an efficient tool for focal liver lesion characterization, given it allows real-time scanning and provides dynamic tissue perfusion information. An accurate diagnosis of liver lesions with CEUS requires a precise interpretation of CEUS images. However, it is a highly experience dependent task which requires amount of training and practice. To help improve the constrains, this study aims to develop an end-to-end method based on deep learning to make malignancy diagnosis of liver lesions using CEUS.

**Methods** A total of 420 focal liver lesions with 136 benign cases and 284 malignant cases were included. A deep learning model based on a two-dimensional convolution neural network, a long short-term memory (LSTM), and a linear classifier (with sigmoid) was developed to analyze the CEUS loops from different contrast imaging phases. For comparison, a 3D-CNN based method and a machine-learning (ML)-based time-intensity curve (TIC) method were also implemented for performance evaluation.

**Results** Results of the 4-fold validation demonstrate that the mean AUC is 0.91, 0.88, and 0.78 for the proposed method, the 3D-CNN based method, and the ML-based TIC method, respectively.

**Conclusions** The proposed CNN-LSTM method is promising in making malignancy diagnosis of liver lesions in CEUS without any additional manual features selection.

**Keywords** Contrast enhanced ultrasound (CEUS), Liver lesion, Malignancy diagnosis, Deep Learning

## Introduction

Hepatocellular carcinoma (HCC) is the most common primary liver malignancy and the 3rd most common cause of cancer-related death worldwide, which occupied 70–90% of various kinds of primary liver cancers [1]. Unfortunately, most patients with HCC are diagnosed at the advanced stage and unsuitable for surgery or local treatment, which leads to poor prognosis with median overall survival (OS) of about 6 months [2]. Therefore, an accurate diagnosis for distinguishing HCC from focal liver lesions (FLLs) at the early stage is essential.

Contrast-enhanced ultrasound (CEUS) is considered as an effective tool in the characterization of focal liver

<sup>†</sup>Hongyu Zhou and Jianmin Ding contributed equally to this work and are co-first authors.

\*Correspondence:

Xiang Jing  
dr.jingxiang@vip.163.com

<sup>1</sup> The Third Central Hospital of Tianjin, 83 Jintang Road, Hedong District, Tianjin 300170, China

<sup>2</sup> Tianjin Key Laboratory of Extracorporeal Life Support for Critical Diseases, Tianjin, China

<sup>3</sup> Artificial Cell Engineering Technology Research Center, Tianjin, China

<sup>4</sup> Tianjin Institute of Hepatobiliary Disease, Tianjin, China

<sup>5</sup> Philips Ultrasound R&D Research, Shanghai, China



lesions [3]. Compared to Contrast-enhanced computed tomography (CE-CT) and Contrast-enhanced Magnetic Resonance Imaging (CE-MRI), CEUS operates in real-time with higher temporal resolution and can provide dynamic information on tissue perfusion during all phases of enhancement. During the CEUS examination, clinicians usually need to carry out a series of tasks with tedious details including size measurements, enhancement characteristics analysis while monitoring the tissue perfusion [4]. An accurate and reliable diagnosis of liver lesions with CEUS is based on a well-defined and precise interpretation of the CEUS loops. However, this is a highly experience dependent task and it requires significant amount of training and practice for a clinician to become proficient. Indeed, clinicians skilled in CEUS diagnosis are limited, in particular in lower level hospitals. To some extent, this issue confines the widespread of CEUS application in Liver. Nonetheless, recent advances in computer aided diagnosis may has the potential to reduce the operator-dependency and to popularize the application of CEUS.

Generally, CAD methods to facilitate CEUS mode based diagnosis can be divided into two categories, the classic machine learning methods relying on feature engineering and the deep learning methods that are capable of feature learning. In tradition machine learning methods, some works were based on perfusion features extracted from the time intensity curve (TIC) [5–7]. For example, in [5], the authors used sparse non-negative matrix factorization to automatically extract the TIC based quantitative parameters and employed a neural network classifier to classify benign and malignant liver lesions with an accuracy of 86.36%. Similar efforts have been made in [6], where authors extracted both images features from B-mode and CEUS images and perfusion features from TICs then fed into Support Vector Machine (SVM) and Artificial Neural Networks (ANN) classifiers to make the class prediction. The reported accuracy of the SVM method was 81.1% in benign/malignancy classification and was similar that of an expert reader, 81.4%. In [7], 28 features were extracted from TICs in the CEUS cine-loops from a case were used as an input to SVM for benign/malignancy classification with an accuracy of 91.8%. Nevertheless, these TIC based methods depend on manual ROI selections for TIC generation, which can be subjective to motion impacts. Other classic machine learning methods could include extract CEUS image features as input. In [8], the authors proposed to manually select three frames from CEUS cine loops during arterial phase (AP), portal venous phase (PVP), and late phases (LP), respectively. Using these images, they extracted tens features and applied a multiple kernel learning classifier to predict liver lesion malignancy (Guo et al, [8]).

Indeed, either the selection of CEUS frames, ROIs, or hand-crafted features can be highly operator experienced dependent, time consuming, and thus limiting the classification performance.

On the other hand, deep learning (DL) methods outperformed many traditional feature engineering based machine learning methods. The advantages of deep learning (DL) are mainly lies in threefold. First, deep learning can automatically uncover features from the training data, hence significantly alleviate efforts of hand-crafted features. The learned features may compensate and even surpass the discriminative power of the conventional feature extraction methods. Second, with deep learning, feature interaction and hierarchy can be exploited jointly within the intrinsic deep architecture of a neural network. Consequently, the feature selection process will be significantly simplified. Third, three steps of feature extraction, selection, and supervised classification can be realized within the optimization of the same deep architecture. This systematic design fashion makes the model easier to be fine-tuned for achieving better performance.

Various DL algorithms have been proposed to facilitate the assessment of focal liver lesions (FLLs) dignity and/or entity classifications in ultrasound using B-mode and/or CEUS image data [9, 10]. DL algorithms developed based on B-mode data typically only use static images as data sources. As such, only spatial information is taking into account and the developed networks are primarily VGG-type 2D CNN, DenseNet 2D CNN, and ResNet-type 2D CNN [11–15]. Unlike B-mode, CEUS provides dynamic perfusion information on the tissue in real-time. CEUS images and cine-loops during the examination contain both spatial and temporal information. Intuitively, DL methods applies to CEUS data for liver lesion classification should have the potential to simultaneously obtain spatio-temporal information. In [16], authors proposed to use a general deep learning-based method to predict personalized responses of HCC to first transarterial chemoembolization (TACE) sessions. They utilized a 3D-CNN model to extract spatio-temporal information from CEUS data and showed it performed better than a pre-defined TIC feature-based model and a pre-defined radiomics feature-based model. Although the objective of this study [16] was not to predict the malignancy of liver lesions, it still demonstrated the robust performance and the potential of adopting deep learning to CEUS for liver applications. In [17], the authors proposed a 3D-CNN based DL network classify benign and HCC tumors. In their work, a fixed number of CEUS video frames was selected with a uniform time sampling and then used to form a 3-dimensional data. Finally, the spatial and temporal features were extracted simultaneously using

the 3D-CNN. However, fusing CEUS temporal dimension and spatial information resulting losing the time-varying information during contrast perfusion, which is critical for lesion classification. In [18], the authors proposed a model based on transfer learning to obtain image features from CEUS images at different perfusion phases for liver lesion classification. In [19] CEUS frames were manually selected from CEUS videos at different perfusion phases then fed to multiple parallel 2D ResNet networks to extract features and fuse them in a fully connected layer, which was used to classify benign and malignant images. However, this approach didn't consider the correlation between perfusion phases. This is again very important for physicians to make diagnosis. A similar parallel multi-network framework was also proposed in [20], this differences between the two works lies in different of number of CEUS video frames selection and different ResNet network chosen. In [21], the authors proposed a framework containing a ResNet network followed by a Views-Related Learning module (VRL) and a Two-Step-Orthogonal-Projection (TSOP) module. The ResNet network served as backbone to extract features from manually selected CEUS video frames from various perfusion phases and the VRL+TSOP was used to learn and fuse dynamic temporal features. In [22], the authors proposed to use a 2D-CNN + LSTM model to classify HCC from FNH based on manually selected CEUS video frames. The LTSM model provides the feasibility to learn temporal correlation between perfusion phases in addition to spatial information extraction using 2D-CNN. However, the current work, only considered CEUS videos from arterial phase. Information from the portal venous and late phases was missing but was critical during liver lesion benign and malignancy classification. Nonetheless, all algorithms above requires experienced radiologist to manually select CEUS video frames from different perfusion phases, which again would be operator-dependent. In this manuscript, we propose an end-to-end DL method using CNN-LSTM model to distinguish between benign and malignant liver lesions using CEUS video cine-loops from all perfusion phases, i.e. AP, PVP and LP We aim to avoid as many manual interactions needed from radiologist as possible to make classification more objective and automated. To our knowledge, our work is the first attempt of applying CNN-LSTM model in CEUS liver lesion diagnosis without any additional manual features selection.

This paper is organized as follows. In [Materials and methods](#) section, we presented the data and described the methods in detail. Results of the proposed method in comparison with a classical 3D-CNN model and a

ML TIC-based method are provided in [Results](#) section. The conclusion are drawn in [Discussion](#) section.

## Materials and methods

### Dataset

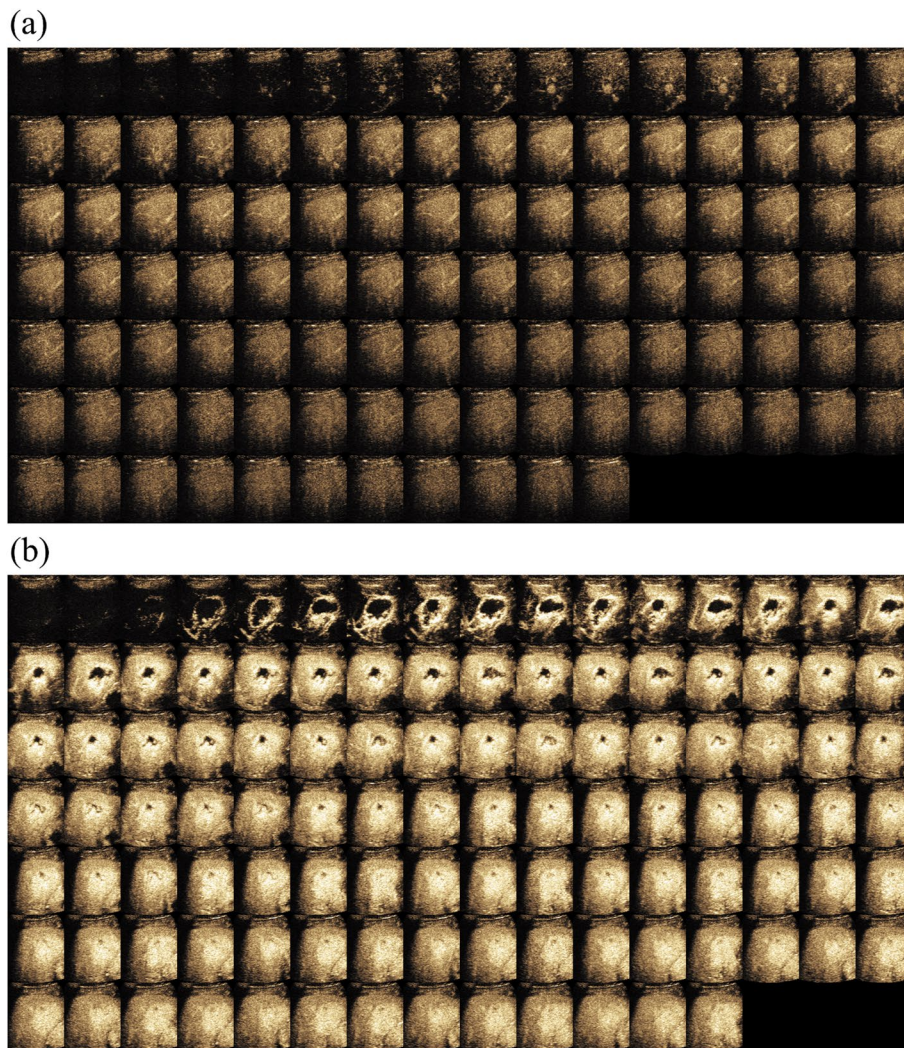
This is a retrospective study and was approved by the Institutional Review Board and Ethics Committee. The requirement for informed consent was waived off. From May 2018 to May 2019, 440 patients with focal liver lesion who underwent CEUS examination were recruited. The CEUS examinations were taken using different ultrasound imaging systems (Philips, Siemens, Mindray and Canon medical systems). During the examination, a bolus of 1.5-2.4 ml of Sonovue (Bracco, Milan, Italy) was injected intravenously through a cubital vein, followed by a flush of NaCl 0.9% 5 ml in bolus. Real-time side by side contrast-enhanced mode was turned on with the mechanical index (MI) was set as low as possible at range of 0.05 to 0.1. Usually, clinicians record several loops from the end of contrast agent injection to about 6 minutes. All CEUS loops were stored in DICOM format. Each frame in loop had both CEUS data and B-mode data in dual display with resolution in 960 (Height) x 1280 (Width) x 3 (RGB). Given this was an retrospective study, the imaging parameter settings on different ultrasound systems were not necessarily the same. Ultrasound doctors adjusted the parameters while scanning to obtain the optimal image quality for them to make the diagnosis. This is more close to common situation during typical clinical practice. In [Fig. 1](#), we showed two typical cases in CEUS.

From the 440 patients above, we further selected with inclusion and exclusion criteria as below. Inclusion criteria were as follows: 1) aged 18-80 years old; 2) diagnosis confirmed by histologic examination, EASL guidelines, contrast MRI scanning, or follow-up beyond 6 months; 3) CEUS loops were completed in the sense that all three phases including AP, PVP and LP were recorded. Exclusion criteria were as follows: 1) lesion size  $\leq 1$ cm; 2) lack of dicom header or fail to load and preprocess; 3) Duration of the first CEUS loop is shorter than 45s. Finally, 420 patients including 136 benign and 284 malignant cases were enrolled in our study, shown as [Fig. 2](#). [Table 1](#) indicates the amounts of benign and malignant cases with their corresponding subtypes. Patients were randomly divided into training and testing cohorts with 3:1 ratio (4-fold cross-validation) for algorithm development and evaluation.

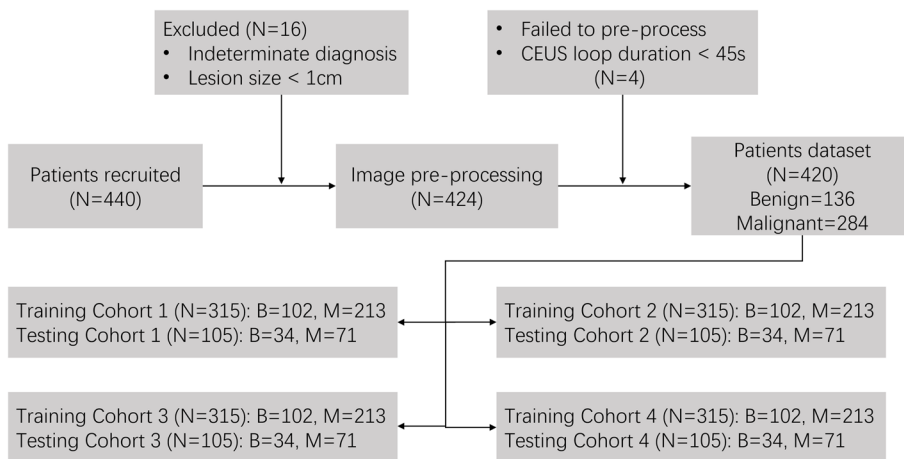
### Method

In this section, we will introduce the proposed method and two methods for comparison.





**Fig. 1** CEUS sequences of two typical cases. Each small-size CEUS image was selected from the CEUS loop at one second interval starting from AP onset time. **a** CEUS frames of a typical malignant liver lesion. **b** CEUS frames of a typical benign liver lesion



**Fig. 2** The inclusion and exclusion criteria for patient cohort

**Table 1** Amount of the benign and malignant cases with their corresponding subtypes

Type	Subtype	Amount	Diagnoses Methods
Benign	HA	55	Histologic examination: 20 cases Clinic follow-up: 35 cases
	HEM	60	Histologic examination: 35 cases Contrast MRI scanning: 25 cases
	FNH	13	Histologic examination
	RN	8	Histologic examination
Malignant	HCC	257	Histologic examination: 94 cases <sup>a</sup> EASL guidelines [23]: 163 cases
	IHCC	7	Histologic examination
	HM	20	Histologic examination

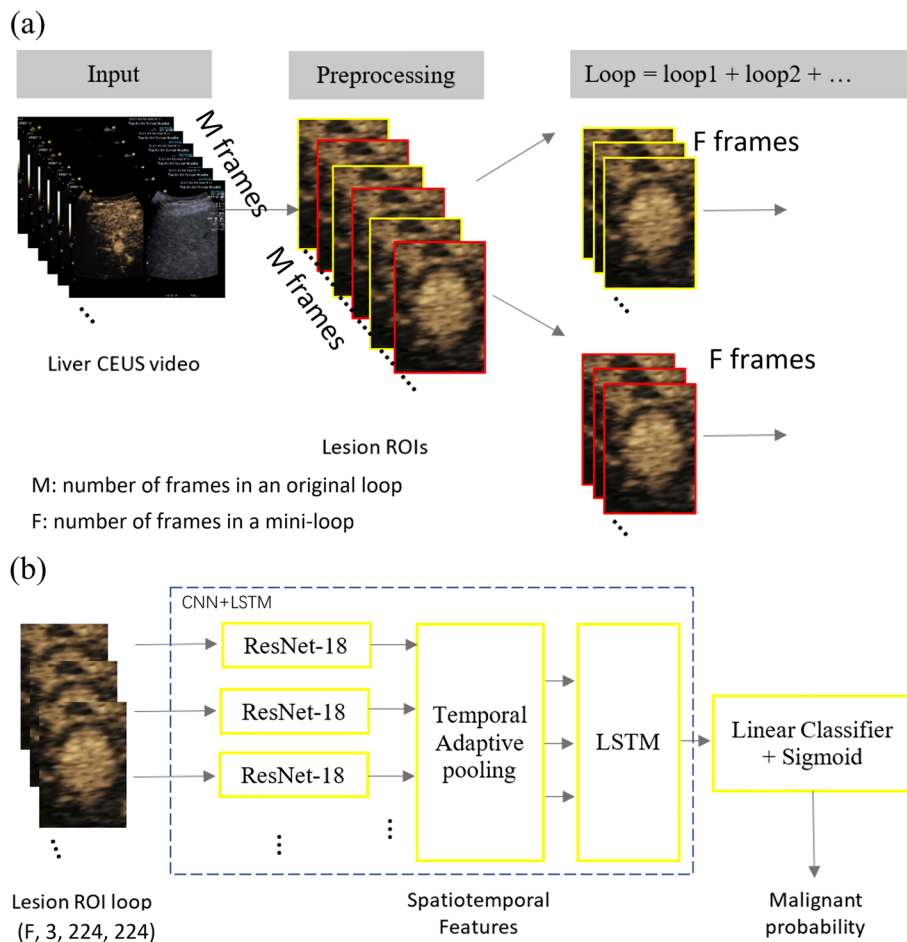
Abbreviations: HA Hepatic abscess, HEM Hemangioma, FNH Focal nodular hyperplasia, RN Regenerative nodule, HCC Hepatocellular Carcinoma, IHCC Intrahepatic Cholangiocarcinoma, HM Hepatic Metastasis

<sup>a</sup>The diagnosis was confirmed by following the EASL guidelines

**Proposed method**

The most intuitive way to employ deep learning technique in the diagnosis of liver lesion with CEUS is to feed the CEUS loops into a 3D CNN model because 3D CNN preserves both spatial and temporal features. However, the 3D CNN has high computational complexity. With limited number of datasets, the model can be easily over-fitted. To overcome this issue, we proposed to use the CNN-LSTM model to automatically make the differential diagnosis of liver lesion based on CEUS loops. In this method, CNN extracts spatial features of each frame, followed by the LSTM to learn the temporal features from the sequences of these spatial features. The flowchart of the method is shown in Fig. 3 and the detailed description are as follows.

*Crop Region of Interest Sequence (ROIs) of liver lesion.* Generally in clinical practice, the clinician would label four markers to denote long and short axis of lesion in a selected frame, which can be utilized as the bounding



**Fig. 3** Flowchart of the proposed method. **a** Flowchart of the pre-processing step. **b** Flowchart of the CNN-LSTM mode

box of lesion. To reduce the effect of respiratory motion, the size of this bounding box is expanded by 50%.

*Split whole CEUS cine-loops into several mini-loops.* As mentioned, an expert clinician can give reliable and accurate diagnosis results for liver lesions mainly based on typical CEUS frames from AP, PVP and LP. Leveraging this domain knowledge, we selected F (e.g. F=4) frames at equal time intervals from each phase (AP, PVP, LP) to reduce the impact from out of plane frames which were caused by respiratory motion. The removal work was automatic with no visual inspection or manual interaction required from a radiologist. It was done through a TIC curve fitting based method. For a given frame, if the absolute difference between its average intensity and the TIC fitted intensity was larger than 2 times of the standard deviation of the average intensity of all frames, it was removed. Since the duration of three phases are not the same, different number of F-frame may be generated. As suggested by guidelines such as CEUS LI-RADS [23], the last few seconds of PVP and LP are the most important frames. With this prior information, we select F frames from back to front on the timeline of PVP and LP. Finally, the entire CEUS loop from each patient was split into multiple mini-loops and each of which has a malignancy probability as the output of the proposed network. The final prediction result is generated by the majority vote of mini-loops. Compared with using the whole loop directly, splitting each CEUS loop into multiple mini-loops can increase sample size and mitigate the risk of overfitting. This can increase the total number of CEUS loops inputting into the network, which in principle equivalent to augment the data temporally.

*Map each mini-loop to malignant probability with CNN-LSTM model.* We employed residual networks because of its good performance and simplicity. Each ROI image feeds to the ResNet-18 [24] network (ResNet) to extract spatial features. The ResNet18 was pretrained based on ImageNet by freezing the first 5 layers. Then the F\*3 dimensional spatial features of mini-loop are smoothed to be 3 dimensional via temporal adaptive pooling in order to mimic the features from AP, PVP and LP and to reduce the respiratory motion at the same time. Afterwards, 3 dimensional spatial features are fed to the LSTM [25] to extract spatio-temporal features through adaptive average pooling layer. The LSTM contains two layers. The first layer is bidirectional LSTM to have sequence information in both directions backwards and forwards. The number of features in the hidden state is 128. Finally spatio-temporal features from every mini-loop are used to predict malignancy probability.

### **Methods for comparison**

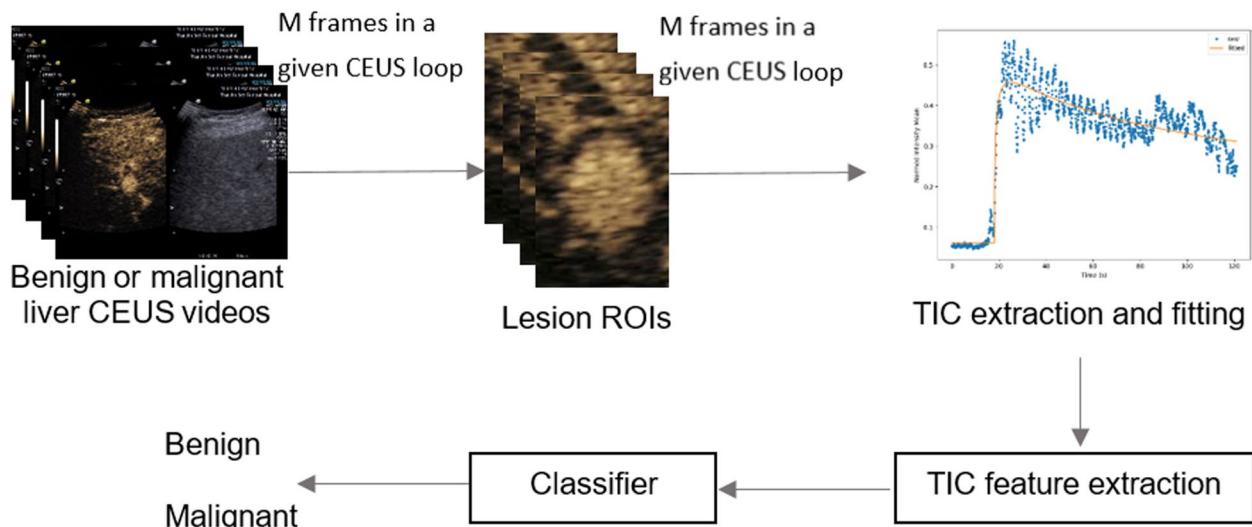
In this paper, we compared the proposed method with a 3D-CNN based method and a classical ML-based TIC method.

*3D-CNN based method.* As mentioned, the most intuitive way to process video based data is to employ a 3D-CNN. With limited number of datasets as well as memory, one can not directly input the CEUS loop into a 3D-CNN. Thus, the same pre-processing step as Fig. 3a was used. Instead of using CNN+LSTM as Fig. 3b, 3D-CNN was employed in this method.

*ML-based TIC method.* The block diagram of this method is shown in Fig. 4. TIC was firstly generated using the signal intensity mean value of a manually drawn ROI within the lesion. Due to respiratory motion, the TIC is usually very noisy and curve fitting is required. There are five different perfusion models commonly used: the log-normal distribution, the gamma variate function, local density random walk (LDRW), first passage time (FPT), and the lagged normal function [26]. In practice, there is no standards or guideline for choosing the best perfusion model. Moreover, using single perfusion model to estimate perfusion parameters is very sensitive to either the initial values for perfusion parameters or the selection of appropriate boundary conditions. Therefore, we have implemented a Multi-Model Framework, which applies nonlinear regression optimizing method to all five perfusion models. Afterwards, the perfusion parameters were chosen from the model with the lowest estimation error by evaluating various curve-fitting models and selecting the optimal solution using an exhaustive search technique. This framework was implemented in-house in Python. The perfusion parameters include AP onset time, peak time, peak intensity, area under the curve, mean transit time, gradient of wash in, gradient of wash out. Finally, all of these parameters plus the lesion size were input into a classifier. XGBoost was used as the classifier and it was chosen based on experiments with best results.

In this work, a NVIDIA GPU GeForce GTX 1080 Ti 11 GB was adopted to train different models. K-fold cross-validation (for example: k=4 here) method was performed to avoid the sampling bias by randomly partitioning dataset stratified by benign and malignant. For the proposed and 3D-CNN based method, models were trained by a stochastic gradient descent solver with a cross-entropy loss function for 6 epochs in total, with learning rate 0.01 and 0.001 in the first 4 epochs and in the last 2 epochs.





**Fig. 4** Diagram of ML-based TIC method

**Evaluation Metrics**

In this study, four evaluation metrics were used to evaluate the model performance, namely: receiver operating characteristic (ROC) curve, the area under the curve (AUC), sensitivity and specificity. Sensitivity and specificity can be calculated as follows:

$$sensitivity = \frac{TP}{TP + FN} \tag{1}$$

$$specificity = \frac{TN}{TN + FP} \tag{2}$$

TP, TN, FP and FN represent the number of true positives, true negatives, false positives and false negatives.

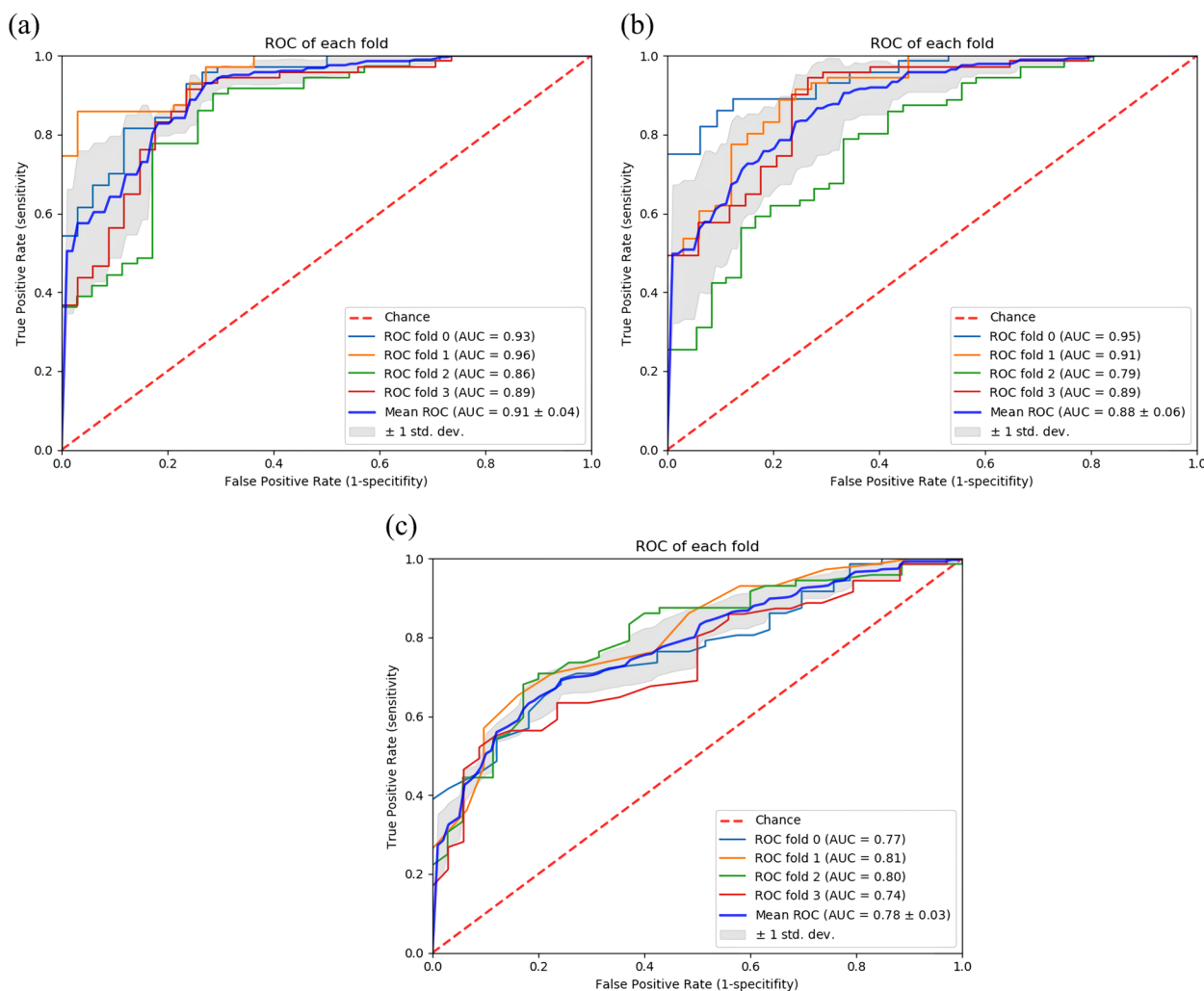
**Results**

The area under the receiver operation characteristic (ROC) curve (AUC) with 4-fold cross-validation was adopted as the quantitative metrics for evaluating the performance of the proposed algorithm for distinguishing between benign and malignant liver lesions. Mean AUCs and their corresponding standard deviations for each fold validation are shown in Fig. 5. The results demonstrate that the mean AUC evaluated is 0.91, 0.88, and 0.78 for the proposed CNN-LSTM based method, the 3D-CNN based method and the ML-based TIC method, respectively. According to previous studies in literature, the AUC for an experienced observer (more than 20 years of performing and interpreting CEUS) and an inexperienced observer (clinician without prior CEUS experience) to make malignancy diagnosis of liver lesions was 84% and 72%, respectively [6]. Although the datasets are different, this result once again confirmed that the deep

learning based methods are promising in providing computer aided diagnosis for liver lesion in CEUS. More specifically, we show in Table 2 the changes in sensitivity and specificity of the three methods when applying different thresholds for the predicted malignancy probabilities. The sensitivity results increased from 0.83 to 0.95 and the specificity decreased from 0.82 to 0.7 with the threshold by using the purposed CNN-LSTM model. The 3D-CNN based method shows the highest sensitivity of 0.96, but with the specificity of 0.55; the ML-based TIC method also shows the lowest specificity of 0.21 when the sensitivity is 0.96.

**Discussion**

This paper demonstrated an end-to-end methodology from data preprocessing to model implementation to successfully make computer aided diagnosis of liver lesions in CEUS. Compared with the traditional ML-based TIC method, the most intuitive 3D CNN based method used as benchmark in this work and the previous literatures [9, 10], it does not require any additional manual features selection and CEUS frames selection but has a higher AUC. With the proposed method, the only manual input that a clinician may need to do is to draw a ROI to indicate the lesion when he/she doesn't measure the lesion before entering CEUS mode, then the lesion benign/malignancy will be automatically prediction by the proposed method. The main contributions of this proposed method include 1) Automatically sample frames from the entire CEUS loops leveraging clinical domain knowledge, cascade them into mini-loops as the input of the model, providing dynamic information of CEUS in a more effective way. 2) An efficient variant



**Fig. 5** ROC curves of the proposed method, 3D-CNN based method and ML-based TIC method with fold parameter ranging from 0 to 3. **a** Proposed method. **b** 3D-CNN based method. **c** ML-based TIC method

**Table 2** The sensitivity and specificity of three methods. Operator 1, 2, 3 represent different thresholds of the predicated malignant probability

Method/ (Sensitivity, Specificity)	Operator 1	Operator 2	Operator 3
CNN-LSTM	(0.83, 0.82)	(0.91, 0.74)	(0.95, 0.70)
3D CNN	(0.83, 0.74)	(0.91, 0.67)	(0.96, 0.55)
TIC based	(0.87, 0.42)	(0.90, 0.36)	(0.96, 0.21)

model of CNN, i.e. CNN-LSTM, was applied to avoid the potential overfitting problem and to extract spatio-temporal features. As far as we know, this is the first time to employ CNN-LSTM in analysis whole contrast enhanced ultrasound video cine-loops (i.e., including all CEUS

perfusion phases, AP, PVP and LP). 3) A temporal adaptive pooling is designed and added in the model to reduce the impact of respiratory motion significantly.

Furthermore, we invited three clinicians with 5, 8 and 10 years of CEUS experience to analyze the misdiagnosed cases predicted from the proposed method. Eight misdiagnosed cases were selected and got in-depth reviews, including four benign cases (misdiagnosed as malignancy) and four malignant cases (misdiagnosed as benign). Among the eight cases, two FNH cases were presented with mild washout in the LP, two HA cases showed fast wash in and fast wash out and two HCC cases demonstrated with classical hyper enhancement in AP but lack of markable washout during LP. All three clinicians confirmed that these patterns are actually atypical in clinical. In other words, six of them presented with features that does not usually



belong to their own class. The remaining two cases are HM, which has the least number of cases in the training data. To conclude, it is noted that all of above cases were considered difficult to diagnosis even for an experienced clinician when only CEUS loops are provided.

We should mention that clinicians usually make a diagnosis decision based on comprehensive information including patient's medical history, related surgical examination, and even medical images from other modalities as well. We also noticed that there are some promising results using regular B-mode [11, 27]. Therefore, our future work include: 1) Increase the variability of dataset including more cases such as HM; 2) Combine patient's information not only in CEUS; 3) compare with the performance of clinicians with different level of experience.

## Conclusions

The study successfully developed an end-to-end deep learning algorithm to classify the liver lesions between benign or malignant effectively. The proposed algorithm with promising results has a high potential to narrow the gap in liver lesion diagnosis between radiologists with different experience levels. Once the proposed model with convincing results is translated into clinical workflow, we believe that it can become a second opinion for the radiologist or the ultrasound doctor to help them make a precise diagnosis.

## Acknowledgements

Not applicable.

## Authors' contributions

HZ: Conceptualization, Data curation, Writing - Original Draft, Validation. JD: Jianmin Ding: Data curation. YZ: Data curation. YW: Data curation. LZ: Data curation. CCS: Writing-Original Draft. JX: Writing-Original Draft. JW: Methodology, Software. LT: Writing - Review and Editing. ZC: Conceptualization, project administration, Review and Editing. QL: Conceptualization, Methodology, Software, Validation. XJ: Conceptualization, Supervision, Funding acquisition. All authors read and approved the final manuscript.

## Funding

This work was supported by National Natural Science Foundation of China (Grant No. 82371986), Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-074C) and Tianjin Health Research Project (Grant No. ZD20014).

## Availability of data and materials

The datasets in this study are not publicly available due to institutional review board.

## Declarations

### Ethics approval and consent to participate

All methods were carried out in accordance with relevant guidelines and regulations. All the experimental protocols were approved by the institutional review board of Ethics Committee of Tianjin Third Central Hospital (IRB2019-005-02) as well as the requirement for informed consent was waived by the institutional review board of Ethics Committee of Tianjin Third Central Hospital since this was a retrospective study.

## Consent for publication

Not applicable.

## Competing interests

Cho-Chiang Shih, Jingping Xu, Jianan Wang, Ling Tong, Zhouye Chen and Qizhong Lin are employees of Philips Research China, which owns the property of the presented program. The other 6 authors declare that they have no competing interests.

Received: 10 March 2023 Accepted: 14 March 2024

Published online: 21 March 2024

## References

- Shiraishi J, Sugimoto K, Moriyasu F, Kamiyama N, Doi K. Computer-aided diagnosis for the classification of focal liver lesions by use of contrast-enhanced ultrasonography. *Med Phys*. 2008;35(5):1734–46.
- Gomaa A, Waked I. Management of advanced hepatocellular carcinoma: review of current and potential therapies. *Hepatoma Res*. 2017;3:112–22.
- Dietrich C, Nolsøe C, Barr R, Berzigotti A, Burn SP, Cantisani V, et al. Guidelines and Good Clinical Practice Recommendations for Contrast-Enhanced Ultrasound (CEUS) in the Liver-Update 2020 WFUMB in Cooperation with EFSUMB, AFSUMB, AIUM, and FLAUS. *Ultrasound Med Biol*. 2020;46(10):2579–604.
- Dietrich CF, Nolsøe CP, Barr RG, Berzigotti A, Burns PN, Cantisani V, et al. Guidelines and good clinical practice recommendations for contrast enhanced ultrasound (CEUS) in the liver-update 2020-WFUMB in cooperation with EFSUMB, AFSUMB, AIUM, and FLAUS. *Ultraschall Med-Eur J Ultrasound*. 2020;41(05):562–85.
- Wu K, Chen X, Ding M. Deep learning based classification of focal liver lesions with contrast-enhanced ultrasound. *Optik*. 2014;125(15):4057–63.
- Ta C, Kono Y, Eghtedari M, Oh Y, Robbin M, Barr R, et al. Focal Liver Lesions: Computer-aided Diagnosis by Using Contrast-enhanced US Cine Recordings. *Radiology*. 2018;286(3):1062–71.
- Kondo S, Takagi K, Nishida M, Iwai T, Kudo Y, Ogawa K, et al. Computer-Aided Diagnosis of Focal Liver Lesions Using Contrast-Enhanced Ultrasonography With Perflubutane Microbubbles. *IEEE Trans Med Imaging*. 2017;36(7):1427–37. <https://doi.org/10.1109/TMI.2017.2659734>.
- Guo LH, Wang D, Qian YY, Zheng X, Zhao CK, Li XL, et al. A two-stage multi-view learning framework based computer-aided diagnosis of liver tumors with contrast enhanced ultrasound images. *Clin Hemorheol Microcirc*. 2018;69(3):343–54.
- Vetter M, Waldner MJ, Zundler S, Klett D, Bocklitz T, Neurath MF, et al. Artificial Intelligence for the classification of focal liver lesions in ultrasound - a systematic review. *Ultraschall Med*. 2023;44:395–407.
- Wei Q, Tan N, Xiong S, Luo W, Xia HX, Luo B. Deep Learning Methods in Medical Image-Based Hepatocellular Carcinoma Diagnosis: A Systematic Review and Meta-Analysis. *Cancers*. 2023;15(23):1–21.
- Nishida N, Yamakawa M, Shiina T, Mekada Y, Nishida M, Sakamoto N, et al. Artificial intelligence (AI) models for the ultrasonographic diagnosis of liver tumors and comparison of diagnostic accuracies between AI and human experts. *J Gastroenterol*. 2022;57(4):309–21.
- Mitrea DA, Brehar R, Nedevschi S, Lupsor-Platon M, Socacia M, Badea R. Hepatocellular Carcinoma Recognition from Ultrasound Images Using Combinations of Conventional and Deep Learning Techniques. *Sensors*. 2023;23(5):1–29.
- Elbashir MK, Mahmoud A, Mohamed MA, Hamouda E, Alruily M, Alotaibi SM, et al. A Transfer Learning Approach Based on Ultrasound Images for Liver Cancer Detection. *Comput Mater Continua*. 2023;75(3):5105–21.
- Zhang Wb, Hou Sz, Chen Yi, Mao F, Dong Y, Chen Jg, et al. Deep learning for approaching hepatocellular carcinoma ultrasound screening dilemma: Identification of a-Fetoprotein-negative hepatocellular carcinoma from focal liver lesion found in high-risk patients. *Front Oncol*. 2022;12:1–9.
- Brehar R, Mitrea DA, Vancea F, Marita T, Nedevschi S, Lupsor-Platon M, et al. Comparison of Deep-Learning and Conventional Machine-Learning Methods for the Automatic Recognition of the Hepatocellular Carcinoma Areas from Ultrasound Images. *Sensors*. 2020;20(11):1–22.

16. Liu D, Liu F, Xie X, Su L, Liu M, Xie X, et al. Accurate prediction of responses to transarterial chemoembolization for patients with hepatocellular carcinoma by using artificial intelligence in contrast-enhanced ultrasound. *Eur Radiol.* 2020;30(4):2365–76.
17. Pan F, Huang Q, Li X. Classification of liver tumors with CEUS based on 3D-CNN. In: 2019 IEEE 4th International Conference on Advanced Robotics and Mechatronics (ICARM). 2019. pp. 845–9. <https://doi.org/10.1109/ICARM.2019.8834190>.
18. Căleanu CD, Sîrbu CL, Simion G. Deep neural architectures for Contrast Enhanced Ultrasound (CEUS) Focal liver lesions automated diagnosis. *Sensors.* 2021;21(12):1–19. <https://www.mdpi.com/1424-8220/21/12/4126>.
19. Hu H, Wang W, Chen L, Ruan S, Chen S, Li X, et al. Artificial intelligence assists identifying malignant versus benign liver lesions using contrast-enhanced ultrasound. *J Gastroenterol Hepatol.* 2021;36(10):2875–83.
20. Kamiyama N, Sugimoto K, Nakahara R, Kakegawa T, Itoi T. Deep learning approach for discrimination of liver lesions using nine time-phase images of contrast-enhanced ultrasound. *J Med Ultrason.* 2024;51(1):83–93.
21. Feng X, Cai W, Zheng R, Tang L, Zhou J, Wang H, et al. Diagnosis of hepatocellular carcinoma using deep network with multi-view enhanced patterns mined in contrast-enhanced ultrasound data. *Eng Appl Artif Intell.* 2023;118:105635. <https://doi.org/10.1016/j.engappai.2022.105635>.
22. Kim N, Lee WJ, Lee HJ. Deep learning classification of focal liver lesions with contrast-enhanced ultrasound from arterial phase recordings. In: 2023 International Conference on Electronics, Information, and Communication (ICEIC). 2023. pp. 1–3. <https://doi.org/10.1109/ICEIC57457.2023.10049872>.
23. Kono Y, Lyshchik A, Cosgrove D, Dietrich CF, Jang HJ, Kim TK, et al. Contrast enhanced ultrasound (CEUS) liver imaging reporting and data system (LI-RADS<sup>®</sup>): the official version by the American College of radiology (ACR). *Ultraschall Med-Eur J Ultrasound.* 2017;38(01):85–6.
24. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Los Alamitos: The Institute of Electrical and Electronics Engineers, Inc.; 2016. p. 770–8.
25. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput.* 1997;9(8):1735–80.
26. Akhbardeh A, Sagreiya H, El Kaffas A, Willmann JK, Rubin DL. A multi-model framework to estimate perfusion parameters using contrast-enhanced ultrasound imaging. *Med Phys.* 2019;46(2):590–600.
27. Yang Q, Wei J, Hao X, Kong D, Yu X, Jiang T, et al. Improving B-mode ultrasound diagnostic performance for focal liver lesions using deep learning: A multicentre study. *EBioMedicine.* 2020;56:102777.

## Publisher's Note

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