Ultrasound multiclass classification of hepatic steatosis using random forests

Ying-Hsiu Lin[1], Shi-Sheng Chen[2], Po-Hsiang Tsui[1]
[1]Department of Medical Imaging and Radiological Sciences, College of Medicine, Chang Gung University, Taoyuan, Taiwan
[2]Institute of Applied Mechanics, National Taiwan University, Taipei, Taiwan

Background, Motivation and Objective

Nonalcoholic fatty liver disease (NAFLD) is a risk factor for advanced liver diseases. Ultrasound is the first-line tool for screening NAFLD. To support clinical decision making, the quantitative assessment of NAFLD is a critical issue in clinical needs. The speckle pattern, attenuation, and echo intensity are sonofeatures that significantly correlate with the progression of hepatic steatosis. Combining the above three sonofeatures may provide more physical interpretations of NAFLD. Differently from the previous investigations, this study improved attenuation estimation using instantaneous frequency analysis based on the Hilbert-Huang transform, which was combined with the speckle and echo intensity analysis for developing a NAFLD multiclass classifier using random forests.

Statement of Contribution/Methods

The NAFLD patients diagnosed by liver biopsies were included (n = 201). The severity of steatosis was scored according to the pathological findings (0: <5%; 1: 5%-33%; 2: >33%). For each patient, a scanner equipped with a 3-MHz transducer was used for data acquisition to calculate the information entropy, attenuation coefficient, and integrated backscatter for characterizing microstructures (speckle pattern), viscoelastic properties, and echogenicities of the liver tissue, respectively. Random forests were applied to combining the sonofeatures with anthropometric (body mass index; BMI) and biochemical characteristics (aspartate aminotransferase and alanine aminotransferase; AST and ALT) for evaluating the performances of multiclass classification by comparisons with the histological score.

Results/Discussion

In contrast to using the single parameter, combining three sonofeatures using random forests allowed multiclass classification of NAFLD with the accuracy of 72.13%. Integrations of BMI, AST, and ALT in the computations of random forests had no contributions to the diagnostic performance. Of three sonofeatures, the information entropy was the principal component (weight ratio: 30.3682%) in the multiclass classification, as shown in the analysis report of random forests. The current results supported ultrasound multiclass classification of NAFLD using random forests, which may benefit clinical staging of NAFLD in the future.

Table. 1 A summary of the results obtained from random forests. The accuracy of multiclass classification was 72.13%. This study is the first to validate the multiclass classification of NAFLD by pathological evidences.

leave-n-out cross-validation	Accuracy	Sensitivity	Specificity	Variable importance dichotomy [%]	
10	72.13%	82.61%	41.18%	Information entropy (no unit)	30.3682
Confusion matrix				Attenuation coefficient (unit: dB/MHz-cm)	25.4943
True\Predicted	*< 5%*	*5%-33%*	*> 33%*	Integrated backscatter (unit: dB)	18.3958
< 5%	20	4	0	BMI	9.71624
5%-33%	7	11	3	AST	9.67458
> 33%	1	2	13	ALT	6.35082