IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 21, 2021, accepted March 4, 2021, date of publication March 9, 2021, date of current version June 29, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3064879

Diagnosis of Significant Liver Fibrosis by Using a DCNN Model With Fusion of Features From US B-Mode Image and Nakagami Parametric Map: An Animal Study

QIANG LIU, ZHONG LIU, WENCONG XU, HUIYING WEN, MING DAI, AND XIN CHEN, (Member, IEEE)

School of Biomedical Engineering, Shenzhen University, Shenzhen 518055, China

Corresponding authors: Zhong Liu (liuzhong@szu.edu.cn) and Xin Chen (chenxin@szu.edu.cn)

This work was supported by the National Science Foundation of China under Grant 81871429 and Grant 61901282.

ABSTRACT The assessment of liver fibrosis is usually required in the diagnostic and treatment procedures for chronic liver disease. In this paper, we propose a deep convolutional neural network (DCNN) with multi-feature fusion (DCNN-MFF) to extract and integrate features from ultrasound (US) B-mode image and Nakagami parametric map for significant liver fibrosis recognition. The DCNN-MFF model mainly consists of three branches, among which the first two branches are used for extracting deep features respectively from the US B-mode image and the Nakagami parametric map, while the remaining branch is designed for extracting quantitative US features from the Nakagami parametric map. At the backend of the DCNN model, the extracted deep and quantitative features are fused together for final decision making. The performance of the DCNN-MFF model was evaluated on an animal dataset collected from 84 rats with 168 liver lobes under various fibrosis stages. Across five-fold cross-validation, the accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC) achieved by the proposed DCNN-MFF model for significant liver fibrosis recognition were respectively 0.827 (95% confidence interval [CI] 0.762-0.881), 0.821 (95% CI 0.729-0.892), 0.836 (95% CI 0.731-0.912), 0.869 (95% CI 0.809-0.916), which are significantly better than those provided by the comparative methods.

INDEX TERMS Liver fibrosis, ultrasound, deep learning, Nakagami imaging, feature fusion.

I. INTRODUCTION

Chronic liver disease (CLDs) is a significant cause of morbidity and mortality on a global scale [1]. The excessive accumulation of extracellular matrix proteins caused by CLDs, such as virus hepatitis, non-alcoholic fatty liver disease, drug-induced liver injury, and immunological liver injury [2], [3], may result in different levels of liver fibrosis. According to the METAVIR scoring system [4], [5], liver fibrosis can be classified into five stages: S0, no fibrosis; S1, minimal fibrosis; S2, moderate fibrosis; S3, severe fibrosis; and S4, cirrhosis. For better managements of patients with CLD, a precise and periodical estimation of the liver fibrosis stage is important for establishing the optimal time and strategy of therapy as well as for predicting the response of treatment. In particular, the identification of significant liver fibrosis, which is defined as fibrosis stages $\geq$S2 [6]–[9], is of significant clinical importance as it signals the necessity for antifibrotic treatments [6], [10] for patients with CLD. Currently, histopathological examination based on liver biopsy is still regarded as the gold standard for liver fibrosis staging. However, liver biopsy is invasive, and may be subject to the sampling error and cause other complications [11], [12]. Moreover, it is impractical to perform liver biopsy repeatedly for monitoring the progress of liver fibrosis state.

In clinical practice, for repeatable examinations of liver fibrosis, conventional reading on ultrasound (US) B-mode images is usually performed to inspect the changes in liver echogenicity and nodularity as well as signs of portal hypertension. However, the diagnosis based on visual inspection of US B-mode images is subjective and empirical. The

The associate editor coordinating the review of this manuscript and approving it for publication was Pengcheng Liu.

misdiagnosis from inexperienced radiologists may lead to an inaccurate treatment to the liver fibrosis. As an alternative to conventional reading, computer aided diagnosis (CAD) systems have been developed for more convenient and objective assessments of liver fibrosis. Conventional CAD systems usually extract handcrafted features from US B-mode images and rely on a traditional classifier for fibrosis classification. For instance, Mojsilovic *et al.* [13] developed a filter bank method to extract wavelet features from liver B-mode images, after which a Bayes classifier was used for liver cirrhosis classification. Yeh *et al.* [14] collected image features based on gray level concurrence and nonseparable wavelet transform and used support vector machine as the classifier for liver fibrosis classification. Wu *et al.* [15] assessed US liver images by making use of a variety of texture features, including the spatial gray-level dependence matrices, the Fourier power spectrum, the gray-level difference statistics, and the Laws texture energy measures. They used Bayes classifier and the Hotelling trace criterion to evaluate the performance of the handcrafted features. In recent years, studies on deep learning, particularly deep convolutional neural network (DCNN), have mushroomed. This can be evidenced by the explosive growth of DCNN-based applications including but not limited to image classification [16], object detection and recognition [17], and natural language processing [18]. In contrast to traditional classification methods, DCNN automates the feature extraction procedure, and the features extracted by a DCNN model is usually in a larger amount. Due to the powerful feature learning ability, there is a trend to apply DCNN methods for solving more specialized problems, such as video surveillance [19], [20], remote sensing image processing [21], [22] and medical image analysis [23]. Particularly, the application of DCNN models for liver fibrosis classification in US B-mode images has been also investigated. For examples: Meng *et al.* [24] concatenated a VGGNet [25] and a fully connected network (FCNet), with the former for deep feature extraction and the latter for classification, for stratifying liver fibrosis in US B-mode images based on transfer learning. Mitani *et al.* [26] investigated the effect of image augmentation on a DCNN in classifying a cirrhosis liver on US B-mode images. Lee *et al.* [27] tested a set of wellknown DCNN models, such as AlexNet [28], GoogleNet [29], and VGGNet, and trained them on liver US B-mode images by leveraging the transfer learning technique. They found that VGGNet had better performance than the other DCNN models for liver fibrosis classification. The above-mentioned intelligent algorithms are developed based on B-mode images, which are constructed from post-processed US radiofrequency (RF) signals, i.e. US echoed waves. The B-mode imaging process involves a series of technical operations including filtering of the RF signals, envelope detection, log compression, and gain/dynamic range adjustments. These imaging operations may lead to information loss or distortion of the raw RF signals, which to some extent restricts the performance of B-mode image-based algorithms.

Rather than using B-mode images, quantitative US (QUS) [30] performs analysis based on raw RF signals. Since the RF signals retain intact frequency information of echoed US waves, better diagnostic results can be expected by exploring the information carried by RF signals for analysis. To date, several QUS techniques with potentials for assessing liver fibrosis have been developed, which can be classified into two main categories: envelope statistics and spectral-based approaches. The approaches based on envelope statistics utilize a particular model to measure the distribution of RF echo envelope, such that the underlying information of tissue microstructures can be explored. Various models, such as Nakagami model [31], Rayleigh model and homodyned-K model [32], have been studied within a spectrum of biomedical US applications [33]. On the other hand, spectral-based approaches are built upon attenuation parameters or backscatter coefficients. The former includes the controlled attenuation parameter [34] and attenuation coefficient slope (ACS) [35], while the latter includes integrate backscatter coefficient (IBC) [36], [37], the effective scatterer diameter (ESD) and the effective acoustic concentration (EAC) of scatterers [38]. The spectral-based approaches also have potentials to uncover tissue microstructures, benefitting from the analysis of amplitude, frequency and phase of RF signals. Among the aforementioned QUS methods, the Nakagami model represented by a shape parameter m and scaling parameter Ω is most frequently used for characterizing liver fibrosis [31]. In particular, images can be constructed based on the estimation of Nakagami-m parameter, which can be used for liver fibrosis analysis, such as the works in [39]–[41]. Compared with B-mode image-based algorithms, QUS techniques have the advantage to describe intrinsic acoustic characteristics associated with tissue microstructures [42], [43]. However, the number of feature variables available for QUS development is typically small (usually less than 10). We hypothesize that the performance could be further improved by increasing the number of feature variables.[1] As mentioned, Nakagami-m parameter can be estimated from the RF signals for liver fibrosis assessment, and more importantly, the parameter can be imaged for more convenient analysis. This suggests that, given the raw RF signals, a Nakagami-m parametric map can be generated parallel to the construction of B-mode images, as shown in Fig. 1. In light of this, it would be natural to apply the deep learning methods to extract and fuse features from both the Nakagami-m parametric map and B-mode image for better assessments of liver fibrosis.

In this paper, we propose a DCNN model with multi-feature fusion (DCNN-MFF) to extract and integrate features from US B-mode image and Nakagami parametric map for recognizing significant liver fibrosis. The proposed DCNN-MFF model mainly consists of three branches, among which the first two branches are used for extracting

[1]The proposed hypothesis has been validated by the ablation test results presented in Section III-A.

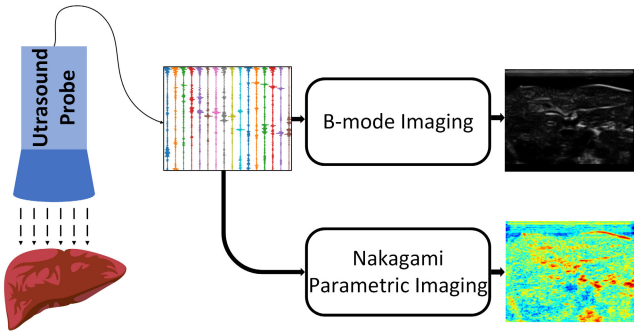


FIGURE 1. Graphical illustration of B-mode and Nakagami parametric imaging.

deep features respectively from the US B-mode image and the Nakagami parametric map, while the remaining branch is designed to extract quantitative US features from the Nakagami parametric map. At the backend of the DCNN model, the extracted deep and quantitative features are fused to a single vector for final decision making of the liver fibrosis state (whether significant or not). Experimental results on an animal dataset show the effectiveness and efficacy of the proposed DCNN-MFF model for significant liver fibrosis recognition.

The main contributions of this paper are as follows:

1) A new deep convolutional neural network (DCNN) is devised to extract and combine features from US B-mode image and Nakagami parametric map for significant liver fibrosis recognition. Particularly, in the developed DCNN model both deep and handcrafted features are extracted from the Nakagami parametric map, which are integrated with the deep features stemming from the US B-mode image for fibrosis classification.
2) The added values of deep and handcrafted features extracted from the Nakagami parametric map to the deep features of US B-mode images for recognizing significant liver fibrosis are validated on a rat model cultured by injecting different doses of dimethylnitrosamine into the rats for forming livers with various fibrosis stages.

The rest of the paper is organized as follows. In section II describes the data acquisition procedures and illustrates the specific details of the proposed DCNN-MFF model. Experimental results are given in Section III, followed by discussions in Section IV. Finally, Section V concludes the paper.

II. MATERIAL AND METHODS

A. DATA ACQUISITION AND PROCESSING

1) ANIMALS

The animal study was approved by the Experimental Animal Ethics Committee of Sun Yat-sen University. In our study, ninety male clean Sprague-Dawley (SD) rats weighing between 220g to 250g were randomly divided into seven groups:10 for each of groups 1 to 3, 15 for groups 4 to 7. All the rats were fed with the same formula and cultured in a sterile and constant environment, where the temperature and humidity were maintained at 20°C to 60°C and 40% to 70%. Particularly, group 1 was without any drug injection, serving as a control group. Groups 2, 3, 4, 5, 6, and 7 were injected intraperitoneally with dimethylnitrosamine (DMN) at a dose of $1\mu L$ (diluted 1: 100 with 0.15 M sterile NaCl) per 100g body weight. The drug injection period for each group were respectively 1, 2, 3, 4, 5 and 6 weeks (for each week, the drug was injected in the first 3 consecutive days). Two weeks after the completion of drug injections for each group, DMN-induced liver fibrosis could be formed and vivo US examinations on the rat livers were conducted to collect US images and RF signals. Immediately after the US examination, the rats in the group were sacrificed, and the left and right liver lobes were excised for separate histopathologic examination.

2) EXAMINATIONS

As mentioned, in our study each rat underwent a US examination *in vivo*, followed by a histopathological examination. Before US examination, the rats were fasted and anesthetized with 1.5% Pentobarbital Sodium 30mg/kg. Then, a Sonix TOUCH US system (Richmond, Canada) equipped with a linear array probe (L14-5/38) was used to capture the US images of the rat liver as well as the corresponding RF signals before US imaging. The central frequency of the probe is 6 MHz and the sampling frequency is 40 MHz. Immediately after the US exam, the examined rats were sacrificed and the left and right liver lobes were excised for separated histopathologic examination. The sampled liver tissues were taken and fixed in 10% formalin solution for at least 24 hours. After washing and dehydration, the samples were embedded in paraffin wax and sections to a thickness of 5 μm. The paraffin slices were stained with hematoxylin-Eosin and Masson and then analyzed under a microscope (BX41, Olympus, Pittsburgh, PA). Finally, according to the METAVIR scoring system, the severity of fibrosis in the rat livers was classified into 5 levels from S0 to S4 as follows: S0 = no fibrosis, S1 = enlarged portal fibrosis, S2 = periportal fibrosis ± periportal septa, S3 = architectural distortion but no obvious cirrhosis and S4 = cirrhosis (probable or definite). Details on the number of rats and liver samples under various fibrosis stages are presented in Table 1.

3) NAKAGAMI-M PARAMETRIC IMAGING

The Nakagami distribution was originally proposed to model radar echoes [44]. The Nakagami distribution also fits well to the probability density function (pdf) of the ultrasonic envelope. Mathematically, the pdf of the ultrasonic backscattered envelope R can be represented by

$$f(R) = \frac{2m^m R^{2m-1}}{\Gamma(m)\Omega^m} \exp(-\frac{m}{\Omega}R^2)U(R), \quad (1)$$

where $\Gamma(\cdot)$, $U(\cdot)$ are respectively the gamma function and the unit step function, m and Ω represent the Nakagami shape parameter and scaling parameter, respectively. The Nakagami

TABLE 1. Number of rats, liver samples and histopathologic results.

Groups†	Number of rats	Number of rat liver lobe samples‡	Fibrosis stage determined via histopathological analysis				
			S0	S1	S2	S3	S4
Group 1 (0/2)*	10	20	11	9			
Group 2 (1/3)	10	20	4	16			
Group 3 (2/4)	10	20		14	6		
Group 4 (3/5)	15	30		11	14	5	
Group 5 (4/6)	15	30		6	18	6	
Group 6 (5/7)	15	22		2	6	12	2
Group 7 (6/8)	15	26			5	11	10
Total	90	168	15	58	49	34	12

Note−†Data in parenthesis are number of weeks for dimethylnitrosamine (DMN) injection / number of weeks before sacrifice of the rats. ‡4 and 2 rats, respectively in groups 6 and 7, died in their rest weeks. *some rat livers in the control group were diagnosed to be at stage S1, which may be caused by non-specific hepatic inflammatory changes of livers for rats in the clean grade.

parameters m and Ω can be calculated according to the following formulas:

$$\Omega = E(R^2), \tag{2}$$

$$m = \frac{[E(R^2)]}{E[R^2 - E(R^2)]}, \tag{3}$$

where $E(\cdot)$ denotes the operation of expectation.

Extensive studies [45]–[48] have shown that Nakagami-m parameter is able to uncover tissue characteristics through the quantification of scatterer concentrations in the tissue. Thus, in our work, parameter imaging was performed to obtain the Nakagami parametric map for assessing liver fibrosis. The imaging procedure goes as follows. For each RF data point, a squared window centered at this point was used to estimate the Nakagami parameter m. First, the envelope for the RF data within the squared window was calculated. Then, by modelling the pdf of the envelope with Nakagami distribution denoted by Eq. (1), a Nakagami-m parameter can be estimated in terms of Eq. (3). Thereafter, the pixel in the Nakagami parametric map, with the same coordinates to the RF data point at the window center, was assigned to be the value of the estimated Nakagami-m parameter. As such, a Nakagami parametric map with the same size of the US B-mode image could be obtained by sliding the window horizontally and vertically in a step of one data point. Note, the size of the sliding window is a key factor to control the imaging quality of the Nakagami parametric map. The use of a larger sliding window yields more stable estimations of the m parameter, but the corresponding image resolution of the Nakagami parametric map is lower. In our study, the side length of the sliding window was empirically set to be a value of ten times the pulse length of the US wave.

4) ROI SEGMENTATION

To limit our focus on the assessment of fusing features from the US image and parametric map for fibrosis classification, we developed a specialized software with graphical user interface to allow users manually delineate on the US image for identifying the region of interest (liver parenchymal area). Particularly, in order to get the region of interest (ROI) of RF signals, we first reconstructed the US images from the RF signals by following the conventional B-mode imaging procedures including envelope detection and log compression. Then, the reconstructed images were viewed by a skilled sonographer to manually delineate the areas of the left and right liver lobes, which were then mapped back to the corresponding RF signals for identifying the ROIs. Since the degrees of liver fibrosis in the left and right liver of the rats may be different, each liver lobe is regarded as an independent data sample. Thus, we finally obtained 168 data samples, each of which contains ROI of US B-mode image as well as the corresponding RF signals.

B. THE PROPOSED DCNN-MFF MODEL

As shown in Fig. 2, the proposed DCNN model with multi-feature fusion (DCNN-MFF) mainly consists of three branches from Branch-1 to Branch-3. Branch-1 and Branch-2 are used for extracting deep features respectively from the segmented US B-mode image and Nakagami parametric map, while Branch-3 is dedicated to collecting handcrafted quantitative features from the Nakagami parametric map. At the backend of the model, the deep features derived from Branch-1 and Branch-2 are integrated to single feature vector, which is then fused with the handcrafted quantitative features from Branch-3 to output a 512-dimensional feature vector for final differential diagnosis between significant liver fibrosis ($\geq$S2) and non-significant liver fibrosis ($<$S2). In what follows of this subsection, we shall describe in depth the various components of the DCNN-MFF model.

1) BASELINE NETWORK

In Branch-1 and Branch-2 of the DCNN-MFF model, deep features are extracted by using a baseline network modified from VGG16 [25]. As shown in Fig. 2, the baseline network contains 13 convolutional layers (CLs, indicated by red arrows) which can be grouped into five convolutional blocks (CBs): each of the first two CBs has 2 CLs, while each of the rest CBs has 3 CLs. The 13 CLs are directly borrowed from VGG16, as the stacking of multiple CLs improves the ability of the DCNN model to extract complex high-level features. Right after the 13 CLs, global average pooling(GAP) [49] is used to integrate the feature maps of the last CLs into a 512-dimensional feature vector. This is different from the original VGG16, which uses fully connected (FC) layers for

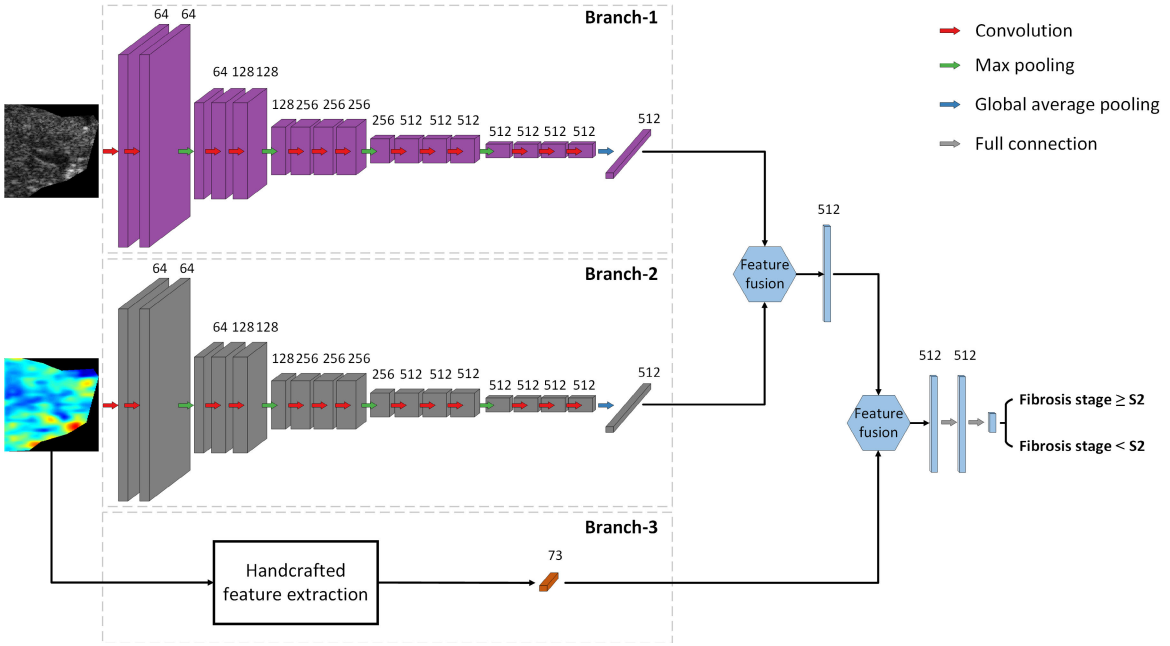


FIGURE 2. Overview of the proposed DCNN model. Branch-1 and Branch-2 are used for feature extraction from US B-mode image and Nakagami parametric map, respectively, and Branch-3 is used for manual feature extraction from the Nakagami parametric map.

feature vector extraction. Such a modification aims to reduce the number of parameters in the model, since a model with the larger number of parameters is prone to overfit, especially when there are no sufficient data to be used for model training. Compared with the FC layer with a large number of parameters, GAP has no parameters to be learned, and thus replacing the FC layers with GAP can alleviate the overfitting problem to some extent. The baseline network is used by Branch-1 and Branch-2 to extract deep features respectively from US B-mode image and the Nakagami parametric map.

2) HANDCRAFTED QUANTITATIVE FEATURES

In order to enhance the performance of the DCNN model, a set of handcrafted quantitative features from the Nakagami parametric map (Branch-3) were extracted to assist the deep learning network for decision making. Specifically, by following the feature definitions in Radiomics [50], a total of 73 handcrafted features were extracted from the Nakagami parametric map, including 14 statistics features (mean, median, maximum, minimum, range, mean absolute deviation, root mean square, variance, standard deviation, skewness, kurtosis, uniformity, energy, entropy) and 59 texture features derived from the Gray-Level Co-occurrence Matrix (GLCM), Gray-Level Run-Length Matrix (GLRLM), Gray-Level Size Zone Matrix (GLSZM), Neighborhood Gray-Tone Difference Matrix (NGTDM). These handcrafted quantitative features extracted in Branch-3 is further fused with the integrated deep features of Branch-1 and Branch-2 for fibrosis classification.

3) FEATURE FUSION

As shown in Fig. 2, there two times of feature fusion in the proposed DCNN-MFF model. The first one is the fusion of the deep features from US B-mode image (Branch-1) and the deep features from the Nakagami parametric map (Branch-2), while the other one is the fusion of handcrafted quantitative features (Branch-3) and the integrated deep features of Branch-1 and Branch-2. In general, concatenation is a straightforward method for feature fusion. However the concatenation method only uses the first-order information of the feature, but fails to take advantage of the correlation between different features. To seek for a better representation capacity, we employed the multi-modal low-rank bilinear pooling approach [51] for feature fusion. Specifically, given two different feature vectors, $\boldsymbol{x} \in \mathbb{R}^m$ and $\boldsymbol{y} \in \mathbb{R}^n$, a multi-modal bilinear model can be formulated to output a o-dimensional feature vector $\boldsymbol{z} = \left(\boldsymbol{U}^T\boldsymbol{x}\right) \odot \left(\boldsymbol{V}^T\boldsymbol{y}\right)$, where $\boldsymbol{U} \in \mathbb{R}^{m\times o}$ and $\boldsymbol{V} \in \mathbb{R}^{n\times o}$ are two factorized tensors to be learned, $\odot$ denotes the element-wise multiplication of two vectors. The approach can be easily implemented by combining some commonly-used layers such as fully connection, element-wise multiplication, dropout, and normalization [52]. as summarized in Fig. 3. After feature fusion, the fused features are further processed by two FC layers (gray arrows indicated in Fig. 2) to output the final recognition result. The first FC layer has 512 neurons activated by rectified linear unit (ReLU) [53], while the last FC layer has 2 neurons (corresponding to two class labels: 0, non-significant liver fibrosis; 1, significant liver fibrosis) activated by the softmax function to output the final class probabilities.

4) MODEL PARAMETERS

The configuration of the proposed DCNN model is summarized in Table 2. As mentioned earlier, the 13 CLs of the baseline model for Branch-1/2 are borrowed directly from the VGG16 network. We suppose that better classification

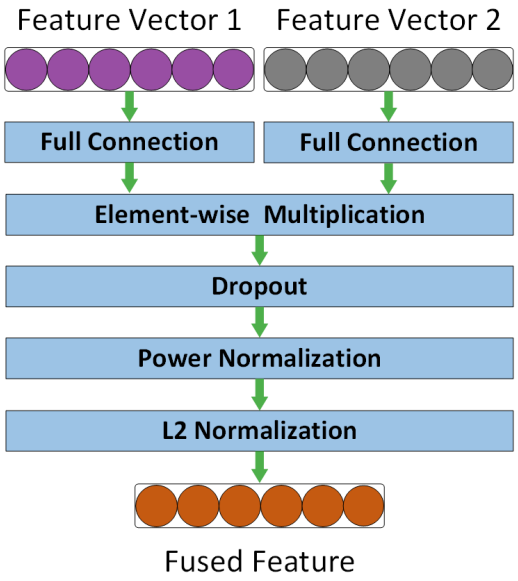


FIGURE 3. Flowchart of multi-modal factorized bilinear pooling for feature fusion.

results might be obtained by properly adjusting the model parameters of the baseline model, but the hypothesis should be evaluated by conducting additional relevant experiments. Since the aim of our study is mainly to examine the effectiveness of fusion of features drawn from various US image resources, the original setting of the 13 CLs in VGG16 is maintained for the development of the baseline model. Right after the 13 CLs, a GAP unit is used to integrate the feature maps of the last CLs in Branch-1/2 into a 512-dimensional feature vector. As previously introduced, GAP is different from FC layers in VGG16 and it has no parameters need to be learned. Therefore, the baseline model with GAP module is less likely to be overfit than the VGG16 network, as the number of model parameters has been reduced. The feature fusion module consists of two FC layers (the size of each is 512, corresponding to the size of feature vector outputted from Branch-1/2), an element-wise multiplication layer connected by a dropout layer with a dropout ratio of 0.5, and two regularization layers respectively for power and L2 normalization. As shown in Fig. 2 and Table 2, the feature fusion module is used to integrate the deep features extracted from both the US B-mode image and the Nakagami map as well as the 73 handcrafted quantitative features extracted from the latter. The fused features are further processed by two FC layers to gradually yield the final class probabilities: the first FC layer has 512 neurons and the last FC layer has 2 neurons corresponding to the binary classification results studied in the paper.

C. TRAINING STRATEGY

Since the number of data samples in our dataset is typically small ($n = 168$), both data augmentation and transfer learning were used for better training of the DCNN model. The training procedure is detailed as follows. First, the US B-mode images and Nakagami parametric maps in the training subset

TABLE 2. Configuration of the DCNN-MFF model.

Branch-1/2			Branch-3
Type	**Filters**	**Size**	
Conv2D	64	3×3	
Conv2D	64	3×3	
Max Pooling		2×2	
Conv2D	128	3×3	
Conv2D	128	3×3	
Max Pooling		2×2	
Conv2D	256	3×3	
Conv2D	256	3×3	73
Conv2D	256	3×3	Handcrafted
Max Pooling		2×2	Features
Conv2D	512	3×3	
Conv2D	512	3×3	
Conv2D	512	3×3	
Max Pooling		3×3	
Conv2D	512	3×3	
Conv2D	512	3×3	
Conv2D	512	3×3	
Global Average Pooling			
	Type	**Size**	
	Full Connection	512	
	Full Connection	512	
Feature Fusion	Element-wise Multiplication		
Module	Dropout		
	Power Normalization		
	L2 Normalization		
Type		**Size**	
Feature Fusion Module			
Full Connection		512	
Full Connection		2	

were flipped horizontally, or rotated between $-10°$ and $10°$ in a random manner, such that the number of training samples can be substantially enlarged. Thereafter, transfer learning was adopted to train the DCNN-MFF model on the enlarged dataset. To be more specific, we firstly assembled a trainable auxiliary network by connecting the baseline model described in subsection-1 with a fully connected layer containing 2 neurons (corresponding to two class labels: 0, non-significant liver fibrosis; 1, significant liver fibrosis). The parameters of the FC layer were initialized by using the Xaiver method [54], while the remaining layers were initialized by transferring the parameters of the VGG16 model trained on ImageNet dataset [55]. Then, the auxiliary network was fine-tuned on the augmented US B-mode images, and the resulted model was used to initialize Branch-1 of the DCNN-MFF model. Similarly, Branch-2 can be initialized by tuning an auxiliary network on the augmented Nakagami parametric maps. Further, by initializing the remaining components (fusion modules and the last two FC layers) with the Xaiver method, the whole DCNN-MFF model could be trained in a layer-wise manner, starting from training only the Xaiver-initialized layers followed by tuning all the layers. The proposed model as well as the auxiliary network was trained by minimizing the following cross-entropy loss function.

$$\mathcal{L} = \frac{1}{N}\sum_{i} -[y_i log(p_i) + (1 - y_i) log(1 - p_i)], \qquad (4)$$

where y_i denotes the true label of i-th rat liver sample, with '0' representing non-significant liver fibrosis and '1' representing significant liver fibrosis, p_i indicates the probability that the i-th sample is predicted to be with label '1'.

D. PERFORMANCE EVALUATION

Five-fold cross-validation was used for performance evaluation. Specifically, we randomly divided the rat liver samples into five portions, each of which has approximately equal number of data samples. Of the five portions, three were used for training, while the remaining two portions were used for validation and testing respectively. Finally, the classification performance of the proposed DCNN model was measured in terms of accuracy (ACC), sensitivity (SEN), specificity (SPE), the receiver operating characteristic (ROC) curve and the area under ROC curve (AUC). The above quantities for performance evaluation were obtained as follows. First, the mean ROC curve was obtained by merging and sorting the scored instances [56] resulted from the cross-validations. Thereafter, the ACC, SEN and SPE were calculated by identifying the optimal cut-point on the ROC curve [57]. The ACC, SEN and SPE are defined as follows:

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \tag{5}$$

$$\mathrm{SEN} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{6}$$

$$\mathrm{SPE} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \tag{7}$$

where, TP (true positive) and TN (true negative) denote respectively the numbers of positive and negative samples that are correctly classified, while FP (false positive), FN (false negative) represent the numbers of negative and positive samples that are incorrectly recognized, respectively. As mentioned, in our study, positive samples are rat livers with fibrosis stage $\geq$S2, while negative samples are rat livers under fibrosis stage $<$S2. The AUC values lie in the range between 0 and 1, representing the probability of correct classifications. Therefore, a good classification model has an AUC value close to 1, and the model with AUC value below 0.5 provides inferior performance worse than random guessing. For statistical analysis, 95% confidence intervals (CIs) of the accuracy, sensitivity and specificity were calculated with the Clopper-Pearson method [58]. In addition, we used the DeLong test [59] to compare the averaged ROC curves.

III. EXPERIMENTAL RESULTS

For each cross-validation, the trained model with the highest AUC value monitored on the validation dataset was selected as the one for model evaluation on the testing fold. The performances of the proposed model on the training, validation and testing datasets are summarized in Table 3, where the results for the validation and testing datasets were obtained by using the performance evaluation method described in Section II-D. For the training dataset, the ACC, SEN, SPE and AUC achieved by the proposed model are all 1.0. For the validation dataset, the ACC, SEN, SPE and AUC values are respectively 0.875 (95% CI 0.815-0.921), 0.863 (95% CI 0.777-0.925), 0.890 (95% CI 0.795-0.952), and 0.929 (95% CI 0.879-0.963), while for the testing dataset, the corresponding values are 0.827 (95% CI 0.762-0.881), 0.821 (95% CI 0.729-0.892), 0.836 (95% CI 0.731-0.912), and 0.869 (95% CI 0.809-0.916), respectively. Next, we shall present in turn the evaluations of various branches of the DCNN-MFF model, and the comparisons against other commonly used DCNN models.

A. ABLATION TEST

To show the contribution of each branch in the DCNN-MFF model, we conducted an ablation test in which the performance of the model was evaluated by trimming the branches alternatively. Specifically, we adjusted the DCNN-MFF model in Fig. 2 by removing the branches in an order as follows: Branch-3, Branch-3 and Branch-1 (Branch-3/1), and Branch-3/2. For the removal of Branch-3, the last two FC layers at the backend of the DCNN-MFF model were connected directly to the fused deep feature vector of Branch-1 and Branch-2 for fibrosis classification, while for the removal of Branch-3/1 and Branch-3/2, the last two FC layers should be linked directly to the remaining branch. Thereafter, the ablated models were fine-tuned on the training data and tested on the testing data for comparisons with the original model. The results are summarized in Table 4. It can be seen the AUC value decreases from 0.869 (95% CI 0.809-0.916) to 0.827 (95% CI 0.762-0.881) along with the removal of branches in the defined order. Compared with the ablated model using only deep features from US B-mode image (ablation of Branch-3/2, bottom row of Table 4), the original DCNN-MFF (no ablation, top row of Table 4) improves the accuracy, sensitivity, specificity and AUC by 6.5, 5.3, 8.3, 4.2 percentages, respectively. The joint use of deep features from Nakagami parametric map (ablation of Branch-3, 2nd row of Table 4) improves the accuracy, sensitivity, specificity and AUC by 4.2, 3.2, 5.5, 2.4 percentages, respectively, while the joint use of handcrafted quantitative features improves the quantities by 2.3, 2.1, 2.8, 1.8 percentages, respectively. The ROC curves of the ablated models are shown in Fig. 4, where it can be seen that the ROC curve of the DCNN-MFF model with no ablation is better than those of the ablated models. These results demonstrate the usefulness and efficacy of the joint use of image features from both US B-mode image and Nakagami parametric map for recognizing significant liver fibrosis.

B. COMPARISON WITH OTHER DCNN MODELS

We further compare the performance of the proposed model against the other commonly used DCNN models including ResNet50 [60], InceptionV3 [61], DenseNet101 [62]. The comparative models were adapted to our study by changing the last FC layer to be with 2 neurons to determine whether the fibrosis is significant or not. Moreover, the parameters of these models were initialized by using the ones that had

TABLE 3. Performance of the DCNN-MFF Model on the Training/Validation/Testing Datasets.

Result	ACC	SEN	SPE	AUC
Training Dataset	1.0	1.0	1.0	1.0
Validation Dataset	0.875 (0.815-0.921)	0.863 (0.777- 0.925)	0.890(0.795-0.952)	0.929 (0.879-0.963)
Testing Dataset	0.827 (0.762-0.881)	0.821 (0.729-0.892)	0.836 (0.731-0.912)	0.869 (0.809-0.916)

Note—ACC: Accuracy; SEN: Sensitivity; SPE: Specificity; AUC: Area Under the Receiver Operating Characteristic Curve.

TABLE 4. Ablation test results.

Ablation	ACC	SEN	SPE	AUC
None	0.827 (0.762-0.881)	0.821 (0.729-0.892)	0.836 (0.731-0.912)	0.869 (0.809-0.916)
Branch-3	0.804 (0.735-0.861)	0.800 (0.705-0.875)	0.808 (0.699-0.891)	0.851 (0.788-0.901)
Branch-1/3	0.786 (0.716-0.845)	0.758 (0.659-0.840)	0.822 (0.715-0.902)	0.833 (0.768-0.886)
Branch-2/3	0.762 (0.690-0.824)	0.768 (0.671-0.849)	0.753 (0.639-0.847)	0.827 (0.762-0.881)

Note—ACC: Accuracy; SEN: Sensitivity; SPE: Specificity; AUC: Area Under the Receiver Operating Characteristic Curve.

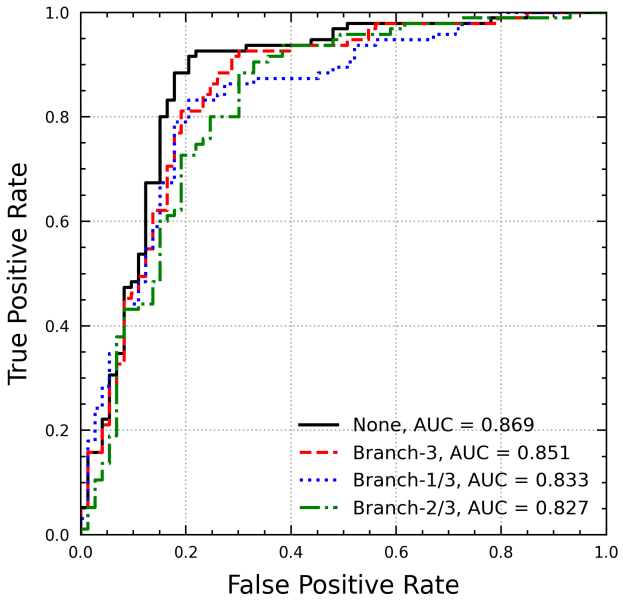


FIGURE 4. ROC curves obtained in the ablation test.

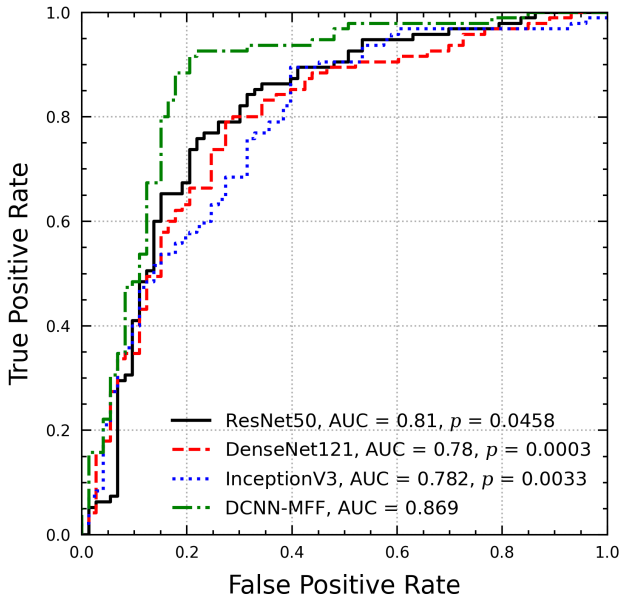


FIGURE 5. ROC curves obtained by different deep learning models.

been trained on the ImageNet dataset, and all of them were fine-tuned on the segmented US B-mode images with the same loss function as defined in the DCNN-MFF model. The results of the comparative models and the proposed one are summarized in Table 5. As expected, the proposed DCNN-MFF model significantly outperforms the other commonly used DCNN models, benefitting from the joint use of US B-mode images and Nakagami parametric map. Moreover, from the upper three rows of Table 5, it can be seen that the performances of the commonly used models are quite similar. Further, by comparing the results in the upper rows of Table 5 with those in the bottom row of Table 4, we found the comparative models also provide performance close to the baseline model of VGG16 (the latter has slightly better performance). These results suggest that similar performances could be obtained by replacing the baseline model in Branch-1 and Branch-2 of DCNN-MFF with the other commonly used DCNN models. The ROC curves of the comparative models and the proposed one are shown in Fig. 5. It can be seen that the curve of the proposed model is significantly better than the others, with all p-values < 0.05, DeLong test.

IV. DISCUSSION

Due to the advantages of real-time imaging, high sensitivity, and non-invasiveness, ultrasonography plays an important role in repeated assessments of liver fibrosis. Nevertheless, the diagnosis based on visual inspection on US images is time-consuming and empirical. In this paper, we propose a DCNN model with multi-feature fusion (DCNN-MFF) to integrate features from US B-mode image and Nakagami parametric map for significant liver fibrosis recognition. The DCNN-MFF model mainly consists of three branches, among which the first two branches are used for extracting deep features respectively from the US B-mode image and the

TABLE 5. Comparison of various DCNN models.

Models	ACC	SEN	SPE	AUC
ResNet50	0.750 (0.678-0.814)	0.716 (0.614-0.804)	0.795 (0.684-0.880)	0.810 (0.742-0.866)
DenseNet121	0.744 (0.671-0.808)	0.758 (0.659-0.840)	0.726 (0.609-0.824)	0.780 (0.722-0.850)
Inceptionv3	0.738 (0.665-0.803)	0.905 (0.828-0.956)	0.521 (0.400-0.639)	0.782 (0.709-0.840)
DCNN-MFF	**0.827 (0.762-0.881)**	**0.821 (0.729-0.892)**	**0.836 (0.731-0.912)**	**0.869 (0.809-0.916)**

Note—ACC: Accuracy; SEN: Sensitivity; SPE: Specificity; AUC: Area Under the Receiver Operating Characteristic Curve.

Nakagami parametric map, while the remaining branch is designed for extracting quantitative US features from the Nakagami parametric map. At the backend of the DCNN model, the extracted deep and quantitative features are integrated together for final decision making. The added values of deep and handcrafted features extracted from the Nakagami parametric map to the deep features of US B-mode images were validated on an animal data comprising 168 rat liver lobes with pathologically confirmed fibrosis stages.

As mentioned in the introduction Section, possible CAD systems could be established by using handcrafted feature-based methods or deep learning-based approaches. In respect to handcrafted feature-based methods: Mojsilovic *et al.* [13] developed a filter bank algorithm based on nonseparable wavelet decomposition to extract features from liver B-mode images. By using a Bayes classification scheme, they achieved an accuracy of 0.95 for differentiating cirrhosis (20 human liver tissue samples) from normal (37) and steatosis (65) human liver tissues. Yeh *et al.* [14] applied gray level concurrence and nonseparable wavelet transform to extract features from 20 human liver US images. In their study, the severity of liver fibrosis was graded from 0 to 5, with Grades 0-3 corresponding to fibrosis stages from S0 to S3 and Grade 4-5 corresponding to S4 in the METAVIR score system. By using support vector machine as the classifier, they achieved an accuracy 0.72 for differentiating the six fibrosis grades. Wu *et al.* [15] evaluated US liver images by using a group of handcrafted features derived from the spatial gray-level dependence matrices, the Fourier power spectrum, the gray-level difference statistics, and the Laws texture energy measures. In their study, the Bayes classifier and the Hotelling trace criterion were employed to evaluate the performance of handcrafted features on classifying human normal liver, hepatoma and cirrhosis. They achieved an accuracy of $\sim$0.9 for three sets of US liver images. In respect to deep learning-based approaches: Meng *et al.* [24] proposed a two-staged DCNN model with transfer learning to classify liver fibrosis US images. The network was evaluated to predict normal, early and late-stage of human liver fibrosis on two testing datasets. Lee *et al.* [27] retrained the VGG16 model on a relatively large dataset containing 13,608 liver US images of 3446 patients to predict the METAVIR scores. The performance of the retrained VGG16 model was evaluated by using both internal (300 images of 266 patients) and external validation (1232 images of 572 patients), resulting accuracies of 0.835 and 0.764, respectively. As previously mentioned, all the above-described intelligent methods used US B-mode images for liver fibrosis classification. Due to the underlying disadvantages of US B-mode images, such as the previously mentioned information loss and distortion problem, the performance of CAD systems developed based on those approaches may be bottlenecked. Therefore, the proposed method can be considered as a promising tool to further improve the diagnosis capacity of the conventional B-mode image-based CAD systems, by exploring and combing the information carried by the Nakagami parametric map.

Finally, we need to acknowledge that our method has limitations. One limitation is that a ROI indicating the liver parenchymal area needs to be manually identified. Moreover, our method requires the processing of raw RF signals before US B-mode imaging for fibrosis classification, while the availability of the RF data on US machines needs the consent of manufacturers. Finally, the performance of our proposed method was only assessed on an animal model, while its effectiveness to assess liver fibrosis on human subjects need to be further evaluated.

V. CONCLUSION

In this paper, we have proposed a new deep convolutional neural network with multiple feature fusion (DCNN-MFF) to extract and combine features from US B-mode image and Nakagami parametric map for recognizing significant liver fibrosis. In the proposed DCNN-MFF model, both deep and handcrafted features are extracted from the Nakagami parametric map, which are integrated with the deep features extracted from the US B-mode image. The added values of deep and handcrafted features extracted from the Nakagami parametric map to the deep features of US B-mode images have been validated on an animal model with a number of rats under various liver fibrosis stages. The proposed method may serve as a good reference for future extension to assess liver fibrosis on human subjects. In the future, we shall improve our model to be a fully automatic network, by which the ROI of liver parenchymal regions could be conveniently identified without any necessity of manual delineation.

REFERENCES

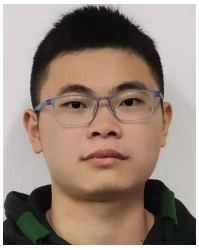
[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, A Cancer J. Clinicians*, vol. 68, no. 6, pp. 394–424, Nov. 2018.

[2] R. Bataller and D. A. Brenner, "Liver fibrosis," *J. Clin. Invest.*, vol. 115, no. 2, pp. 209–218, 2005.
[3] G. Marchesini, "Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome," *Hepatology*, vol. 37, no. 4, pp. 917–923, Apr. 2003.
[4] K. Ishak, A. Baptista, L. Bianchi, F. Callea, J. De Groote, F. Gudat, H. Denk, V. Desmet, G. Korb, R. N. M. MacSween, M. J. Phillips, B. G. Portmann, H. Poulsen, P. J. Scheuer, M. Schmid, and H. Thaler, "Histological grading and staging of chronic hepatitis," *J. Hepatology*, vol. 22, no. 6, pp. 696–699, Jun. 1995.
[5] P. Bedossa and T. Poynard, "An algorithm for the grading of activity in chronic hepatitis C," *Hepatology*, vol. 24, no. 2, pp. 289–293, Aug. 1996.
[6] L. Petitclerc, G. Sebastiani, G. Gilbert, G. Cloutier, and A. Tang, "Liver fibrosis: Review of current imaging and MRI quantification techniques," *J. Magn. Reson. Imag.*, vol. 45, no. 5, pp. 1276–1295, May 2017.
[7] T. Ren, H. Wang, R. Wu, and J. Niu, "Gamma-glutamyl transpeptidase-to-platelet ratio predicts significant liver fibrosis of chronic hepatitis b patients in China," *Gastroenterology Res. Pract.*, vol. 2017, pp. 1–7, Jul. 2017.
[8] W. Li, Y. Huang, B.-W. Zhuang, G.-J. Liu, H.-T. Hu, X. Li, J.-Y. Liang, Z. Wang, X.-W. Huang, C.-Q. Zhang, S.-M. Ruan, X.-Y. Xie, M. Kuang, M.-D. Lu, L.-D. Chen, and W. Wang, "Multiparametric ultrasomics of significant liver fibrosis: A machine learning-based analysis," *Eur. Radiol.*, vol. 29, no. 3, pp. 1496–1506, Mar. 2019.
[9] W.-Y. Kao, I.-W. Chang, C.-L. Chen, C.-W. Su, S. U. Fang, J.-H. Tang, C.-C. Chang, Y.-J. Chang, and W. Wang, "Fibroscan-based score to predict significant liver fibrosis in morbidly obese patients with nonalcoholic fatty liver disease," *Obesity Surg.*, vol. 30, no. 4, pp. 1249–1257, Apr. 2020.
[10] P. Lamb, D. V. Sahani, J. M. Fuentes-Orrego, M. Patino, A. Ghosh, and P. R. S. Mendonca, "Stratification of patients with liver fibrosis using dual-energy CT," *IEEE Trans. Med. Imag.*, vol. 34, no. 3, pp. 807–815, Mar. 2015.
[11] L. B. Seeff, G. T. Everson, T. R. Morgan, T. M. Curto, W. M. Lee, M. G. Ghany, M. L. Shiffman, R. J. Fontana, A. M. Di Bisceglie, H. L. Bonkovsky, and J. L. Dienstag, "Complication rate of percutaneous liver biopsies among persons with advanced chronic liver disease in the HALT-C trial," *Clin. Gastroenterology Hepatology*, vol. 8, no. 10, pp. 877–883, Oct. 2010.
[12] B. Stotland and G. Lichtenstein, "Liver biopsy complications and routine ultrasound," *Amer. J. Gastroenterol.*, vol. 91, no. 7, pp. 1295–1296, 1996.
[13] A. Mojsilovic, M. Popovic, S. Markovic, and M. Krstic, "Characterization of visually similar diffuse diseases from B-scan liver images using nonseparable wavelet transform," *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 541–549, 1998.
[14] W.-C. Yeh, S.-W. Huang, and P.-C. Li, "Liver fibrosis grade classification with B-mode ultrasound," *Ultrasound Med. Biol.*, vol. 29, no. 9, pp. 1229–1235, Sep. 2003.
[15] C.-M. Wu, Y.-C. Chen, and K.-S. Hsieh, "Texture features for classification of ultrasonic liver images," *IEEE Trans. Med. Imag.*, vol. 11, no. 2, pp. 141–152, Jun. 1992.
[16] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural Comput.*, vol. 29, no. 9, pp. 2352–2449, Sep. 2017.
[17] X. Jiang, A. Hadid, Y. Pang, E. Granger, and X. Feng, *Deep Learning in Object Detection and Recognition*. Cham, Switzerland: Springer, 2019.
[18] L. Deng and Y. Liu, *Deep Learning in Natural Language Processing*. Cham, Switzerland: Springer, 2018.
[19] Z. Shao, L. Wang, Z. Wang, W. Du, and W. Wu, "Saliency-aware convolution neural network for ship detection in surveillance video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 3, pp. 781–794, Mar. 2020.
[20] R. Zhang, Z. Shao, X. Huang, J. Wang, and D. Li, "Object detection in UAV images via global density fused convolutional network," *Remote Sens.*, vol. 12, no. 19, p. 3140, Sep. 2020.
[21] Z. Shao and J. Cai, "Remote sensing image fusion with deep convolutional neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1656–1669, May 2018.
[22] Z. Shao, J. Cai, P. Fu, L. Hu, and T. Liu, "Deep learning-based fusion of Landsat-8 and Sentinel-2 images for a harmonized surface reflectance product," *Remote Sens. Environ.*, vol. 235, Dec. 2019, Art. no. 111425.
[23] S. Liu, Y. Wang, X. Yang, B. Lei, L. Liu, S. X. Li, D. Ni, and T. Wang, "Deep learning in medical ultrasound analysis: A review," *Engineering*, vol. 5, no. 2, pp. 261–275, Apr. 2019.
[24] D. Meng, L. Zhang, G. Cao, W. Cao, G. Zhang, and B. Hu, "Liver fibrosis classification based on transfer learning and FCNet for ultrasound images," *IEEE Access*, vol. 5, pp. 5804–5810, 2017.
[25] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: http://arxiv.org/abs/1409.1556
[26] Y. Mitani, R. B. Fisher, Y. Fujita, Y. Hamamoto, and I. Sakaida, "Effect of an augmentation on CNNs in classifying a cirrhosis liver on B-mode ultrasound images," in *Proc. IEEE 2nd Global Conf. Life Sci. Technol. (LifeTech)*, Mar. 2020, pp. 253–254.
[27] J. H. Lee, I. Joo, T. W. Kang, Y. H. Paik, D. H. Sinn, S. Y. Ha, K. Kim, C. Choi, G. Lee, J. Yi, and W.-C. Bang, "Deep learning with ultrasonography: Automated classification of liver fibrosis using a deep convolutional neural network," *Eur. Radiol.*, vol. 30, no. 2, pp. 1264–1273, Feb. 2020.
[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
[29] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
[30] M. L. Oelze and J. Mamou, "Review of quantitative ultrasound: Envelope statistics and backscatter coefficient imaging and contributions to diagnostic ultrasound," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 63, no. 2, pp. 336–351, Feb. 2016.
[31] J.-J. Lin, J.-Y. Cheng, L.-F. Huang, Y.-H. Lin, Y.-L. Wan, and P.-H. Tsui, "Detecting changes in ultrasound backscattered statistics by using Nakagami parameters: Comparisons of moment-based and maximum likelihood estimators," *Ultrasonics*, vol. 77, pp. 133–143, May 2017.
[32] D. P. Hruska and M. L. Oelze, "Improved parameter estimates based on the homodyned k distribution," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 56, no. 11, pp. 2471–2481, Nov. 2009.
[33] F. Destrempes and G. Cloutier, "A critical review and uniformized representation of statistical distributions modeling the ultrasound echo envelope," *Ultrasound Med. Biol.*, vol. 36, no. 7, pp. 1037–1051, Jul. 2010.
[34] W.-K. Chan, N. R. Nik Mustapha, and S. Mahadeva, "Controlled attenuation parameter for the detection and quantification of hepatic steatosis in nonalcoholic fatty liver disease," *J. Gastroenterology Hepatology*, vol. 29, no. 7, pp. 1470–1476, Jul. 2014.
[35] K. Samimi and T. Varghese, "Ultrasonic attenuation imaging using spectral cross-correlation and the reference phantom method," in *Proc. IEEE Int. Ultrason. Symp.*, Oct. 2011, pp. 53–55.
[36] G. Ghoshal, R. J. Lavarello, J. P. Kemmerer, R. J. Miller, and M. L. Oelze, "*Ex vivo* study of quantitative ultrasound parameters in fatty rabbit livers," *Ultrasound Med. Biol.*, vol. 38, no. 12, pp. 2238–2248, Dec. 2012.
[37] V. Roberjot, S. L. Bridal, P. Laugier, and G. Berger, "Absolute backscatter coefficient over a wide range of frequencies in a tissue-mimicking phantom containing two populations of scatterers," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 43, no. 5, pp. 970–978, Sep. 1996.
[38] F. L. Lizzi, M. Astor, T. Liu, C. Deng, D. J. Coleman, and R. H. Silverman, "Ultrasonic spectrum analysis for tissue assays and therapy evaluation," *Int. J. Imag. Syst. Technol.*, vol. 8, no. 1, pp. 3–10, 1997.
[39] M.-C. Ho, J.-J. Lin, Y.-C. Shu, C.-N. Chen, K.-J. Chang, C.-C. Chang, and P.-H. Tsui, "Using ultrasound nakagami imaging to assess liver fibrosis in rats," *Ultrasonics*, vol. 52, no. 2, pp. 215–222, Feb. 2012.
[40] M.-C. Ho, P.-H. Tsui, Y.-H. Lee, Y.-S. Chen, C.-N. Chen, J.-J. Lin, and C.-C. Chang, "Early detection of liver fibrosis in rats using 3-D ultrasound Nakagami imaging: A feasibility evaluation," *Ultrasound Med. Biol.*, vol. 40, no. 9, pp. 2272–2284, Sep. 2014.
[41] P.-H. Tsui, M.-C. Ho, D.-I. Tai, Y.-H. Lin, C.-Y. Wang, and H.-Y. Ma, "Acoustic structure quantification by using ultrasound Nakagami imaging for assessing liver fibrosis," *Sci. Rep.*, vol. 6, no. 1, pp. 1–9, Dec. 2016.
[42] M. F. Insana, R. F. Wagner, D. G. Brown, and T. J. Hall, "Describing small-scale structure in random media using pulse-echo ultrasound," *J. Acoust. Soc. Amer.*, vol. 87, no. 1, pp. 179–192, Jan. 1990.
[43] J. C. Bamber and C. R. Hill, "Acoustic properties of normal and cancerous human liver—I. Dependence on pathological condition," *Ultrasound Med. Biol.*, vol. 7, no. 2, pp. 121–133, Jan. 1981.
[44] M. Nakagami, *Statistical Methods in Radio Wave Propagation*. Oxford, U.K.: WC Hoffman, 1960.
[45] M. Han, N. Wang, S. Guo, N. Chang, S. Lu, and M. Wan, "Nakagami-m parametric imaging for characterization of thermal coagulation and cavitation erosion induced by HIFU," *Ultrason. Sonochemistry*, vol. 45, pp. 78–85, Jul. 2018.
[46] N. Bahbah, A. Novell, A. Bouakaz, and H. Djelouah, "Linear and nonlinear characterization of microbubbles and tissue using the nakagami statistical model," *Ultrasonics*, vol. 76, pp. 200–207, Apr. 2017.

[47] P.-H. Tsui, Y.-L. Wan, D.-I. Tai, and Y.-C. Shu, "Effects of estimators on ultrasound Nakagami imaging in visualizing the change in the backscattered statistics from a Rayleigh distribution to a pre-Rayleigh distribution," *Ultrasound Med. Biol.*, vol. 41, no. 8, pp. 2240–2251, Aug. 2015.
[48] S. Zhang, S. Shang, Y. Han, C. Gu, S. Wu, S. Liu, G. Niu, A. Bouakaz, and M. Wan, "*Ex vivo* and *in vivo* monitoring and characterization of thermal lesions by high-intensity focused ultrasound and microwave ablation using ultrasonic Nakagami imaging," *IEEE Trans. Med. Imag.*, vol. 37, no. 7, pp. 1701–1710, Jul. 2018.
[49] M. Lin, Q. Chen, and S. Yan, "Network in network," 2013, *arXiv:1312.4400*. [Online]. Available: http://arxiv.org/abs/1312.4400
[50] S. Rizzo, F. Botta, S. Raimondi, D. Origgi, C. Fanciullo, A. G. Morganti, and M. Bellomi, "Radiomics: The facts and the challenges of image analysis," *Eur. Radiol. Experim.*, vol. 2, no. 1, pp. 1–8, Dec. 2018.
[51] J.-H. Kim, K.-W. On, W. Lim, J. Kim, J.-W. Ha, and B.-T. Zhang, "Hadamard product for low-rank bilinear pooling," 2016, *arXiv:1610.04325*. [Online]. Available: http://arxiv.org/abs/1610.04325
[52] Z. Yu, J. Yu, J. Fan, and D. Tao, "Multi-modal factorized bilinear pooling with co-attention learning for visual question answering," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1821–1830.
[53] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. Int. Conf. Artif. Intell. Stat.*, 2011, pp. 315–323.
[54] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. Int. Conf. Artif. Intell. Stat.*, 2010, pp. 249–256.
[55] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
[56] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006.
[57] W. J. Youden, "Index for rating diagnostic tests," *Cancer*, vol. 3, no. 1, pp. 32–35, 1950.
[58] C. J. Clopper and E. S. Pearson, "The use of confidence or fiducial limits illustrated in the case of the binomial," *Biometrika*, vol. 26, no. 4, pp. 404–413, 1934.
[59] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach," *Biometrics*, vol. 44, pp. 837–845, Sep. 1988.
[60] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[61] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.
[62] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.

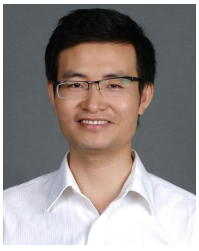
QIANG LIU received the B.S. degree in biomedical engineering from the Taiyuan University of Technology, in 2017, and the M.S. degree in biomedical engineering from Shenzhen University, in 2020. He is currently a Research Assistant with the School of Biomedical Engineering, Shenzhen University, Shenzhen, China. His research interests include medical image processing and ultrasonic signal processing analysis.

ZHONG LIU received the B.Eng. degree from the Huaihai Institute of Technology, Lianyungang, China, in 2010, the M.Eng. degree from the Zhongyuan University of Technology, Zhengzhou, China, in 2013, and the Ph.D. degree from the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, in 2018. He is currently a Research Fellow with the Department of Biomedical Engineering, Shenzhen University, Shenzhen, China. His research interests include signal processing, pattern recognition, and human–computer interaction.



WENCONG XU received the bachelor's degree in biomedical engineering from Nanchang Hangkong University, Nanchang, China, in 2017. He is currently pursuing the master's degree in biomedical engineering with Shenzhen University, Shenzhen, China. From 2018 to 2020, he was with the Ultrasound Laboratory, Shenzhen University, Shenzhen. His research interests include ultrasound signal processing and image processing, application of quantitative ultrasound parameters to diagnose diseases in ultrasound images, and the method of using deep learning to diagnose in ultrasound images.

HUIYING WEN received the M.S. degree from the Department of Biomedical Engineering, Shenzhen University, where she is currently pursuing the Ph.D. degree. Her research interests include image processing and machine learning for medical diagnostics.



MING DAI received the B.S. degree in computer science and technology from Hainan Tropical Ocean University, Sanya, Hainan, in 2012, the dual M.S. degree in electronics and communication engineering from Nanchang Hangkong University, Nanchang, Jiangxi, and Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, in 2015, and the Ph.D. degree in information and communication engineering from Shenzhen University, Shenzhen, Guangdong, in June 2019. He was a Visiting Scholar with The Chinese University of Hong Kong, from 2019 to 2020. He currently holds a postdoctoral position with Jinan University and is studying with Shenzhen University. His research interests include electrical conductivity imaging, biomedical multimodal data fusion, image processing, artificial intelligence, things engineering, wearable medical applications, and design of electrical systems.

XIN CHEN (Member, IEEE) received the B.S. degree in electrical engineering and the Ph.D. degree in biomedical engineering from the University of Science and Technology of China, in 1998 and 2003, respectively. He then joined the Department of Health Technology and Informatics, The Hong Kong Polytechnic University, as a Research Associate. In 2008, he joined the Department of Biomedical Engineering, Shenzhen University, as an Assistant Professor, where he became an Associate Professor, in 2011, and a Professor, in 2017. His research interests include ultrasound elasticity imaging and measurement, ultrasound assessment of musculoskeletal tissues, biomedical conductivity imaging, and biomedical instrumentation.