



Machine Learning-Based Ultrasonics Improves the Diagnostic Performance in Differentiating Focal Nodular Hyperplasia and Atypical Hepatocellular Carcinoma

Wei Li^{1†}, Xiao-Zhou Lv^{2†}, Xin Zheng¹, Si-Min Ruan¹, Hang-Tong Hu¹, Li-Da Chen¹, Yang Huang¹, Xin Li³, Chu-Qing Zhang⁴, Xiao-Yan Xie¹, Ming Kuang^{1,5}, Ming-De Lu^{1,5}, Bo-Wen Zhuang^{1*} and Wei Wang^{1*}

OPEN ACCESS

Edited by:

Youyong Kong,
Southeast University, China

Reviewed by:

Yuming Jiang,
Stanford University, United States
Jinghao Duan,
Shandong University, China
Yan Wu,
Second Affiliated Hospital of Harbin
Medical University, China

*Correspondence:

Wei Wang
wangw73@mail.sysu.edu.cn
Bo-Wen Zhuang
zhuangbw3@mail.sysu.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

Received: 23 March 2020

Accepted: 03 March 2021

Published: 26 March 2021

Citation:

Li W, Lv X-Z, Zheng X, Ruan S-M,
Hu H-T, Chen L-D, Huang Y, Li X,
Zhang C-Q, Xie X-Y, Kuang M,
Lu M-D, Zhuang B-W and Wang W
(2021) Machine Learning-Based
Ultrasonics Improves the Diagnostic
Performance in Differentiating Focal
Nodular Hyperplasia and Atypical
Hepatocellular Carcinoma.
Front. Oncol. 11:544979.
doi: 10.3389/fonc.2021.544979

¹ Department of Medical Ultrasonics, Institute of Diagnostic and Interventional Ultrasound, Ultrasonics Artificial Intelligence X-Lab, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China, ² Department of Traditional Chinese Medicine, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, ³ Research Center, GE Healthcare, Shanghai, China, ⁴ Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China, ⁵ Department of Hepatobiliary Surgery, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China

Background: The typical enhancement patterns of hepatocellular carcinoma (HCC) on contrast-enhanced ultrasound (CEUS) are hyper-enhanced in the arterial phase and washed out during the portal venous and late phases. However, atypical variations make a differential diagnosis both challenging and crucial. We aimed to investigate whether machine learning-based ultrasonic signatures derived from CEUS images could improve the diagnostic performance in differentiating focal nodular hyperplasia (FNH) and atypical hepatocellular carcinoma (aHCC).

Patients and Methods: A total of 226 focal liver lesions, including 107 aHCC and 119 FNH lesions, examined by CEUS were reviewed retrospectively. For machine learning-based ultrasonics, 3,132 features were extracted from the images of the baseline, arterial, and portal phases. An ultrasonics signature was generated by a machine learning model. The predictive model was constructed using the support vector machine method trained with the following groups: ultrasonics features, radiologist's score, and combination of ultrasonics features and radiologist's score. The diagnostic performance was explored using the area under the receiver operating characteristic curve (AUC).

Results: A total of 14 ultrasonics features were chosen to build an ultrasonics model, and they presented good performance in differentiating FNH and aHCC with an AUC of 0.86 (95% confidence interval [CI]: 0.80, 0.89), a sensitivity of 76.6% (95% CI: 67.5%, 84.3%), and a specificity of 80.5% (95% CI: 70.6%, 85.9%). The model trained with a combination of ultrasonics features and the radiologist's score achieved a significantly higher AUC (0.93, 95% CI: 0.89, 0.96) than that trained with the radiologist's score (AUC: 0.84, 95% CI: 0.79, 0.89, $P < 0.001$). For the sub-group of HCC with normal AFP value, the model trained with a combination of ultrasonics features, and the radiologist's score remain achieved the highest AUC of 0.92 (95% CI: 0.87, 0.96) compared to that with the

ultrasonics features (AUC: 0.86, 95% CI: 0.74, 0.89, $P < 0.001$) and radiologist's score (AUC: 0.86, 95% CI: 0.79, 0.91, $P < 0.001$).

Conclusions: Machine learning-based ultrasonics performs as well as the staff radiologist in predicting the differential diagnosis of FNH and aHCC. Incorporating an ultrasonics signature into the radiologist's score improves the diagnostic performance in differentiating FNH and aHCC.

Keywords: ultrasonography, machine learning, focal nodular hyperplasia, hepatocellular carcinoma, ultrasonics

INTRODUCTION

The typical enhancement pattern of hepatocellular carcinoma (HCC) on contrast-enhanced ultrasound (CEUS) is characterized by hyper-enhancement in the arterial phase and wash out during the portal venous and late phases (1). However, atypical variations occur, especially in some well-differentiated tumors, accounting for 5–41% of HCC cases; such lesions may show sustained hyper-/iso-enhancement in the portal venous and late phases and are defined as atypical HCC (aHCC) (2–4). Meanwhile, most benign focal liver lesions show complete hyper- or iso-enhancement in the portal venous and late phases, making differential diagnosis both crucial and challenging (5, 6). This diagnostic difficulty could be resolved using CEUS techniques, such as micro-flow imaging to further characterize the enhancement features in the arterial phase, e.g., a spoke-wheel artery for focal nodular hyperplasia (FNH) and chaotic vessel for HCC (7–10). However, the interpretation of features involves the experience of radiologists, making inter-reader variability inevitable.

In contrast to the traditional practice of treating medical images as pictures intended solely for visual interpretation, radiomics features could reflect not only the macroscopic manifestation but also the cellular and molecular nature of tissues (11–13). Radiomics offers a vast scale of imaging biomarkers that could potentially assist in detecting and diagnosing, evaluating the prognosis and predicting the therapeutic response, and monitoring the disease status of cancer (11, 12, 14–16). Machine learning-based ultrasonics approaches, derived from radiomics, involve the analysis and transformation of ultrasound images into large sets of quantitative data and have been identified as potential alternatives to detect and classify lesions (17, 18).

Recently, few applications of machine learning in HCC diagnosis have been reported (19, 20). Most machine learning systems have demonstrated excellent diagnostic performance, with the area under the receiver operating characteristic curve (AUC) of 0.89–0.97 for HCC characterization (19, 20). Gatos et al. applied radiomics to segment and classify focal liver lesions on non-enhanced T2-weighted images, providing a non-invasive method for assessing liver lesions (21). Some studies have shown that multi-modal ultrasound images also perform well for the detection and classification of focal liver lesions (19, 22, 23). However, most studies have only compared the diagnostic performance between machine learning systems and radiologists. The influence of the performance of these systems on radiologists when used in clinical practice has not been

evaluated. Thus, the added clinical value of machine learning systems to observers is necessary to determine and validate.

The purpose of our research was to develop a machine learning-based ultrasonics approach to assess ultrasonics features for improving the diagnostic performance in differentiating FNH and aHCC.

PATIENTS AND METHODS

Patients

This retrospective analysis obtained ethical approval and waived the informed consent requirement. From December 2013 to January 2018, 119 patients with FNH and 107 patients with aHCC lesions were included in the study based on the inclusion and exclusion criteria. The inclusion criteria were as follows: (a) CEUS was performed; (b) lesions were visually hyper-enhanced during the arterial phase and sustained hyper- or iso-enhanced during the portal venous and late phases; (c) HCC was diagnosed by pathological examinations and FNH was confirmed by pathological examinations or supported by CT or MRI findings with a minimum 1 year follow-up; and (d) no treatment was conducted before CEUS. Patients were excluded if they had multiple tumors. Baseline clinical trial data, including age, gender, and some blood test, such as hepatitis background and alpha-fetoprotein (AFP), were performed no more than 7 days before or after the CEUS examination.

Image Acquisition

US examinations were performed using an Aplio 500 scanner (Canon Medical Systems, Tokyo, Japan), equipped with a 375BT convex transducer (frequency, 3.5 MHz) and an Aixplorer scanner (Supersonic, Paris, France) with an SC6-1 curvilinear transducer (frequency, 1–6 MHz). Contrast harmonic imaging (CHI) and contrast pulse sequencing (CPS) were used with a mechanical index of 0.06–0.10. Baseline ultrasonography was performed to scan the liver thoroughly before CEUS. Additionally, the target lesions were identified and observed carefully during the baseline observation in B mode. The imaging settings, such as the gain, depth, and focus, were optimized for each examination. After the CHI or CPS mode was activated, a bolus intravenous injection of 2.4 mL of SonoVue (Bracco, Milan, Italy) was administered, followed by flushing with 5 mL of saline. The targeted lesion was observed

continuously for 5 minutes. The arterial, portal venous, and late phases were defined as 10–30 seconds, 31–120 seconds, and 121–300 seconds after injection, respectively. CEUS examinations were performed by one of two radiologists (WW and X-YX) with at least 10 years' experience of performing CEUS. Three images from the same section, which showed the maximum observation of the target lesion, were taken from each patient, (a) a baseline ultrasound; (b) an arterial phase image at the enhancement time of 25–30 seconds; and (c) a portal phase image at the enhancement time of 60–70 seconds.

Radiologist's Scoring

Two staff radiologists (B-WZ and L-DC) reviewed ultrasound images and videos retrospectively, and they had more than 5 years of experience in assessing liver CEUS data. The radiologists were not involved in the feature extraction process below. All patient identification information from the images was removed, and the researchers were unaware of all the clinicopathological

information. The diagnostic criteria for HCC and FNH were based on the 2012 guidelines issued by the European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) (2). The diagnostic criteria for HCC were the manifestation of basket pattern and/or chaotic vessels (**Figure 1A**) and non-enhanced areas (**Figure 1B**). The diagnostic criteria of CEUS feature for FNH were centrifugal enhancement (Video 1), spoke-wheel artery (**Figure 1C**), unenhanced central scar (**Figure 1D**), and feeding artery.

For each group, the diagnostic confidence was scored using a subjective three-point scale (grade 1, definitely or most likely FNH; grade 2, indeterminate; and grade 3, most likely or definitely HCC). If there was inconsistency, we performed a consensus reading, and the consensus data were used for subsequent analysis.

Ultrasonics Feature Extraction

Digital imaging and communications in medicine (DICOM) images were used to extract ultrasonics features using the in-

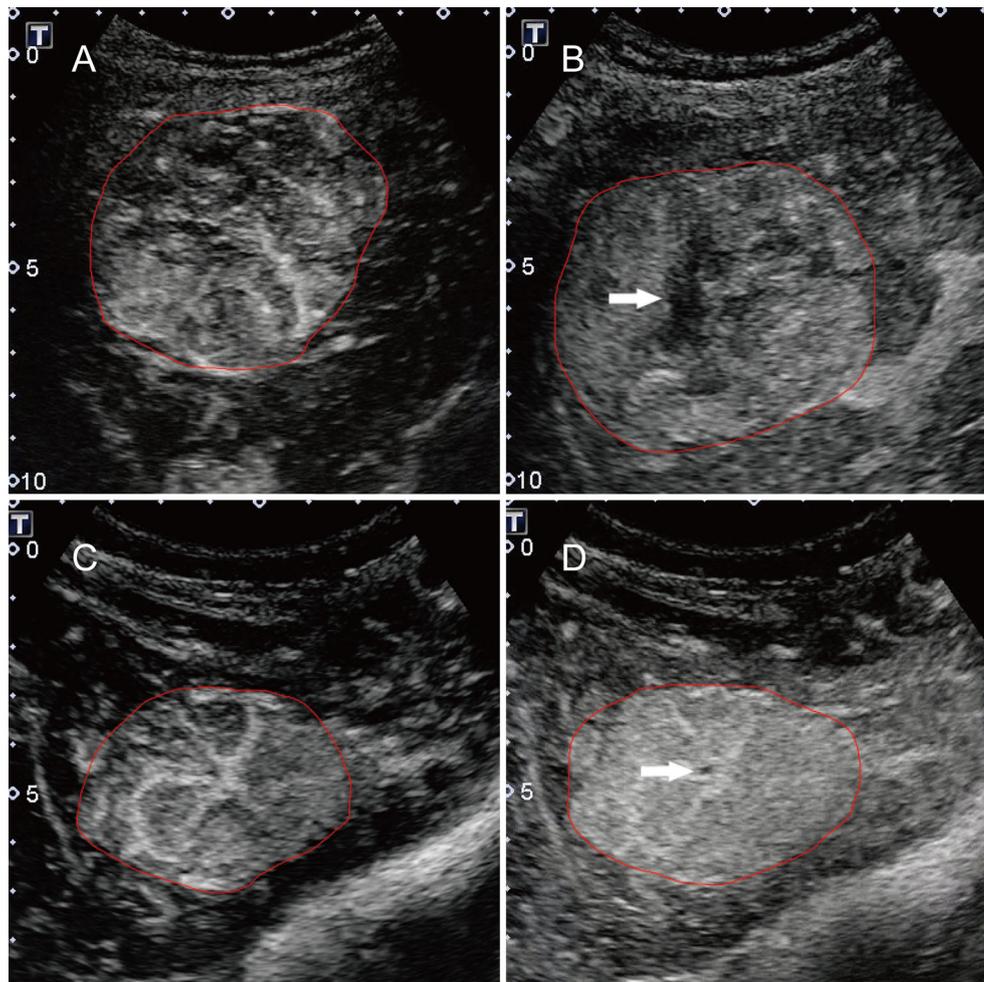


FIGURE 1 | Typical features for HCC and FNH lesions. **(A)** the basket pattern and/or chaotic vessels; **(B)** non-enhancing areas (arrow); **(C)** spoke-wheel arteries; and **(D)** unenhanced central scar (arrow). Annotations of the ROI generated by the radiologists around the tumor outline are delineated in red.

house designed Ultrasonics-Platform software (Version 1.0; Ultrasonics Artificial Intelligence X-lab, Guangzhou, China). After an image was imported, the radiologist drew a region of interest (ROI) on the largest cross-section along the tumor contour. Next, the software automatically extracted the features from the ROIs. In total, 1,044 features could be extracted from a single image. These 1,044 features extracted from a single image consisted of five categories of features: histogram parameters, textural parameters, form factor parameters, grey-level co-occurrence matrix (GLCM) parameters, and run length matrix (RLM) parameters. Detailed information on the features is provided in **Supplementary Material S1**. Finally, 3,132 features were extracted from the baseline US, arterial phase and portal phase of CEUS images of each patient. Initially, two radiologists (WL and YH, with at least 5 years of experience in performing US examinations) were required to trace out ROIs on the selected images. The inter-observer and intra-observer reproducibility in feature extraction were assessed and are described in **Supplementary Material S2**. The remaining images were delineated by the first radiologist.

Feature Selection and Model Development for Prediction

Of the 3,132 features from each patient, many were highly redundant, which could degrade the classification. We eliminated redundant features by using a two-step feature selection method. First, if two features were highly-correlated with a correlation coefficient higher than 0.95, one of the features was removed. Second, we eliminated features with an AUC less than 0.6. According to the Harrell guidelines for multivariate analysis, the number of events should be at least 10 times greater than that of the included covariates (24). The least absolute shrinkage and selection operator (LASSO) regression was used to perform the ultrasonics features selection in the training dataset. All ultrasonics feature values were normalized by using the mean and variance of the feature values to be within similar dynamic ranges.

A support vector machine (SVM) based on the radial basis function (RBF) kernel was trained from the selected feature subset produced by the preceding steps. The entire data set was randomly divided into a training dataset (comprising 80% of subjects) and a validation dataset (comprising the remaining 20% of subjects). The training dataset was used to construct a model, which was then evaluated using the validation dataset. A 10-fold cross-validation method was adopted to ensure the robustness of the classifiers to training and testing data. All processes were repeated 10 times with random seeds, generating 10 different training and validation datasets. We built the model using the training dataset and then evaluated it using the validation dataset repeatedly. Subsequently, the model with the best classification performance was selected as the best model.

Statistical Analysis

Descriptive statistics are summarized as the mean \pm standard deviation (SD) or median and interquartile range. Comparisons between groups were tested using Student's *t* test or the Mann-Whitney test for quantitative variables and the chi-squared test or Fisher's test for qualitative variables.

A weighted kappa statistics test was used to assess the two radiologists' scores. We evaluated the reproducibility of the ultrasonics feature extraction using the "irr" package in R. The LASSO regression was performed using the "glmnet" package.

All observations of patients with known outcomes were classified into three datasets (1): radiologist's score, (2) ultrasonics features, and (3) a combination of ultrasonics features and radiologist's score. The diagnostic performance of the radiologist's score was evaluated by plotting receiver operating characteristic (ROC) curves. The diagnostic performance in discriminating between FNH and aHCC is expressed as the AUC. The ultrasonics features and the combination of the ultrasonics features and radiologist's score were further compared through an SVM classifier using the "rattle" package in R. The performance of the SVM model was tested using the AUC. Paired comparisons of AUC values were performed by a two-sided Wilcoxon signed-rank test at a significance level of 5%. The predictive sensitivity (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (+LR), and negative likelihood ratio (-LR) were calculated at a cut-off point that maximized the value of the Youden index. Comparisons among the three datasets were performed using the Delong test. Decision curve analysis (DCA) was performed with the "dca.R" function. All statistical tests were two-sided tests, and $P < 0.05$ indicated statistical significance. All Statistical analyses were performed using R version 3.3.3 (<http://www.r-project.org/>).

RESULTS

Clinical Characteristics

The clinical characteristics are listed in **Table 1**. The study included 226 patients; 107 (47.3%) patients (mean age, 54.0 \pm 11.9 years old) had a final diagnosis of HCC; and the remaining 119 (52.7%) patients (mean age, 34.5 \pm 11.7 years old) had a final diagnosis of FNH. 20 FNH lesion were confirmed by pathological examinations (11 by biopsy, 9 by surgery), while 99 cases were supported by CT or MRI findings with a minimum one-year follow-up. No significant difference was found in the tumor number between the two groups ($P=0.118$). The average lesion size of FNH and HCC was 3.3 \pm 1.8 cm (range: 0.8-10.2 cm) and 4.8 \pm 3.4 cm (range: 0.8-18.6 cm), respectively.

Ultrasonics Signature Construction and Validation

After the feature selection and dimensional reduction process, 14 selected features were taken as the input of the SVM to train a prediction model, including 6 features derived from baseline US images, 3 from arterial phase images, and 4 from portal phase images (**Figure 2, Supplementary Material S3**). All feature values were normalized to achieve similar dynamic ranges. The parameter *C* which is used to control the error-margin trade-off was set at 1, and the kernel width sigma was 0.012. Next, the training and validation procedures for tumor classification were employed with 10-fold cross-validation.

TABLE 1 | Clinical Characteristics and Laboratory Information of the Patients.

Patients	FNH(N=119)	aHCC(N =107)	P value
Gender (male/female)	61/58	92/15	<0.001
Age (years)	34.5 ± 11.7	54.0 ± 11.9	<0.001
HBsAg (IU/ml)			<0.001
≤0.05	116 (97.5)	15 (14.0)	
0.05–250	1 (0.8)	34 (31.8)	
>250	2 (1.7)	58 (54.2)	
HBV-DNA (IU/mL)			<0.001
<100	118 (99.2)	46 (43.0)	
100–10 ⁵	0	44 (41.1)	
>10 ⁵	1 (0.8)	17 (15.9)	
HCV-Ab (S/CO)			0.212
<1.0	119 (100)	105 (98.1)	
≥1.0	0	2 (1.9)	
AFP (µg/L)			<0.001
<20	117 (98.3)	38 (35.5)	
20–400	2 (1.7)	40 (37.4)	
>400	0	29 (27.1)	
Tumor number			0.118
1	112 (94.1)	92 (86.0)	
2	4 (3.4)	8 (7.5)	
≥3	3 (2.5)	7 (6.5)	
Tumor size (cm)			<0.01
<3	3.34 ± 1.80	4.76 ± 3.35	
3–5	64 (53.8)	37 (34.6)	
>5	35 (29.4)	35 (32.7)	

Data are the number of patients, with the percentage in parentheses unless indicated. aHCC, atypical hepatocellular carcinoma.

Diagnostic Performance of Ultrasonics Features and Radiologist's Score Models

A total of 83 FNHs were correctly identified while 36 lesions were incorrectly identified as HCC, leading to a specificity of 69.8% by the radiologists. For combined model, 100 FNHs were correctly classified whereas 19 lesions were incorrectly assigned to HCC, resulting in a specificity of 84.0%. Comparing the performance of the radiologists' score and the combined model, twenty-four cases have a different result; consequently, the combined model leads to an additional 15 FNHs and 5 HCCs being correctly classified.

The model trained with the combination of the ultrasonics features and radiologist's score performed significantly better (AUC: 0.93, 95% CI: 0.89, 0.96) than that trained with the ultrasonics features (AUC: 0.86, 95% CI: 0.80, 0.89, $P < 0.001$) and radiologist's score (AUC: 0.84, 95% CI: 0.79, 0.89, $P < 0.001$). Overall, the model based on the radiologist's score had the highest diagnostic SEN of 94.4% (95% CI: 88.2%, 97.9%) but the lowest SPE of 69.8% (95% CI: 60.7%, 77.8%) compared with the combined model (SEN: 93.5% [95% CI: 87.0%, 97.3%], SPE: 84.9% [95% CI: 77.1%, 90.8%]), and the ultrasonics features model (SEN: 76.6% [95% CI: 67.5%, 84.3%], SPE: 80.5% [95% CI: 70.6%, 85.9%]). Furthermore, when the ultrasonics features were combined with the radiologist's score, the diagnostic performance was significantly improved in terms of the AUC, SPE, and PPV and +LR (AUC: 0.93, SPE: 84.9%, and PPV: 84.7%, +LR 6.2) compared with the performance of the other two models. The performance measurements of each dataset are reported in **Table 2** based on each ROC curves to distinguish between FNH and aHCC (**Figure 3**).

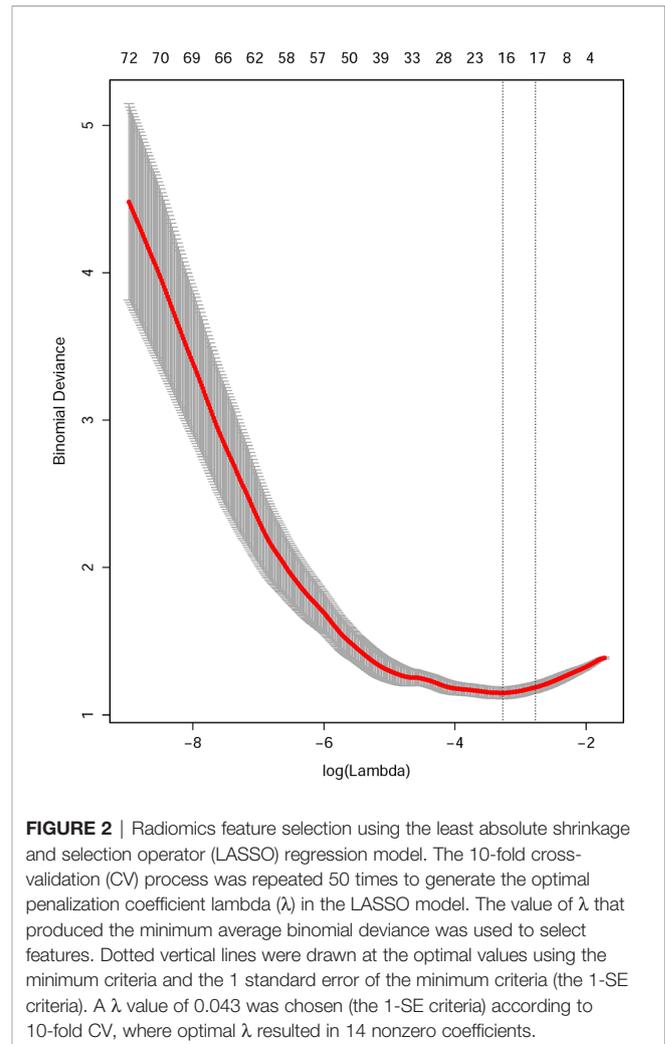


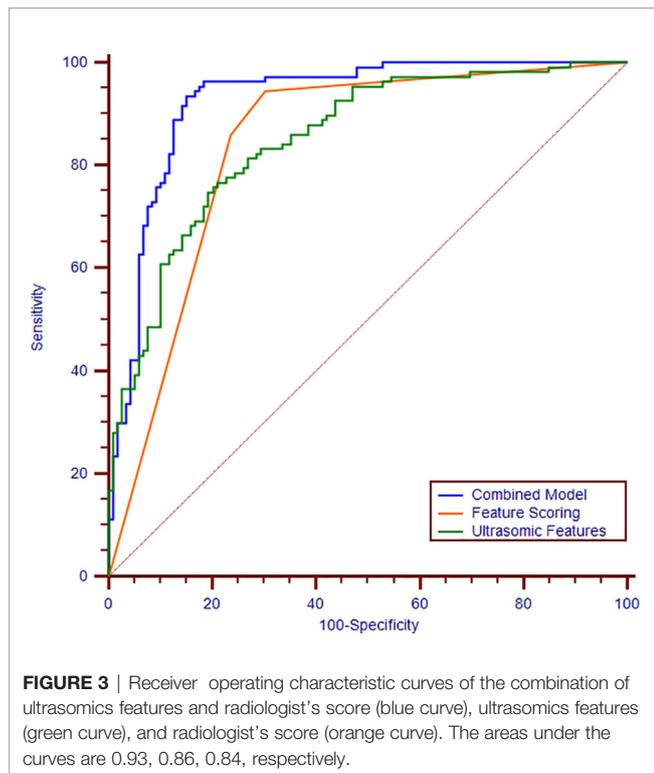
FIGURE 2 | Radiomics feature selection using the least absolute shrinkage and selection operator (LASSO) regression model. The 10-fold cross-validation (CV) process was repeated 50 times to generate the optimal penalization coefficient lambda (λ) in the LASSO model. The value of λ that produced the minimum average binomial deviance was used to select features. Dotted vertical lines were drawn at the optimal values using the minimum criteria and the 1 standard error of the minimum criteria (the 1-SE criteria). A λ value of 0.043 was chosen (the 1-SE criteria) according to 10-fold CV, where optimal λ resulted in 14 nonzero coefficients.

TABLE 2 | Diagnostic Performance of the Three Models in Differentiating Focal Nodular Hyperplasia and Atypical Hepatocellular Carcinoma.

	Ultrasonics score	Radiologist's score	Combined
Sensitivity (%)	76.6 (67.5-84.3)	94.4 (88.2-97.9)	93.5 (87.0-97.3)
Specificity (%)	80.5 (70.6-85.9)	69.8 (60.7-77.8)	84.9 (77.1-90.8)
PPV (%)	76.6 (67.5-84.3)	73.7 (65.5-80.9)	84.7 (77.0-90.7)
NPV (%)	79.0 (70.6-85.9)	93.3 (85.9-97.5)	93.5 (87.1-97.3)
+LR	3.7 (3.2-4.2)	3.1 (2.7-3.5)	6.2 (5.6-6.8)
-LR	0.3 (0.2-0.5)	0.1 (0.04-0.2)	0.1 (0.03-0.2)
AUC of training set	0.94 (0.89-0.99)	0.93 (0.85-0.98)	0.99 (0.94-1.00)
AUC of validation set	0.86 (0.80-0.89)	0.84 (0.79-0.89)	0.93 (0.89-0.96)

Data in parentheses are 95% confidence interval. PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; AUC, area under the curve.

The DCA shows that within most reasonable threshold probability ranges, the combined model showed the highest overall net benefit than the radiologist's score or ultrasound feature model. The DCA results for the three models are presented in **Figure 4**.



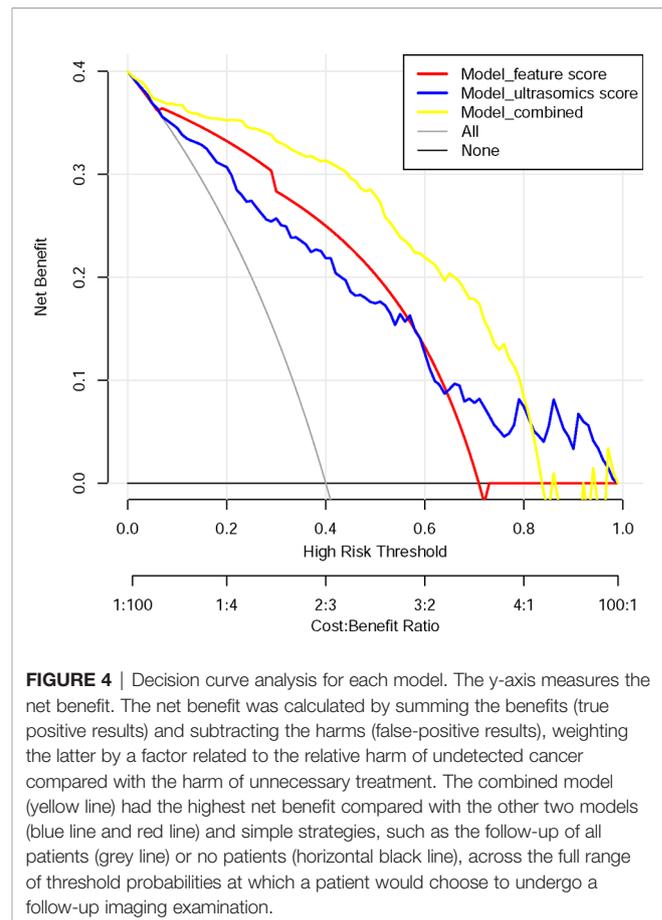
Validation in the Sub-group of HCC With Normal AFP Level

Forty patients were confirmed to develop HCC with normal AFP level. The model trained above performed significantly better (AUC: 0.92, 95% CI: 0.87, 0.96) with the combination of the ultrasonics features and radiologist's score than that with the ultrasonics features (AUC: 0.86, 95% CI: 0.74, 0.89, $P < 0.001$) and radiologist's score (AUC: 0.86, 95% CI: 0.79, 0.91, $P < 0.001$) (**Table 3**).

DISCUSSION

In this study, we derived and validated an ultrasonics-based machine learning approach to analyze ultrasound images for the preoperative individualized diagnosis of FNH and aHCC. Our analysis reveals that the diagnostic performance of ultrasonics is comparable to that of a staff radiologist in differentiating between FNH and aHCC. Furthermore, when adding ultrasonics to the radiologist's classification, the diagnostic performance was improved significantly with an AUC ranging from 0.84 to 0.93 ($P < 0.001$). Our study shows that ultrasonics may increase the diagnostic confidence of radiologists in CEUS examinations and potentially improve their accuracy when facing atypical features. Thus, clinicians would benefit from this decision-making process in the diagnosis of HCC.

In clinical practice, when radiologists face a lesion that shows hyper-enhancement in the arterial phase and sustained enhancement in the portal vein and late phases, it is difficult to diagnose HCC. However, the high sensitivity and negative



predictive values of the radiologist would be useful in clinical practice for excluding disease; thus, HCC would be excluded by the radiologist if the result was considered to be FNH by ultrasonics. Therefore, this system would help to reduce unnecessary biopsies or active clinical treatment requested by experienced radiologists.

In contrast, ultrasonics, referred to as high-throughput computing, extracts innumerable quantitative features from US images (18). By transforming digital medical images into mineable high-dimensional data, ultrasonics yields features, such as textural features, that could objectively reflect the homogeneity or heterogeneity of an image. These patterns could represent enhancement features just as heterogeneity might represent chaotic vessels and necrosis. Focal liver lesions can be featured by typical features in the arterial phase and wash-out during the portal and late phases. In this study, the features displayed in the arterial phase could provide a major benefit for the diagnosis of liver tumors. In previous studies, we utilized a maximum intensity projection technique of micro-flow imaging and achieved higher spatial resolution and higher temporal resolution when detecting vessel contours. Compared to conventional CEUS features (AUC: 0.84), micro-flow imaging technology provided significant improvements over the detection rates achieved for the staff radiologists (AUC: 0.89) (10). In this study, ultrasonics features alone can achieve a similar diagnostic performance (AUC: 0.86) as

TABLE 3 | Validation in the sub-group of HCC with normal AFP value.

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	+LR	-LR	AUC
Ultrasonics score	77.5 (61.5-89.1)	80.5 (72.2-87.2)	60.0 (43.1-75.3)	89.8 (81.5-95.2)	4.4 (3.4-5.8)	0.3 (0.2-0.6)	0.86 (0.74- 0.89)
Radiologist's score	92.5 (79.6-98.3)	77.1 (68.5-84.3)	57.8 (44.7-70.2)	96.8 (90.9-99.3)	4.0 (3.5-4.6)	0.1 (0.03-0.3)	0.86 (0.79-0.91)
Combined	95.0 (83.0-99.2)	82.2 (74.1-88.6)	64.4 (50.9 - 76.4)	98.0 (92.9-99.7)	5.3 (4.8-6.0)	0.06 (0.01-0.2)	0.92 (0.87- 0.96)

Data in parentheses are 95% confidence intervals. PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio; AUC, area under the curve.

micro-flow imaging (AUC: 0.868–0.873). Ultrasonics could reach such achievement because it analyzes textural features objectively and quantitatively to describe the intrinsic characteristics of tumors, in particular heterogeneous tumors. Ultrasonics analysis has already been applied to various types of disease, such as HCC, liver fibrosis, and breast cancer (18, 25–27). The potential of ultrasonics has already been demonstrated for liver imaging in some studies (17, 18).

In this study, we additionally evaluated the benefit of ultrasonics in assisting doctors with the interpretation of medical images. As an interdisciplinary technology, it combines elements of imaging generation, digital image processing, statistical imaging, and knowledge engineering to manage the volume of information related to the diagnostic process and outcome prediction (17, 18, 25). In our study, the additional information in the combined model led to improved diagnostic performance (AUC: 0.93) and higher specificity of 84.9% compared with the ultrasonics (AUC: 0.86, specificity: 80.5%) and radiologist's score (AUC: 0.84, specificity: 69.8%) models. The combined model was also comparable to the ML model based on multi-modal ultrasound images (AUC: 0.94, sensitivity: 91.0%, specificity: 86.0%) (19). The results are also comparable to and even better than those of MRI (AUC: 0.89, sensitivity: 82.2%, specificity: 71.4%), as previously reported (20). However, the use of artificial intelligence is not intended to replace expert diagnosticians because no solution is guaranteed and knowledge-based maintenance is required. Artificial intelligence is also affected by several elements, such as the source of images and the cognition of disease. Presently, most domains of large data have not tapped the full potential of artificial intelligence technology. However, rapid developments in the area will add more potential to the advantages. Therefore, the most important role of artificial intelligence is to help improve diagnostic accuracy and assist rather than replace clinicians in making treatment decisions. It is worth noting that the combined model greatly improved the diagnostic ability of radiologists. A similar conclusion was obtained in another study (28).

Our research has some limitations. First, this study was retrospective and conducted in one center. This may cause potential variations and selection bias in the patient population and imaging methods, which is difficult to generalize the outcomes to other agencies. Second, due to the relative rarity of aHCC, the sample size is relatively small, which may cause over-fitting to this particular population. Hence, large-scale multicenter studies are necessary for the future to validate the results. Third, only two radiologists were involved in the assessment of the basic imaging features and feature extraction. All outcomes were based on the features extracted by one radiologist, which may not be generalizable to all radiologists.

Fourth, the machine and imaging settings in this study were inconsistent, which may affect the ultrasonics features (29, 30).

In conclusion, an ultrasonics approach was developed to investigate the association between the quantitative ultrasound features and pathological characteristics of tumors effectively and objectively. We evaluated the added value of ultrasonics to the radiologist, and this approach improved the performance of CEUS by providing quantitative and standardized criteria to radiologists, thereby enabling the more confident application of CEUS in detecting HCC to achieve better treatment planning. Our findings can assist clinicians in the differential diagnosis between FNH and aHCC accurately using CEUS images, and this allows for early and precise medical management and treatment.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of the First Affiliated Hospital of Sun Yat-Sen University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

WW, M-DL, MK, and X-YX contributed conception and design of the study. XZ collected data from clinical trials. WL and B-WZ drafted and revised the manuscript. X-ZL, WW, L-DC, and S-MR revised the manuscript for important intellectual content. X-ZL, XL, H-TH, YH, and C-QZ gave technical support. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from the National Natural Science Foundation of China (No. 81601500 and 81701701).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2021.544979/full#supplementary-material>

REFERENCES

- Martie A, Sporea I, Popescu A, Sirli R, Danila M, Serban C, et al. Contrast enhanced ultrasound for the characterization of hepatocellular carcinoma. *Med Ultrason* (2011) 13(2):108–13.
- Claudon M, Dietrich CF, Choi BI, Cosgrove DO, Kudo M, Nolsoe CP, et al. Guidelines and good clinical practice recommendations for contrast enhanced ultrasound (CEUS) in the liver—update 2012: a WFUMB-EFSUMB initiative in cooperation with representatives of AFSUMB, AIUM, ASUM, FLAUS and ICUS. *Ultraschall Med* (2013) 34(1):11–29. doi: 10.1055/s-0032-1325499
- Nicolau C, Catala V, Vilana R, Gilabert R, Bianchi L, Sole M, et al. Evaluation of hepatocellular carcinoma using SonoVue, a second generation ultrasound contrast agent: correlation with cellular differentiation. *Eur Radiol* (2004) 14(6):1092–9. doi: 10.1007/s00330-004-2298-0
- Jang HJ, Kim TK, Burns PN, Wilson SR. Enhancement patterns of hepatocellular carcinoma at contrast-enhanced US: comparison with histologic differentiation. *Radiology* (2007) 244(3):898–906. doi: 10.1148/radiol.2443061520
- Dietrich CF. Contrast-Enhanced Ultrasound of Benign Focal Liver Lesions. *Ultraschall der Med* (2018) 40(1):12–29. doi: 10.1055/a-0668-5746
- Friedrich-Rust M, Klopffleisch T, Nierhoff J, Herrmann E, Vermehren J, Schneider MD, et al. Contrast-Enhanced Ultrasound for the differentiation of benign and malignant focal liver lesions: a meta-analysis. *Liver Int* (2013) 33(5):739–55. doi: 10.1111/liv.12115
- Wang W, Chen LD, Lu MD, Liu GJ, Shen SL, Xu ZF, et al. Contrast-enhanced ultrasound features of histologically proven focal nodular hyperplasia: diagnostic performance compared with contrast-enhanced CT. *Eur Radiol* (2013) 23(9):2546–54. doi: 10.1007/s00330-013-2849-3
- Yen YH, Wang JH, Lu SN, Chen TY, Changchien CS, Chen CH, et al. Contrast-enhanced ultrasonographic spoke-wheel sign in hepatic focal nodular hyperplasia. *Eur J Radiol* (2006) 60(3):439–44. doi: 10.1016/j.ejrad.2006.06.007
- Xu HX, Liu GJ, Lu MD, Xie XY, Xu ZF, Zheng YL, et al. Characterization of focal liver lesions using contrast-enhanced sonography with a low mechanical index mode and a sulfur hexafluoride-filled microbubble contrast agent. *J Clin Ultrasound* (2006) 34(6):261–72. doi: 10.1002/jcu.20234
- Li W, Wang W, Liu GJ, Chen LD, Wang Z, Huang Y, et al. Differentiation of Atypical Hepatocellular Carcinoma from Focal Nodular Hyperplasia: Diagnostic Performance of Contrast-enhanced US and Microflow Imaging. *Radiology* (2015) 275(3):870–9. doi: 10.1148/radiol.14140911
- Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
- Sun R, Limkin EJ, Vakilopoulou M, Dercle L, Champiat S, Han SR, et al. A radiomics approach to assess tumor-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* (2018) 19(9):1180–91. doi: 10.1016/S1470-2045(18)30413-3
- Banerjee S, Wang DS, Kim HJ, Sirlin CB, Chan MG, Korn RL, et al. A computed tomography radiogenomic biomarker predicts microvascular invasion and clinical outcomes in hepatocellular carcinoma. *Hepatology (Baltimore Md)* (2015) 62(3):792–800. doi: 10.1002/hep.27877
- Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuze S, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol* (2017) 28(6):1191–206. doi: 10.1093/annonc/mdx034
- Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, et al. Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer. *Radiology* (2016) 281(3):947–57. doi: 10.1148/radiol.2016152234
- Prasanna P, Patel J, Partovi S, Madabhushi A, Tiwari P. Radiomic features from the peritumoral brain parenchyma on treatment-naïve multi-parametric MR imaging predict long versus short-term survival in glioblastoma multiforme: Preliminary findings. *Eur Radiol* (2017) 27(10):4188–97. doi: 10.1007/s00330-016-4637-3
- Wang K, Lu X, Zhou H, Gao Y, Zheng J, Tong M, et al. Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. *Gut* (2018) 68(4):729–41. doi: 10.1136/gutjnl-2018-316204
- Li W, Huang Y, Zhuang BW, Liu GJ, Hu HT, Li X, et al. Multiparametric ultrasonics of significant liver fibrosis: A machine learning-based analysis. *Eur Radiol* (2018) 29(3):1496–506. doi: 10.1007/s00330-018-5680-z
- Yao Z, Dong Y, Wu G, Zhang Q, Yang D, Yu JH, et al. Preoperative diagnosis and prediction of hepatocellular carcinoma: Radiomics analysis based on multi-modal ultrasound images. *BMC Cancer* (2018) 18(1):1089. doi: 10.1186/s12885-018-5003-4
- Wu J, Liu A, Cui J, Chen A, Song Q, Xie L. Radiomics-based classification of hepatocellular carcinoma and hepatic haemangioma on precontrast magnetic resonance images. *BMC Med Imaging* (2019) 19(1):23. doi: 10.1186/s12880-019-0321-9
- Gatos I, Tsantis S, Karamesini M, Spiliopoulos S, Karnabatidis D, Hazle JD, et al. Focal liver lesions segmentation and classification in nonenhanced T2-weighted MRI. *Med Phys* (2017) 44(7):3695–705. doi: 10.1002/mp.12291
- Gatos I, Tsantis S, Spiliopoulos S, Skouroliaou A, Theotokas I, Zoumpoulis P, et al. A new automated quantification algorithm for the detection and evaluation of focal liver lesions with contrast-enhanced ultrasound. *Med Phys* (2015) 42(7):3948–59. doi: 10.1118/1.4921753
- Acharya UR, Koh JE, Hagiwara Y, Tan JH, Gertych A, Vijayanathan A, et al. Automated diagnosis of focal liver lesions using bidirectional empirical mode decomposition features. *Comput Biol Med* (2018) 94:11–8. doi: 10.1016/j.combiomed.2017.12.024
- Helmreich JE. *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. 2nd edn. Switzerland: Springer International Publishing (2015). doi: 10.1007/978-3-319-19425-7
- Hu HT, Wang Z, Huang XW, Chen SL, Zheng X, Ruan SM, et al. Ultrasound-based radiomics score: a potential biomarker for the prediction of microvascular invasion in hepatocellular carcinoma. *Eur Radiol* (2019) 29(6):2890–901. doi: 10.1007/s00330-018-5797-0
- Chen LD, Ruan SM, Liang JY, Yang Z, Shen SL, Huang Y, et al. Differentiation of intrahepatic cholangiocarcinoma from hepatocellular carcinoma in high-risk patients: A predictive model using contrast-enhanced ultrasound. *World J Gastroenterol* (2018) 24(33):3786–98. doi: 10.3748/wjg.v24.i33.3786
- Guo B, Ouyang F, Ouyang L, Liu Z, Lin S, Meng W, et al. Development and validation of an ultrasound-based nomogram to improve the diagnostic accuracy for malignant thyroid nodules. *Eur Radiol* (2018) 29(3):1518–26. doi: 10.1007/s00330-018-5715-5
- Wang J, Wu CJ, Bao ML, Zhang J, Wang XN, Zhang YD. Machine learning-based analysis of MR radiomics can help to improve the diagnostic performance of PI-RADS v2 in clinically relevant prostate cancer. *Eur Radiol* (2017) 27(10):4082–90. doi: 10.1007/s00330-017-4800-5
- Yu X, Guo Y, Huang SM, Li ML, Lee WN. Beamforming effects on generalized Nakagami imaging. *Phys Med Biol* (2015) 60(19):7513–31. doi: 10.1088/0031-9155/60/19/7513
- Byra M, Wan L, Wong JH, Du J, Shah SB, Andre MP, et al. Quantitative Ultrasound and B-Mode Image Texture Features Correlate with Collagen and Myelin Content in Human Ulnar Nerve Fascicles. *Ultrasound Med Biol* (2019) 45(7):1830–40. doi: 10.1016/j.ultrasmedbio.2019.02.019

Conflict of Interest: XL was employed by GE Healthcare.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Li, Lv, Zheng, Ruan, Hu, Chen, Huang, Li, Zhang, Xie, Kuang, Lu, Zhuang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.