2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI) | 978-1-6654-0004-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/CISP-BMEI53629.2021.9624402

Evaluation of liver fibrosis based on ultrasound radio frequency signals

Jinli Xu
Department of Electronic Engineering, Fudan University
ShangHai, China

Liyun Xue
Department of Ultrasound, Huashan hospital, Fudan University
ShangHai, China

Jinhua Yu
Department of Electronic Engineering, Fudan University
ShangHai, China

Hong Ding
Department of Ultrasound, Huashan hospital, Fudan University
ShangHai，China

***Abstract*—Liver fibrosis is a common consequence of almost all chronic liver diseases. Precise and timely evaluation of liver fibrosis progression is essential for the treatment of liver disease. In this paper, we present a deep learning-based framework for evaluating the degree of liver fibrosis, using ultrasound radio frequency signals. The deep learning model based on long short-term memory (LSTM) network and attention cell. The dataset consisted of 96 sets of ultrasound radio frequency signals of rat livers, with five fibrosis stages ranging from 0 to 4. The accuracy and the areas under the receiver operating characteristic curve of the four classification models were greater than 0.81 and 0.89, respectively. This study indicated that evaluation system of liver fibrosis stage based on deep learning approaches and ultrasound RF signals was promising and it would be of great value in monitoring liver fibrosis precisely and non-invasively.**

Keywords-component; Liver fibrosis; Ultrasound radio frequency; Deep learning

I. INTRODUCTION *(HEADING 1)*

Liver fibrosis is a common consequence of almost all chronic liver diseases. If the etiology cannot be discovered and eliminated in time, liver fibrosis will progress to cirrhosis and eventually to liver failure or malignancy. Effective intervention in the process of liver fibrosis or early cirrhosis can stop or revert the progression of hepatic fibrosis and improve liver function [1][2]. At present, liver biopsy is recognized as the gold standard for evaluating of staging of liver fibrosis [3]. However, despite its universal use, liver biopsy has many disadvantages that patients cannot accept. Liver biopsies are expensive, risky and invasive, requiring a puncture needle through the skin to the parenchymal portion of the liver, which cause tremendous pain to the patient. Furthermore, needle liver biopsy samples only about 1/50,000 of the liver and so causes sampling error. Significant complications occur in 1-5% of patients [4].

Ultrasound B-mode imaging is widely used to examine human tissue structures because of its convenience, affordability and safety. However, due to the massive postprocessing operations, such as the dynamic range setting or filtering operations, the information that can be extracted from B-mode data is limited [5]. The radio frequency (RF) signal is a raw type of data, and the ultrasonic B-mode image is demodulated from this data through. Therefore, RF data contain more information than B-mode images. Studies have shown that the acoustic information contained in the radio frequency signal, including attenuation, backscatter, sound velocity, phase, etc., could further reflect the changes in the microscopic level of biological tissues [6]. Ultrasound radio frequency signal was applied to Breast calcification detection achieving an accuracy of 84% and an F1 score of 91% [7]. Han et al. used ultrasound radio frequency data to quantitatively analyze liver fat [8]. Therefore, it may be helpful to diagnose the procession of liver fibrosis with ultrasound radio frequency data.

With the growth of computer computing power, deep learning has played an increasingly important role in the medical field, especially convolution neural network (CNN) and recurrent neural network (RNN). The RNN does a good job at handling sequence data with internal connections. In this paper, rats were used as experimental subjects. We developed a deep learning model based on ultrasound radio frequency data to evaluate the grade of liver fibrosis. By combining the recurrent neural network and the attention cell, we can generate good classification results.

II. RELATE WORK

In the task of evaluating the degree of liver fibrosis, there are mainly three methods according to the type of ultrasound data source: the ultrasound B-mode images, the ultrasound radio frequency signals, the ultrasound radio frequency time series.

The principle of the method based on the ultrasound B-mode images is that the morphological manifestations of the diseased tissue and the normal tissue are different, and the morphological related features can be extracted through ultrasound imaging to achieve the purpose of identifying the tissue. Zhou G et al. [9] used ultrasound images of the liver to distinguish fibrotic livers from healthy livers. This study used a gray-level co-occurrence matrix to calculate 13 texture feature parameters from the image, constructed a Fisher linear classifier and used 83 As a training set, the results showed that the sensitivity of the classifier was 81.0% and the specificity was 70.6%. However, ultrasound-based image analysis faces many limitations. For example, research methods based on image texture characteristics cannot avoid the effects of speckle noise, and only present the different characteristics of liver fibrosis in various stages to a certain extent.

The analysis method based on the time series of ultrasound radio frequency signals combines the data of the same position in multiple frames of radio frequency signals to construct a radio frequency signal time series at that position, then calculates the feature parameters of the time series, and finally uses a classifier to classify. C.-Y. Lin et al. [10] combined the spectral and fractal features with the time-domain features of ultrasound radio frequency time series and got an average classification accuracy of 77.33%. However, in the existing research, the time series of ultrasound signals often need to store more than two hundred frames of signal data. The current hardware level can hardly accept such high computing and storage costs. Moreover, the research on the time series of ultrasound signals is still in its infancy, and the technology is still immature. There is still a long way to go before the time series of ultrasound signals can be used in clinical applications.

The analysis method based on the single-frame ultrasound radio frequency signals has been confirmed by many studies that it can reflect the changes of tissue microstructure [11][12]. As it has gone through many years, the storage cost and calculation cost have obvious advantages over the time series analysis method, it has more mature conditions for application.

III. METHODS

A. Dataset

The data used in this experiment came from rats with different stages of liver fibrosis. The rats were induced by thioacetamide, and the operating procedure was established as previously reported [13]. Table I gave the details of the data, the dataset contained a total of 96 cases, of which 80 cases were used to train and 16 cases were used to validate. The stage of liver fibrosis was evaluated according to the Scheuer fibrosis scoring system [14]: S0, absence of fibrosis; S1, fibrous portal expansion; S2, periportal or rare portal–portal septa; S3, fibrous septa with architectural distortion; S4, cirrhosis.

TABLE I. DISTRIBUTION OF DATA.

	Total	***S0***	***S1***	***S2***	***S3***	***S4***
Dataset	96	21	28	28	10	9
Training	80	17	24	24	8	7
Validating	16	4	4	4	2	2

B. Method overview

The experimental process included the following parts：

1. The ultrasound image was rebuilt from the RF data.
2. The liver parenchymal area (regions of interest, ROI) was labeled on the ultrasound image.
3. The ROI at the same position in the RF data was extracted.
4. The RF data in the ROI area were cropped into many one-dimensional signal segments. The length of the signal segment was 200. The excess part was truncated, and the insufficient part was padded with zeros.
5. Deep learning models were trained based on one-dimensional RF signal segments.
6. In the prediction stage, voting strategy was adopted to judge the overall degree of liver fibrosis.

C. Data preprocessing

First, after logarithmic compression and Hilbert transform, the ultrasound RF data were converted into B-mode images that we could recognize. Then we labeled the regions of interest (the liver parenchyma area) on the images. The labeling process follows the following rules:

1) Used a rectangle with an unfixed size
2) Avoided the blood vessel area as much as possible
3) Kept the same depth as much as possible.

As shown in Fig. 1, the part inside the dark blue rectangle was the region of interest for deep learning analysis.

The position of the ROI area in the RF data was the same as the position of the ROI area in the ultrasound image. The RF data in the ROI area were cropped into many one-dimensional signal segments. The length of the signal segment was 200. The excess part was truncated, and the insufficient part was padded with zeros. Fig. 2 showed a one-dimensional RF signal segment extracted from the region of interest in Figure 1.

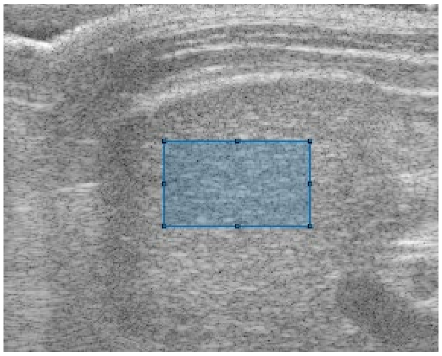

Fig 1. Ultrasound image with blue outline superimposed to indicate the region of interest

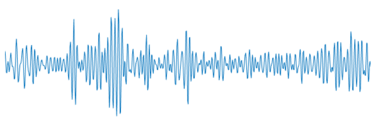

Fig 2. A one-dimensional RF signal segment

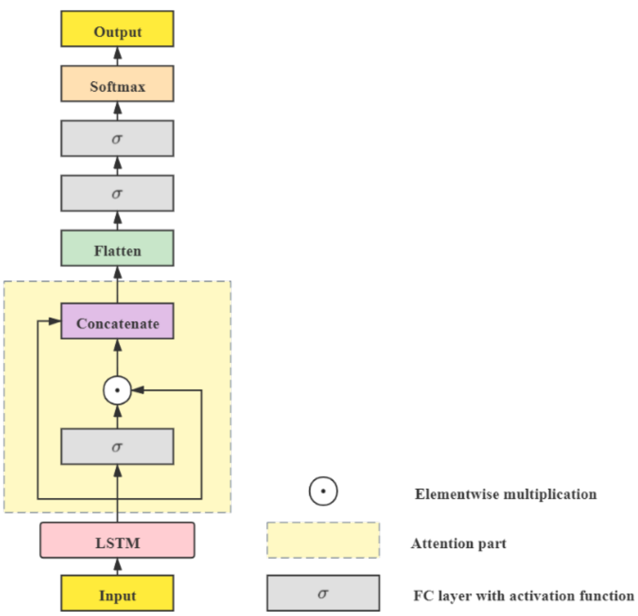

Fig 3. Network architecture

D. Network architecture

In traditional neural networks, the input layer, hidden layer , and output layer were fully connected, and the nodes in the layer were not connected. Therefore, the traditional neural network model cannot find the internal connections in the data. For example, in the task of predicting the next word in a sentence, the previous word need to be used due to the words in the sentence were not independent. The recurrent neural network (RNN) [15] was designed to process this type of sequence data. In the RNN network, the current output of a sequence was also related to the previous output. However, traditional RNN suffered from the problem of vanishing gradients, which made it difficult for the model to learn the internal connections of the data.

Long short-term memory networks (LSTM) [16] were a special type of recurrent neural network that were good at learning long-term dependencies. A memory cell was introduced to record additional information in LSTM networks. Additionally, in order to control the memory cell three different gate structures were introduced into the model: Forget gate F_t determined what information would be thrown away from the cell state; Input gate I_t determined what new information would be stored in the cell state; Output gate O_t determined what entries would be read out from the cell.

Figure 4 had a graphical illustration of LSTM cell structure. X_t was the input at the current time step. H_{t-1} was the hidden state of the previous time step. C_{t-1} was the memory cell of the previous time step.

$$I_t = \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i) \tag{1}$$

$$F_t = \sigma(X_t W_{xf} + H_{t-1} W_{hf} + b_f) \tag{2}$$

$$O_t = \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o) \tag{3}$$

$$\tilde{C}_t = tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c) \tag{4}$$

$$C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \tag{5}$$

$$H_t = O_t \odot \tanh(C_t) \tag{6}$$

Where $W_{xi}, W_{xf}, W_{xo}, W_{xc}, W_{hi}, W_{hf}, W_{hc}, W_{ho}$ were weight parameters and b_i, b_f, b_o, b_c were bias parameters. The above eight variables were trainable parameters that can be adaptively adjusted during training. $\tilde{C}_t$ was the candidate memory cell.

The mechanism for managing input and forgetting in LSTM was implemented in the calculation process of C: the input gate I_t controled how much new data from $\tilde{C}_t$ was taken into account and the forget gate F_t governed how much of the old memory cell content C_{t-1} was retained. This design could effectively alleviate the problem of the vanishing gradient and better capture long range dependencies in the sequence.

The light yellow part in Figure 3 was the attention cell. Note that it was not the standard attention module proposed in Attention all your need [17]. A fully connected layer with an activation function of softmax was introduced, which outputed

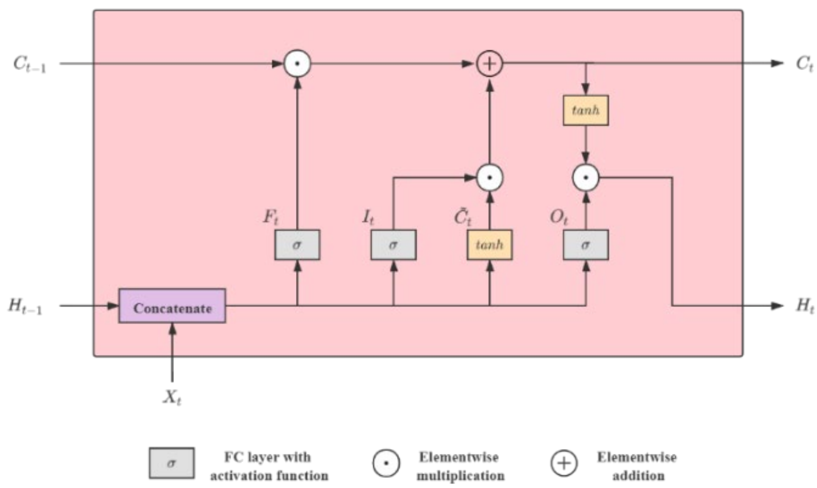

Fig 4. Long short-term memory networks

a set of weights A_h to represent the attention of each column (Because of the matrix dimensions, a flatten layer between the fully connected layer and the activation function softmax is not drawn in the figure, neither the repeating vector and permuting operation after the activation function).

$$A_h = W_h H + b_h \tag{7}$$

$$B_h = A_h \odot H \tag{8}$$

$$O_h = [H, B_h] \tag{9}$$

Where W_h and b_h were weight and bias that can be leaned during the training. H was the output of the LSTM cell. B_h was the data allocated by attention. O_h was the final output of the attention cell.

The training process of the network mainly included forward propagation and backpropagation. Forward propagation, the input data passed through LSTM cell, attention cell, and two fully connected layers in turn, and finally used the Softmax activation function to output the probability value of each category. The learnable parameters in the model were adjusted through backpropagation. In training, the cross-entropy loss function was selected as the loss function. Adaptive moment estimate (Adam) [18] was selected as the optimization algorithm.

IV. Results

In the experiment, four binary classifiers were built using the same network structure: S0<, S1<, S2<, and S3< (S0< denoted the binary classifier dividing the five classes of liver fibrosis into two categories: one for S0 and the other for S1, S2, S3, and S4). In the model training phase, 80 training data were cropped into 13,349 RF signal segments as the input of the model. In the model evaluation stage, each verification case was first cropped into RF signal segments to evaluate the category of each signal segment. A score was then given to indicate the overall degree of liver fibrosis of the case by a voting strategy

We trained model from scratch and used the trained network for evaluation on the validation set. We used four schemes for evaluation: Sensitivity, specificity, area under the receiver operating characteristic curve and accuracy. Sensitivity was also called the true positive (TPR). It was explained in equation 10:

$$Sensitivity = \frac{true\ positives}{true\ positives + false\ negatives} \tag{10}$$

Specificity was also called the true negative (TNR) and it was explained by equation 11:

$$Specificity = \frac{true\ negatives}{true\ negatives + false\ positives} \tag{11}$$

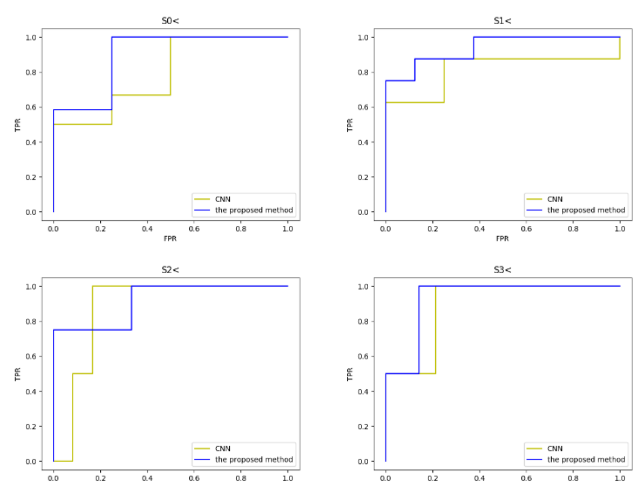

Fig 5. Area under the receiver operating characteristic curve of the four classifiers of CNN and the proposed method

Identify applicable sponsor/s here. *(sponsors)*

TABLE II. EVALUATION RESULT FOR CLASSIFICATION

	AUROC	*Accuracy*	*Sensitivity*	*Specificity*
S0 <	0.896	87.5%	91.7%	75%
S1 <	0.936	81.3%	87.5%	75%
S2 <	0.917	93.7%	75%	100%
S3 <	0.929	87.5%	50%	100%

Note: AUROC, Area under the receiver operating characteristic curve

TABLE III. PERFORMANCE OF CNN AND THE PROPOSED METHOD

	Method	*S0 <*	*S1 <*	*S2 <*	*S3 <*
AUROC	CNN	0.791	0.815	0.876	0.893
	Proposed method	0.896	0.936	0.917	0.929
ACC	CNN	75.0%	81.3%	87.5%	87.5%
	Proposed method	87.5%	81.3%	93.7%	87.5%

Table II gave the evaluation results of the four classification models on the validation set, the accuracy of the four networks were around 90%. S1< had the highest AUROC of 0.936, and the rest of the networks were also higher than 0.89. The results showed that our models were well trained and had good generalization performance.

In addition, we compared the proposed method with the traditional CNN network proposed by Dai [19]. Traditional CNN networks were built by stacking convolutional layers with one-dimensional convolutional kernels followed by pooling layers. Fig. 5 compared the area under the receiver operating characteristic curve of the two methods. Table III showed the comparison results of two methods, the proposed method's AUROC and accuracy were better than traditional CNN in the four classification cases.

V. CONCLUSIONS

Based on long short-term memory networks, this paper proposed a model to classify the grade of liver fibrosis based on ultrasound radio frequency signals. Attention cell was used as a supplement to LSTM to improve the accuracy. Our method receives one-dimensional RF signals extracted from the ROI area as inputs. The experiment was conducted on a data set with 96 rat cases. The accuracy and the areas under the receiver operating characteristic curve of the four classification models were greater than 0.81 and 0.89, respectively. Our preliminary results indicated that evaluation method of liver fibrosis stage based on deep learning approaches and ultrasound RF signals was promising.

REFERENCES

[1] H. M. Atta, "Reversibility and heritability of liver fibrosis: Implications for research and therapy," WJG, vol. 21, no. 17, p. 5138, 2015.

[2] Y. K. Jung and H. J. Yim, "Reversal of liver cirrhosis: current evidence and expectations," *Korean J Intern Med*, vol. 32, no. 2, pp. 213–228, Mar. 2017.

[3] C. Lucero and R. S. B. Jr, "Noninvasive Measures of Liver Fibrosis and Severity of Liver Disease," p. 8.

[4] N. H. Afdhal and D. Nunes, "Evaluation of Liver Fibrosis: A Concise Review," *Am J Gastroenterology*, vol. 99, no. 6, pp. 1160–1174, Jun. 2004.

[5] C. Carvalho *et al.*, "Deep Learning Approaches for Gynaecological Ultrasound Image Segmentation: A Radio-Frequency vs B-mode Comparison," in *Image Analysis and Recognition*, vol. 11663, F. Karray, A. Campilho, and A. Yu, Eds. Cham: Springer International Publishing, 2019, pp. 295–306.

[6] C. Abbey, N. Nguyen, and M. Insana, "Effects of frequency and bandwidth on diagnostic information transfer in ultrasonic B-Mode imaging," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 59, no. 6, pp. 1115–1126, Jun. 2012.

[7] M. Qiao, Z. Fang, Y. Guo, S. Zhou, C. Chang, and Y. Wang, "Breast calcification detection based on multichannel radiofrequency signals via a unified deep learning framework," *Expert Systems with Applications*, vol. 168, p. 114218, Apr. 2021.

[8] A. Han *et al.*, "Noninvasive Diagnosis of Nonalcoholic Fatty Liver Disease and Quantification of Liver Fat with Radiofrequency Ultrasound Data Using One-dimensional Convolutional Neural Networks," *Radiology*, vol. 295, no. 2, pp. 342–350, May 2020.

[9] G. Zhou, Y. Wang, and W. Wang, "Diagnosis of Hepatic Fibrosis by Ultrasonic Image Analysis," in *2012 International Conference on Biomedical Engineering and Biotechnology*, Macau, Macao, May 2012.

[10] C.-Y. Lin, T. Yi, Y.-Z. Gao, J.-H. Zhou, and Q.-H. Huang, "Early Detection and Assessment of Liver Fibrosis by using Ultrasound RF Time Series," J. Med. Biol. Eng., vol. 37, no. 5, pp. 717–729, Oct. 2017.

[11] G. Cheng *et al.*, "Quantitative evaluation of liver fibrosis based on ultrasound radio frequency signals: An animal experimental study," *Computer Methods and Programs in Biomedicine*, vol. 199, p. 105875, Feb. 2021.

[12] A. Han *et al.*, "Noninvasive Diagnosis of Nonalcoholic Fatty Liver Disease and Quantification of Liver Fat with Radiofrequency Ultrasound Data Using One-dimensional Convolutional Neural Networks," *Radiology*, vol. 295, no. 2, pp. 342–350, May 2020.

[13] Z. Yue MD *et al.*, "Histological Reference for Shear Wave Elastography in Liver Fibrosis: Collagen Quantification and Scoring System," *Advanced Ultrasound in Diagnosis and Therapy*, vol. 3, no. 3, p. 87, 2019.

[14] J. Schcuer, "Classification of chronic vira! hepatitis: a need for reassessment," p. 3, 1991.

[15] K. Funahashi and Y. Nakamura, "Approximation of dynamical systems by continuous time recurrent neural networks," *Neural Networks*, vol. 6, no. 6, pp. 801–806, Jan. 1993.

[16] S. Hochreiter, J. Schmidhuber, Long short-term memory. Neural Comput. 9(8), 1735–1780, 1997.

[17] A. Vaswani *et al.*, "Attention Is All You Need," *arXiv:1706.03762 [cs]*, Dec. 2017, Accessed: Oct. 12, 2021.

[18] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv:1412.6980 [cs]*, Jan. 2017, Accessed: Sep. 06, 2021.

[19] M. Dai, S. Li, Y. Wang, Q. Zhang, and J. Yu, "Post-processing radio-frequency signal based on deep learning method for ultrasonic microbubble imaging," *BioMed Eng OnLine*, vol. 18, no. 1, p. 95, Dec. 2019.