AUTHOR QUERY FORM

AIP American Institute of Physics	**Journal: Med. Phys.** **Article Number: 022207MPH**	**Please provide your responses and any corrections by annotating this PDF and uploading it to AIP's eProof website as detailed in the Welcome email.**

Dear Author,

Below are the queries associated with your article. Please answer all of these queries before sending the proof back to AIP. Author please indicate the correct color processing option from the list below:

1. Author, please confirm Figure number(s) that should appear as color in print. Please know that any associated mandatory fees will apply for figures printed in color.
2. Author, please confirm Figure number(s) that should appear as color online only. There will be no fees applied.
3. Author, your paper currently does not include any color figures for online or print. If color is needed, please indicate which figures it should be applied to and whether it is color in print or online.

Location in article	Query / Remark: click on the Q link to navigate to the appropriate spot in the proof. There, insert your comments as a PDF annotation.
Q1	AU: Please provide complete affiliation details for the author "Ricardo Ribeiro."
Q2	AU: Please provide the zip code for the affiliations of authors "Ganapathy Krishnamurthi" and "Jasjit S. Suri" and postal codes for the affiliations of authors "U. Rajendra Acharya," "Ricardo Ribeiro," "Rui Tato Marinho," and "João Sanches."
Q3	AU: Please verify the edits made in the sentence beginning as "These random distributions. . . ."
Q4	AU: Please provide the publisher location for Refs. 2, 6, and 20.
Q5	AU: Please provide a digital object identifier (doi) for Ref(s) 9, 27, 33, 39, and 45. For additional information on doi's please select this link: http://www.doi.org/. If a doi is not available, no other information is needed from you.
Q6	AU: Please provide the publisher name and location and thus verify the edited details for Ref. 21 and 40.
Q7	AU: Please verify the volume number for Ref. 22.
Q8	AU: Please provide the city of the publisher for Ref. 25.
Q9	AU: Please provide the location for the proceedings as well as publisher name and location for Refs. 34, 35, 38, 42, and 43.

Thank you for your assistance.

Data mining framework for fatty liver disease classification in ultrasound: A hybrid feature extraction paradigm

U. Rajendra Acharya[a)]
Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore

S. Vinitha Sree[b)]
Global Biomedical Technologies Inc., California 95661

Ricardo Ribeiro[c)]
Institute for Systems and Robotics and Escola Superior de Tecnologia da Saúde de Lisboa

Ganapathy Krishnamurthi[d)]
Mayo Clinic, Rochester, Minnesota

Rui Tato Marinho[e)]
Liver Unit, Department of Gastroenterology and Hepatology, Hospital de Santa Maria, Medical School of Lisbon, Lisbon, Portugal

João Sanches[f)]
Institute for Systems and Robotics and Instituto Superior Técnico, Lisbon, Portugal

Jasjit S. Suri[g)]
Global Biomedical Technologies, California 95661 and Biomedical Engineering Department, Idaho State University, Idaho

(Received 25 September 2011; revised 30 April 2012; accepted for publication 14 May 2012; published XX XX XXXX)

Purpose: Fatty liver disease (FLD) is an increasing prevalent disease that can be reversed if detected early. Ultrasound is the safest and ubiquitous method for identifying FLD. Since expert sonographers are required to accurately interpret the liver ultrasound images, lack of the same will result in interobserver variability. For more objective interpretation, high accuracy, and quick second opinions, computer aided diagnostic (CAD) techniques may be exploited. The purpose of this work is to develop one such CAD technique for accurate classification of normal livers and abnormal livers affected by FLD.
Methods: In this paper, the authors present a CAD technique (called Symtosis) that uses a novel combination of significant features based on the texture, wavelet transform, and higher order spectra of the liver ultrasound images in various supervised learning-based classifiers in order to determine parameters that classify normal and FLD-affected abnormal livers.
Results: On evaluating the proposed technique on a database of 58 abnormal and 42 normal liver ultrasound images, the authors were able to achieve a high classification accuracy of 93.3% using the decision tree classifier.
Conclusions: This high accuracy added to the completely automated classification procedure makes the authors' proposed technique highly suitable for clinical deployment and usage. © *2012 American Association of Physicists in Medicine*. [http://dx.doi.org/10.1118/1.4725759]

Key words: fatty liver disease, computer aided diagnostic technique, texture, higher order spectra, discrete wavelet transform

I. INTRODUCTION

Fatty liver disease (FLD) or hepatic steatosis is a condition which is characterized by the presence of vacuoles of triglyceride fat in liver cells. This accumulation of fat happens through a process called steatosis in which there is abnormal retention of lipids in the cells. Some of the key causes of FLD are chronic alcohol consumption, obesity due to insulin resistance, and metabolic syndrome.[1,2] There are two major types of FLD based on the contribution of alcohol, namely, alcoholic steatosis and nonalcoholic fatty liver disease (NAFLD). NAFLD is progressively prevalent in Western countries and affects people of all ages and ethnicities.[3,4] Both alcoholic and nonalcoholic FLD, if left undetected and untreated, will progress to advanced liver diseases like inflammation (steatohepatitis), cirrhosis, and liver cancer. However, if found and treated early, FLD may be reversible. Therefore, early detection is of utmost importance in order to save patients from unwanted anxiety and also to reduce costs associated with providing treatments for advanced liver diseases. Liver biopsy is currently the standard for the assessment of steatosis. It is, however, invasive, uncomfortable, and prone to sampling errors.[5–7] The noninvasive techniques include ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI). Even though these methods have shown promise in detecting fatty infiltration in the liver, they

are insensitive in detecting steatosis of less than 25%–30%.[8] In the case of ultrasound, FLD causes increased echogenicity on ultrasound causing the liver to appear brighter than the cortex of the ipsilateral kidney.[9] Ultrasound has a sensitivity of around 82%–94% and specificity greater than 82% for detecting a fatty liver.[10–13] Even though ultrasound is more sensitive than CT,[14] it is less specific and also has poor visualization in obese patients.

In the case of CT imaging, hepatic steatosis produces an attenuation that is lower for the hepatic parenchyma than for the surrounding blood vessels, spleen, and the kidneys, thereby enabling the visualization of the presence of the steatotic liver. CT is limited by the calibrations required for different scanners and interobserver variabilities. MRI presents anatomical information of the imaged liver and magnetic resonance spectroscopy (MRS) provides a biochemical component.[15] MR has demonstrated that it can detect small fat fractions of less than 33%.[16] However, MR techniques, especially MRS, can be technically challenging. A good MRS of the liver requires good spatial resolution, high SNR, and adequate compensation for or elimination of patient motion.

Among all these modalities, ultrasound is the most commonly used modality due to its widespread availability in current clinical practice. In order to improve the specificity of ultrasound and also to address the interobserver variability issue that is common in medical image interpretation, computer aided diagnostic (CAD) or data mining techniques can be developed to more objectively and accurately detect the presence or absence of FLD in ultrasound images of liver. These techniques use the acquired ultrasound images to extract meaningful and discriminative features that are capable of adequately distinguishing a normal liver from an abnormal liver that is affected by FLD. These features are fed to supervised learning-based classifiers to train the classifiers to enable them to learn the parameters that effectively differentiate the patterns belonging to either of the classes. Thereafter, these learned parameters are used to classify new liver images into normal and abnormal categories. FLD affects the entire liver or a lobe of the liver, and hence, causes changes in the texture of the liver in the B-mode ultrasound (US) images. As indicated earlier, due to failure in fat metabolism, there is an increase in the deposition of fat in the liver which in turn gives rise to an increased brightness in the ultrasound and results in changes in the image texture.[2] Therefore, texture of the image has been studied as one of the key distinguishing features in this work. We have also extracted higher order spectral (HOS) features that characterize the nonlinearity in the images and discrete wavelet transform (DWT)-based features that quantify the changes occurring in the time-frequency domain of the images. We have demonstrated that a combination of these three types of features which extract significant information from the liver images is capable of discriminating normal and abnormal liver images with high classification accuracy. These features have been briefed in Sec. II.B.

The block diagram of the proposed CAD technique, named Symtosis, is shown in Fig. 1. In the offline training system, the acquired ultrasound liver images in the training set are preprocessed, and three sets of features, namely, HOS, texture, and

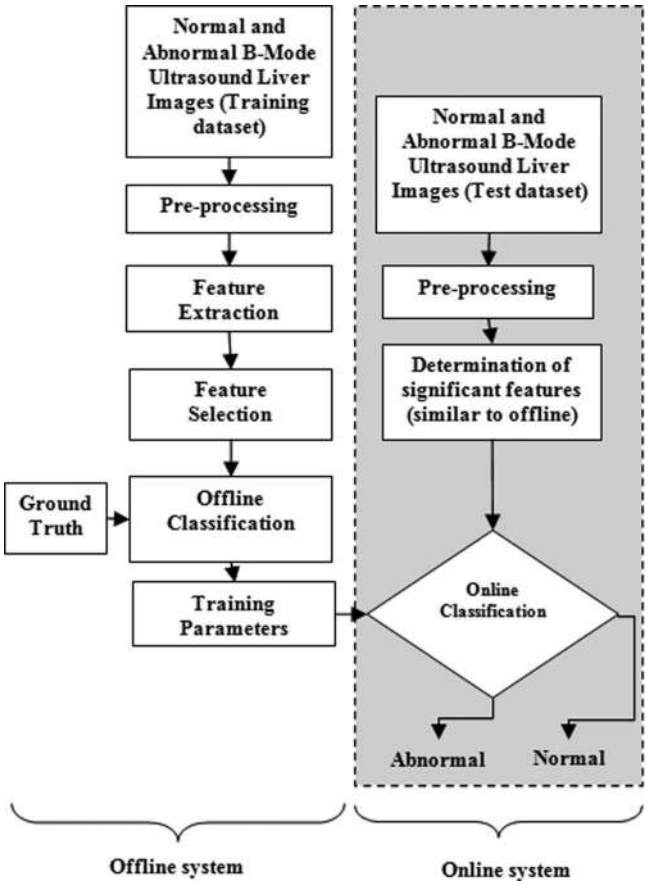

FIG. 1. Block diagram of the proposed Symtosis system for fatty liver disease detection; the blocks outside the dotted shaded rectangular box represent the flow of offline training system, and the blocks within the dotted box represent the online real-time system.

DWT, are extracted from the images in the *Feature Extraction* step. In the *Feature Selection* step, in order to reduce the dimensionality of the extracted feature set and to select only unique and highly discriminating features, the extracted features are subjected to the Student's *t-test* and only significant features are selected to form the final feature set. During *Offline Classification*, the significant feature set and the ground truth of whether the images belong to normal or abnormal cases (as predicted by doctors or by lab results) are used as inputs to several supervised learning-based classifiers in order to train them to determine appropriate parameters for differentiating both classes based on the features. The obtained *Training Parameters* are the output of the offline training system. In the online real-time system, which is the one that will be used by the end-user, the test images are preprocessed and the features reported as significant by the offline system are calculated from the test images. Subsequently, in the *Online Classification* step, the training parameters from the offline system are used on the calculated features to determine the class of the images. The resultant class labels are used to determine five performance measures, namely, accuracy, sensitivity, specificity, and positive predictive value (PPV), and area under the receiver operating characteristic curve (AUC). Definitions of these measures are given in Sec. II.D.

The key contributions of this paper are the (a) development of a completely automated CAD technique for detection of FLD in ultrasound liver images and (b) determination of a powerful combination of highly representative features for achieving high accuracy for automatic classification of the liver disease.

The paper is organized as follows. In Sec. II, we describe the data acquisition, extracted features, statistical techniques, and the classifiers. In Sec. III, we present the significant features and classification results. Section IV summarizes key findings in related studies in literature and compares and discusses the results obtained in this work. We conclude the paper in Sec. V.

II. MATERIALS AND METHODS

In this section, we describe the data used in this work, and present brief descriptions of the features extracted, statistical techniques used, and the classifiers evaluated. The MATLAB software was used for coding and analysis in this work.

II.A. Patient data

One hundred ultrasound liver images were used for classifier development and evaluation in this work. Among these 100 cases, 58 were abnormal (affected by FLD) and 42 were normal images. The ultrasound images of normal and fatty livers were acquired by expert operators with the ultrasound equipment in a hospital facility. All the images were collected from routine cases and were consecutively recruited. No challenges were faced during patient recruitment. The ultrasound images were obtained by a Philips CX© 50 ultrasound machine. All images were captured with 1024 × 1024 pixels with a gray level resolution of 8 bits/pixel. Images were stored in the Digital Imaging and Communications in Medicine (DICOM) format. The default computer interface given by the manufacturer was used for the input of patient data and further ultrasound image acquisition. The broadband curved array transducer C5-1 from Philips© was used. It is composed by 160 piezoelectric elements with a curved array shape, and had the operating frequency range from 1 to 5 MHz.

A calibration procedure and an ultrasound machine settings preset were developed before the data collection phase, in order to obtain reproducible results. To perform the calibration and consequent preset, images from 20 normal liver patients, according to the laboratorial analysis and with a body mass index (BMI) within the normal range (18.5–24.9), were scanned. Different imaging conditions were trained, mainly transducer frequency, gain, time gain compensation (TGC), dynamic range, focus, and depth. Using this procedure it is believed that standardization in all image acquisition is achieved, as reported by Kadah *et al.*[17]

The established ultrasound machine preset for this study, after the calibration step, was set by using a fundamental frequency of 3.5 MHz, an image depth of 15 cm, and two focal zones were used and set at the central portion of the image (7.5 cm). The dynamic range was set at 70 dB and the gain was variable, according to the patient biotype. TGC was set to its central position and kept constant through the examinations, eliminating this variable parameter. Acquiring US images only from the right liver lobe also allowed standardizing the acquisition protocol. According to patient biotype, different transducer orientation angles were performed, using as protocol the same liver anatomical landmarks. Patients were positioned in supine, comfortable, and asked to breathe gently, avoiding major patient motion.

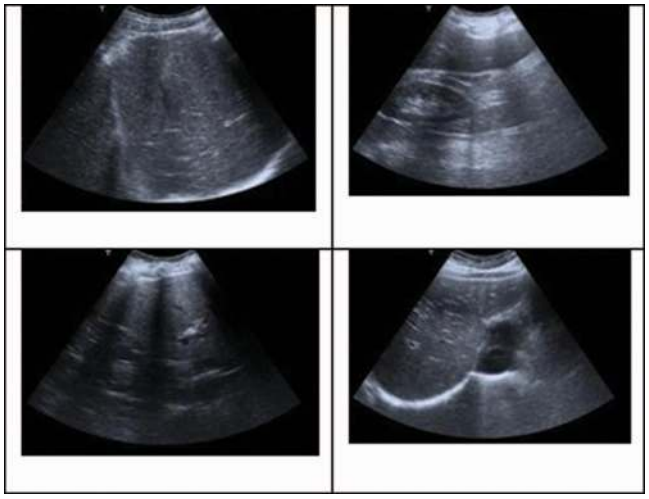

FIG. 2. Normal liver images (left column) and abnormal liver images (right column).

The ground truth as to whether each image was normal or abnormal was determined manually by the operators and confirmed by indicators obtained from laboratory analysis. A region of interest (ROI) of 128 × 128 pixels along the medial axis was extracted from each image. Typical images of normal and abnormal liver are shown in Fig. 2.

II.B. Grayscale feature extraction

II.B.1. HOS-based features

Higher order spectra-based features quantify the nonlinear behavior of a process.[18] Pixels in the ultrasound images are very randomly distributed with possible nonlinear interactions among the frequency components and perhaps some form of phase coupling. These random distributions cannot be fully described by second-order measures, but the HOS features are capable of capturing these distributions. They are useful in detecting nonlinear coupling and deviation from Gaussianity, and features derived from HOS can be made invariant to shift, rotation, and amplification. The HOS of Gaussian signals are statistically zero thus making HOS more robust to Gaussian noise.[19] Therefore, we have chosen HOS as one of the key features for quantifying the subtle changes in the normal and abnormal images.

Higher order statistics deal with higher order moments (order greater than two) and nonlinear combinations of these higher order moments, called the higher order cumulants. The bispectrum, which is the spectrum of the third order cumulants, is one of the most commonly used HOS features. Prior to the calculation of the bispectrum, the preprocessed images

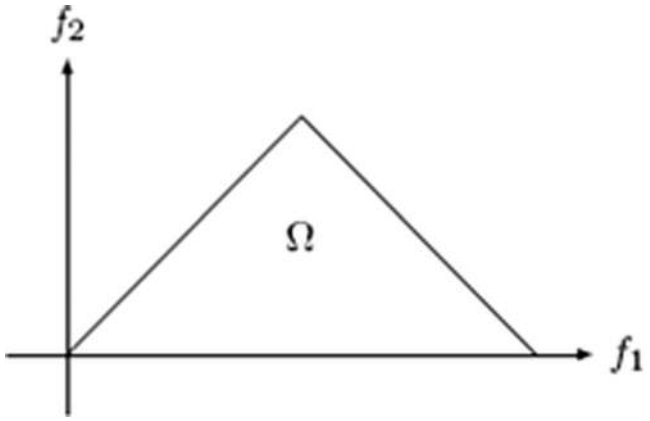

FIG. 3. Principal domain region (Ω) used for the computation of the bispectrum for real signals.

were first subjected to Radon transform.[20] This transform determines the line integrals along many parallel paths in the image from different angles θ by rotating the image around its center. Hence, the intensities of the pixels along these lines are projected into points in the resultant transformed signal. Thus, the Radon transform converts a 2D image into a 1D signal at various angles. This 1D signal is then used to determined the bispectrum, which is a complex valued function of two frequencies f_1 and f_2 given by

$$B(f_1, f_2) = E[X(f_1)X(f_2)X^*(f_1 + f_2)], \tag{1}$$

where $X(f)$ is the Fourier transform of the signal studied, $E[.]$ stands for the expectation operation, and * stands for the conjugate operator. As per the equation, the bispectrum is the product of the three Fourier coefficients. The function exhibits symmetry, and is computed in the nonredundant/principal domain region Ω as shown in Fig. 3.

The bispectrum phase entropy [21–23] obtained from the bispectrum is used as one of the features in this work. This bispectrum phase entropy (ePRes) is defined as

$$\text{eP Res} = \sum\nolimits_n p(\psi_n)\log(p(\psi_n)), \tag{2}$$

where

$$p(\psi_n) = \frac{1}{L}\sum_{\Omega} l(\phi(B(f_1, f_2)) \in \psi_n), \tag{3}$$

$$\psi_n = \{\phi | -\pi + 2\pi n/N \le \phi < -\pi + 2\pi(n+1)/N\},$$
$$n = 0, 1, \ldots, N-1, \tag{4}$$

where L is the number of points within the region Ω, ϕ is the phase angle of the bispectrum, and $l(.)$ is an indicator function which gives a value of 1 when the phase angle is within the range depicted by ψ_n in Eq. (4). In this work, we calculated the Radon transformed signals for every 1° step size and then determined the phase entropy of these signals. Entropies are generally used to characterize the regularity or irregularity of the pixels in the image. If the resulting Radon transformed signal obtained from the liver image at a particular angle is perfectly periodic and predictable, then the consequent phase entropy would be zero. As the signal becomes more random, the entropy increases.[24] In this work, it was observed that the normal images had more randomness than the abnormal images (Sec. III.A).

II.B.2. Texture-based features

The presence of various granular structures in the liver ultrasound images makes the use of image texture analysis techniques suitable for liver image classification. In most image processing applications, assumptions are made regarding the uniformity of gray-level intensity values in the image. In real applications, most images have a variation in gray levels which are repetitive and these variations are characterized as the texture of the image.[25] The most commonly used texture matrices are the gray level co-occurrence matrix (GLCM) and the run length matrix. We have calculated one homogeneity feature from the GLCM (Ref. 26) and three features from the run length matrix.[27] These features are described briefly below.

Texture homogeneity: The gray level co-occurrence matrix of an image of size $m \times n$ is defined as follows:

$$C_d(i, j) = \left| \left\{ \begin{array}{c} (p, q), (p + \Delta x, q + \Delta y) : I(p, q) = i \\ I(p + \Delta x, q + \Delta y) = j \end{array} \right\} \right|, \tag{5}$$

where (p, q), $(p + \Delta x, q + \Delta y)$ belong to $m \times n$, $d = (\Delta x, \Delta y)$, and $|\ldots|$ denotes the set cardinality. The probability of a pixel with a gray level intensity value i having a pixel with a gray level intensity value j at a distance $(\Delta x, \Delta y)$ away in an image is defined as

$$P_d(i, j) = \frac{C_d(i, j)}{\sum_{<i>}\sum_{<j>} C_d(i, j)}. \tag{6}$$

The homogeneity of the image is now defined as

$$C_h = \sum_i \sum_j \left(\frac{P_d(i, j)}{1 + |i - j|} \right). \tag{7}$$

The homogeneity measures the closeness of the distribution of the co-occurrence matrix elements to the main diagonal. A homogenous image will give rise to a $P_d(i, j)$ clustered around the main diagonal. In other words, the similarity between two pixels that are $(\Delta x, \Delta y)$ apart is measured by the homogeneity feature.

Texture run percentage (TexRL): The run percentage is a texture property derived from the run length matrix of an image. The run length matrix P_θ contains all the elements, where the gray level value i has the run length j continuous in direction θ.[27] Often the direction θ is set as 0°, 45°, 90°, or 135°. The run percentage is defined as the total number of runs in the image divided by the total number of pixels in the image as depicted in Eq. (8):

$$\text{TexRL} = \frac{\sum_{i=1}^{Ng}\sum_{j=1}^{N_r} P_\theta(i, j)}{N_p}. \tag{8}$$

Run percentage has the lowest value for images with the most linear structure. Here, $P_\theta(i,j)$ is the element of the run length matrix, N_p is the total number of pixels in the image,

N_g is the number of gray levels in the image, and N_r is the number of different run lengths that occur.

Short run emphasis (SRE): Based on the run length matrix, the short run emphasis is defined as

$$\mathrm{SRE} = \frac{\sum_{i-1}^{N_g} \sum_{j=1}^{N_r} \frac{P_\theta(i,j)}{j^2}}{\sum_{i-1}^{N_g} \sum_{j=1}^{N_r} P_\theta(i,j)}, \tag{9}$$

where the index i runs over the gray level values in the image and the index j runs over the run length. Higher values of j, i.e., long run lengths will contribute less to the sum in Eq. (9) and consequently higher sum emphasizes short runs.

Gray level nonuniformity (GLNU): The gray level nonuniformity is defined as

$$\mathrm{GLNU} = \frac{\sum_{i=1}^{N_g} \left(\sum_{j=1}^{N_r} P_\theta(i,j) \right)^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P_\theta(i,j)}. \tag{10}$$

The gray level nonuniformity squares the run lengths for each gray value. Hence, longer run lengths will make significant contributions to the summation, i.e., uniform images will have higher values of this sum as compared to images that are nonuniform in their gray levels.

II.B.3. DWT-based features

A wavelet transform is the representation of a function by wavelets, which are scaled and translated copies of a basic wavelet shape called the "mother wavelet." Mother wavelets are functions that are localized in both time and frequency and have varying amplitudes during a limited time period and very low or zero amplitude outside that time period. Wavelet transforms such as continuous wavelet transform (CWT), DWT, and wavelet packet decomposition (DWT) determine a limited number of wavelets coefficients that adequately describe the image. Two-dimensional DWT was used in this work. DWT analyzes the image at different frequency bands with different resolutions by decomposing the image into coarse approximation and detail information. The approximation coefficients are obtained by passing the image through a low pass filter (LPF), and the detail coefficients are obtained by filtering the image using a high pass filter (HPF). This decomposition is done recursively on the low pass approximation coefficients obtained at each level until the desired number of iterations is reached.

An illustration of DWT is given in Fig. 4. The rows of the image I are convolved using a LPF and the columns of the convolved output are down-sampled, i.e., only the even indexed columns are retained for further filtering. Next, the down-sampled columns are passed through another LPF, the output of which is again sampled to keep the even indexed rows alone. These are the approximation coefficients cA_1 at level 1. Similarly, the down-sampled columns are passed through a HPF, sampled to retain the even indexed rows alone to get the horizontal detail coefficients cH_1. In a similar fashion, the rows of the image I are high passed filtered and processed through a set of low pass and high pass filters to get the vertical detail coefficients cV_1 and diagonal detail coefficients cD_1, respectively. In our work, we calculated the averages of each set of coefficients cA_1, cH_1, cV_1, and cD_1 at level 1, and again found the average of these individual averages. This overall average value was used as a feature.

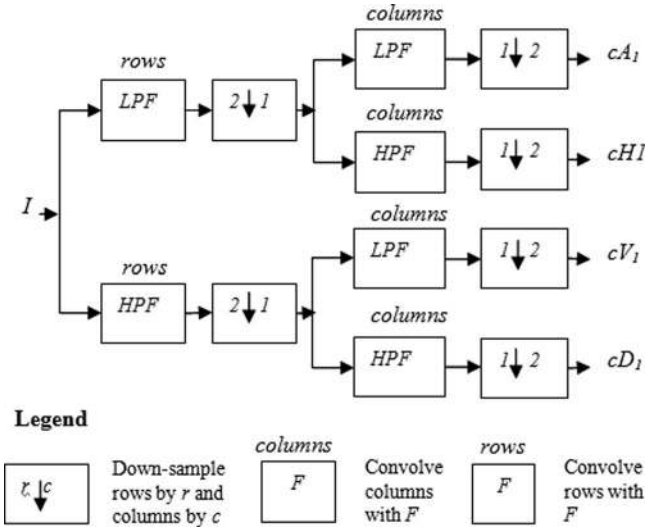

FIG. 4. DWT decomposition.

II.C. Classification paradigm in Symtosis system

Most of the supervised learning-based classifiers have a black box approach to determining the end results, i.e., the end-user would not be able to comprehend how the classifier determined the output class label from the input features. On the contrary, both decision tree (DT) and Fuzzy classifiers output feature-based rules for classifying future samples, and hence, are more comprehendible to the end-user. Medical practitioners, who are the end-users of such CAD-based diagnostic software, would prefer the classification protocol to be more transparent in order to have confidence in the output. Therefore, we chose these two classifiers in this work.

Decision Tree: In the case of DT, the input features are used to construct a tree, and then a set of rules for the different classes are derived from the tree. More details on how to construct a decision tree using features can be found in Refs. 28 and 29. The obtained rules are used to predict the class of a new data.

Fuzzy classifier: In the case of Fuzzy classifier, a subtractive clustering technique was used to generate a Fuzzy inference system (FIS).[30] The FIS structure contains if–then rules that specify a relationship between the input and output fuzzy sets. Each input and output has as many membership functions as the number of clusters. The clustering technique estimates the number of clusters and the cluster centers in the examined dataset. Radius parameter is used to indicate a cluster center's range of influence in each of the data dimensions. The determined is used to perform fuzzy inference

TABLE I. Mean ± standard deviation (SD) values of the significant features for the normal and abnormal classes using Symtosis system.

Features	Normal (mean ± SD)	Abnormal (mean ± SD)	p-value
SRE	$0.869 \pm 3.105 \times 10^{-2}$	$0.821 \pm 4.125 \times 10^{-2}$	<0.0001
ePRes(12°)	$4.770 \pm 3.993 \times 10^{-2}$	$4.504 \pm 6.623 \times 10^{-2}$	<0.0001
$\text{DWTMean1}_{\text{sym4}}$	19.1 ± 8.35	11.7 ± 5.06	<0.0001

calculations of the test data. In this work, we implemented a Sugeno-type fuzzy inference system.[31]

II.D. Statistical analysis

In order to select unique and highly discriminating features, the Student's *t-test* was used to select the features that were significantly different between the normal and abnormal cases. In this test, initially, for each feature, the null hypothesis is assumed to consider that the mean of the feature from the normal class is equal to the mean of the feature from the abnormal class. Subsequently, the *t-statistic*, which is the ratio of difference between the means of two classes to the standard error between class means, and the corresponding *p-value* are calculated. The *p-value* is the probability of rejecting the null hypothesis given that the null hypothesis is true. A low *p-value* (less than 0.01 or 0.05) indicates rejection of null hypothesis, which implies that the means are not equal in both classes and are significantly different, and hence, the feature is significant.

Sensitivity, specificity, positive predictive value, and accuracy were calculated to evaluate the performance of the classifiers. True negative (TN) is the number of normal samples identified as normal. True positive (TP) is the number of abnormal samples identified as abnormal. False negative (FN), on the other hand, is the number of abnormal samples identified as normal and False positive (FP) is the number of normal samples identified as abnormal. Sensitivity, which is the probability that a test will produce a positive result when used on abnormal population, is calculated as TP/(TP + FN) and specificity, which is the probability that a test will produce a negative result when used on normal disease-free population, is determined as TN/(TN + FP). PPV, which is the probability that the patient is abnormal when restricted to those patients who test positive, is calculated as TP/(TP + FP), and accuracy, which is the ratio of the number of correctly classified samples to the total number of samples, is calculated as (TP + FP)/(TP + FP + TN + FN).

Another important performance measure is the area under the receiver operating characteristic (ROC) curve, called AUC. The ROC curve is obtained by calculating the sensitivity and specificity of a classifier at different cut-off values and plotting sensitivity vs (1-specificity).[32] (1-specificity) is called the false positive rate (FPR). A classifier that perfectly discriminates between the two classes would yield a curve that coincides with the left and top sides of the plot. This means that sensitivity is high and the FPR is low. A classifier that is completely useless would give a straight line that follows a diagonal path from the bottom left corner to the top right corner. Generally, the curve will lie somewhere between these extremes because of the overlap of the values in the two classes. The goodness of a classifier is assessed by determining the AUC. For an ideal test, the AUC would be 1. For a useless classifier, which follows the diagonal ROC curve, the AUC would be 0.5 which is equivalent to having sensitivity and specificity of 0.5 (50%). Hence, in practice, the closer the AUC is to 1.0, the better the classifier is, and the closer the AUC is to 0.5, the worse the classifier is.[33]

III. RESULTS

III.A. Significant features

As shown in Table I, all the three selected features had statistically significant differences between the abnormal and normal classes, as indicated by the low *p-value* (<0.01). The table also presents the mean and standard deviation of all the features. In the case of HOS-based features, one phase entropy-based feature obtained for Radon transform angle $\theta = 12°$, denoted in Table I as ePRes(12°), was found to be significant. In the case of texture features, only the short run emphasis (SRE) was found to be significant. To obtain the DWT features, around 54 mother wavelets were studied to find the mean value of the level 1 coefficients. Among them, the mean of the coefficients obtained at level one of decomposition using the sym4 mother wavelet was found to be significantly different between the two classes. In the case of abnormal images, all the features have registered lower values compared to that of the normal cases.

III.B. Symtosis classification results

In view of the low sample size, threefold stratified cross-validation was employed to obtain robust classifiers. In this resampling technique, the entire dataset is randomly split into three equal parts, each part containing the same proportion of samples from both the classes. No image is repeated in any of the parts. In the first fold, two parts of the data are used for training the classifier, and the remaining one part is used for testing the trained classifier and to obtain the performance measures. This procedure is repeated twice, using a new test set each time. The average of the performance measures obtained during each fold is taken to be the final values of the performance measures. To be specific, ~10 normal and 15 abnormal cases are used in each fold. Classification accuracy, sensitivity, specificity, PPV, and AUC were used as the

TABLE II. Symtosis classification results (the listed values are average of values obtained in the three folds) TN: true negatives, FN: false negatives, TP: true positives, FP: false positives, A: accuracy, PPV: positive predictive value, Sn: sensitivity, Sp: specificity.

	TN	FN	TP	FP	A (%)	PPV (%)	Sn (%)	Sp (%)
Feature combination A: All features except the HOS feature ePRes(12°)								
DT	8	6	9	2	65.3	78.5	57.8	76.7
Fuzzy	9	4	11	1	77.3	88.8	71.1	86.7
Feature combination B: All features except the DWT feature $DWTMean1_{sym4}$								
DT	10	2	13	0	93.3	100	88.9	100
Fuzzy	9	2	13	1	86.7	92.7	84.4	90.0
Feature combination C: All features except the texture feature SRE								
DT	10	2	13	0	93.3	100	88.9	100
Fuzzy	9	3	12	1	84.0	94.4	77.8	93.3
Feature combination D: All features								
DT	10	2	13	0	93.3	100	88.9	100
Fuzzy	9	2	13	1	86.7	90.8	86.7	86.7

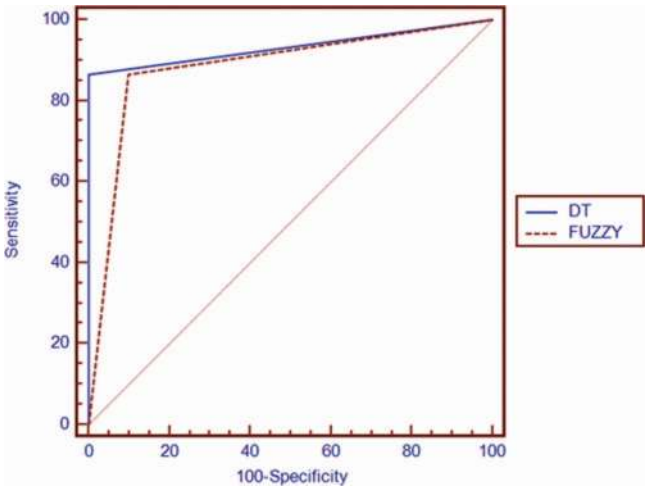

FIG. 5. ROC curves of the DT and Fuzzy classifiers using Symtosis.

performance measures to select the optimal classifier for this work. Table II presents the classification results obtained. In order to study the effect of each of the features on the performance measures, in Table II, we have presented the performance measures obtained using all features except the HOS feature ePRes(12°), measures obtained using all features except the DWT feature $DWTMean1_{sym4}$, measures obtained using all features except the texture feature SRE, and also those measures obtained using all the features.

The maximum accuracy that could be achieved using all the features except the HOS feature (ePRes(12°)) was only 65.3% using the DT classifier and 77.3% using the Fuzzy classifier (feature combination A in Table II). However, this accuracy increased significantly to 93.3% using DT classifier and 86.7% using the Fuzzy classifier on inclusion of the HOS feature during training (feature combination D in Table II). This significant increase in the accuracy demonstrates the capabilities of the HOS feature that were highlighted in Sec. II.B. The significant difference in the value of this phase entropy HOS feature for both classes of images (Table I) indicates that there are variations in the nonlinear dynamics in the image captured from a normal liver and that from a liver affected by FLD. The phase entropy feature has clearly captured these different nonlinear interactions in both the normal and abnormal liver images. The DWT feature ($DWTMean1_{sym4}$), on the other hand, did not have such a significant impact on the accuracy as evident from Table II. It can be seen that the performance measures obtained with and without the DWT feature are almost the same for both the classifiers (Feature Combinations B and D in Table II). A similar case was observed in case of the SRE feature (Feature Combinations C and D in Table II). Moreover, we also performed classification with the inclusion of four individual DWT coefficients (averages of each set of coefficients cA_1, cH_1, cV_1, and cD_1 at level 1) instead of using their average. The classification accuracy was still lower than 90% (results not shown in Table II). Furthermore, when we trained the classifiers with only the HOS feature, we obtained a low accuracy of around 64% for both classifiers (not shown in Table II). This indicates that either all three features or ePRes(12°) and $DWTMean1_{sym4}$ features or ePRes(12°) and SRE features should be used in the DT classifier to obtain the highest accuracy of 93.3%. This is because classifiers present different class separability based on the features input to them. From our experience, we inferred that the DT classifier provides good separability between the two classes with Feature Combinations B, C, and D in the table.

Moreover, the average AUC of the DT classifier was 0.933 and that of the Fuzzy classifier was 0.883. These values indicate the excellent performance of these classifiers. The ROC curves are depicted in Fig. 5.

IV. DISCUSSION

A few studies have been carried out to automatically classify diffuse liver diseases. We present a summary of these studies here and in Table III. Kyriacou *et al.*[34] used the texture feature algorithms such as fractal dimension texture analysis (FDTA), the spatial gray level dependence matrices (SGLDM), the gray level difference statistics (GLDS), the gray level run length statistics (RUNL), and first order gray level parameters (FOP) to classify three sets of ultrasound liver images, namely, fatty, cirrhosis, and normal (30 samples each). A ROI of 32 × 32 pixels in size was selected by an expert physician before feature extraction was done. The combination of FDTA and SGLDM features in a KNN classifier resulted in an accuracy of 82.2%. In another study by the same group,[35] they applied the algorithms on four sets of images, namely, normal, fatty, cirrhosis, and hepatoma. They obtained the highest accuracy of 80% using a combination of RUNL, SGLDM, and FDTA in the KNN classifier. On using a novel neural network classifier based on geometrical fuzzy sets, the same group[36] demonstrated an accuracy of 82.67% in classifying normal, fatty, and cirrhotic liver images.

In a study by Badawi *et al.*,[37] eight features, namely, the mean gray level, the percentile 10%, the contrast, the angular

TABLE III. Summary of studies that presented various CAD techniques for liver image classification.

Authors	Modality/classes	Features/classifier	Accuracy
Kyriacou *et al.*[34]	Ultrasound/fatty, cirrhosis, normal	11 texture features/KNN classifier	82.2%
Kyriacou *et al.*[35]	Ultrasound/fatty, cirrhosis, normal, hepatoma	10 texture features/KNN classifier	80.0%
Kyriacou *et al.*[36]	Ultrasound/fatty, cirrhosis, normal	12 texture feature algorithms/neural network classifier based on geometrical fuzzy sets	82.7%
Badawi *et al.*[37]	Ultrasound/normal, fatty, cirrhotic	8 numerical quantitative features from ultrasound/Fuzzy classifier	Accuracy not reported; Sensitivity: 96.0%
Wan and Zhou[38]	Ultrasound/normal, cirrhotic	32 wavelet packet transform-based features/SVM classifier	85.8%
Lee *et al.*[39]	Ultrasound/normal, hepatoma, cirrhosis	Fractal feature vector based on M-band wavelet transform/hierarchical classifier	96.7%
Ribeiro and Sanches[40]	Ultrasound envelope RF image/normal, fatty	3 intensity and texture features/Bayes classifier	95%
Yeh *et al.*[41]	Ultrasound images of fresh human liver samples/steatosis and nonsteatosis	Gray-level concurrence and nonseparable wavelet transform/support vector machine classifier	90.5%
Mougiakakou *et al.*[42]	CT/normal, cyst, hemangioma, hepatocellular carcinoma	5 sets of texture-based features/multiple classifier system using five neural networks	93.8%
Lin.[44]	Age, blood tests/normal, liver disease	CART to detect presence of liver disease, CBR to diagnose the type of liver disease	90.0%
Lin[45]	Age, blood tests/normal, liver disease	ANN to detect presence of liver disease, CBR + AHP to diagnose the type of disease	94.6%
In this work	Ultrasound/normal, fatty	3 texture, wavelet transform and higher order spectra features; decision tree classifier	93.3%

second moment, the entropy, the correlation, the attenuation, and the speckle separation, were extracted from 140 ultrasound images belonging to either normal, fatty, and cirrhotic livers and fed to a fuzzy classifier. Ninety six percent sensitivity was obtained for classification of the fatty livers. These results were higher than those obtained by the same group on using other classifiers.[17] Wan and Zhou[38] extracted the mean and energy from the subimages obtained from wavelet packet transform applied images. Thirty two such features from 390 normal and 200 cirrhosis samples were used in a SVM classifier and an accuracy of 85.79% was obtained. Wavelet transform resulted in only 77.65% accuracy.

Lee *et al.*[39] classified normal, hepatoma, and cirrhosis ultrasound images using fractal feature vector based on M-band wavelet transform. Having tested their methodology using various classifiers, they observed that a hierarchical classifier was 96.7% accurate in the classifying normal and abnormal liver images. Ribeiro and Sanches[40] used original RF signal generated by the ultrasound probe, and used the resulting RF image to estimate a despeckled image from which one intensity feature was extracted and a speckle image from which two texture features were obtained. On evaluating the technique with 10 normal and 10 fatty samples, in a Bayes classifier, they obtained an accuracy of 95%. Yeh *et al.*[41] developed a CAD technique to determine the steatosis grade in high frequency ultrasound liver images of 19 samples obtained surgically. They extracted image features from gray-level concurrence and nonseparable wavelet transform and fed them to a support vector machine classifier. An accuracy of 90.5% was registered for the classification of steatosis and nonsteatosis samples. On evaluating Haralick's statistical texture features extracted from 76 normal and 24 fatty ultrasound liver images, two features, namely, maximum probability and uniformity were found to be highly significant.[42]

Mougiakakou *et al.*[43] have used CT liver images to classify normal liver, cyst, hemangioma, and hepatocellular carcinoma. They extracted several texture-based features from 147 ROIs and used genetic algorithm to select significant features. On classifying the samples using a system of five neural networks, they obtained 93.75% accuracy for the validation set and 90.63% for the test set. They also incorporated their algorithm in diagnosis software called DIAGNOSIS.[44]

An intelligent model that detects the presence of liver disease using classification and regression tree (CART) and classifies the type of liver disease in the detected cases using a case-based reasoning (CBR) technique was developed by Lin.[45] The model was developed using 340 samples and comparative study was done using 170 samples. It was found that CART had an accuracy of 92.94% in the detection of the presence of liver disease. A 90% diagnostic accuracy was registered by CBR in classifying the type of disease. They concluded that the CART rules can help the physician in liver disease detection, whereas CBR had the capability of retrieving the most similar case in the database in order to solve new cases. Lin and Chuang[46] developed a similar intelligent liver diagnosis model using artificial neural network (ANN) instead of CART for detecting the presence of liver disease and integrated analytic hierarchy process (AHP) with CBR for diagnosing the type of disease. Using 39 clinical features from 300 patients as inputs to a three-layer backpropagation ANN, 98.04% accuracy was obtained in detecting the presence or absence of liver disease. AHP integrated with CBR could detect the type of disease with 94.57% accuracy.

A review of the literature indicates that ultrasound is the most commonly used modality for liver imaging, and most data mining-based studies use the ultrasound images[34–40] to characterize the liver tissue. Hence, we used ultrasound images in our work. Among the ultrasound-based studies[34–40], it is evident that the accuracy obtained in Refs. 34–38 is not as high as what we have obtained in our work. The limitation of the work by Lee *et al.*[39] is that prior to the image analysis, the region of interest covering the liver parenchyma without major blood vessels, acoustic shadowing, or any type of distortion was chosen manually by a physician. Hence, the process is not completely automated unlike our work. In Ref. 40, all the features were derived from the images obtained from envelope RF images, whereas in our work we used the B-mode ultrasound images directly for feature extraction. This reduces the computational complexity of the algorithm. Moreover, in all these studies except Ref. 40, the proposed algorithms were for classifying normal livers from other abnormal classes like fatty, cirrhosis, and hepatoma. In our current work, Symtosis system is used for classification of normal vs fatty liver disease. We intend to extend our technique for other abnormal classes in our future studies. Two of the major advantages of the proposed technique over the other studies in the literature are the determination of the unique and promising combination of these three features for obtaining high classification accuracy and the demonstration of the powerful capabilities of the HOS feature in improving liver image classification accuracy. Moreover, to obtain robust classification accuracies, we have used threefold cross-validation technique unlike most studies in the literature which used hold-out technique that results in less robust performance measures.

In our Symtosis design work, on using all the features including the HOS feature, the DT classifier resulted in high accuracy of 93.3% and balanced sensitivity (88.9%) and specificity (100%) values. The classification results indicate that the classification accuracy is influenced not only by the choice of features (type and number) but also on the choice of the classifier. We believe that by adding more relevant features we can improve the overall performance of our classifier. In future, a larger dataset from a multiethnic population would be studied. A variety of texture features and WPT-based features would be analyzed to improve the accuracy. We also intend to incorporate, in future, the information about the aggressiveness of the disease in the abnormal cases in order to more clearly understand how the features discriminate the normal and abnormal cases. We plan to use the speckle images obtained from the envelope radio frequency (RF) images to extract the features to investigate if the accuracy may be improved further at the expense of a slighter higher computational cost.

In spite of these limitations, the following are the key features of the proposed Symtosis CAD-based technique. (a) The technique is fully automated and does not require any segmentation to select the region of interest. Traditional ultrasound liver images are the only input required. (b) The novelty of the work lies in the fact that this is the first study that has exploited the HOS features and the combination of three features for FLD detection. We have demonstrated the utility and power of these features by evaluating the performance of the classifiers by training them without and with the HOS feature. It is evident that the accuracy significantly increased from 65.3% to 93.3% on including the HOS feature for classifier development and evaluation. (c) A high classification accuracy has been obtained (93.3%) with 100 samples. This emphasizes the discriminating capability of the significant features used. To account for the small sample size, we have employed the cross-validation data resampling technique in order to build robust classifiers. (d) The high classification accuracy has been achieved using only three features, making the entire process computationally less complex and cost-effective. (e) No additional cost is needed to incorporate the built classifier into a physician's computer. Executable software can be written and it can be downloaded from the internet easily. (f) No expert training is necessary to operate the software. The user has to only input the acquired liver ultrasound image, and the software will output the class label.

V. CONCLUSIONS

In this paper, we explored the possibility of a CAD-based technique called Symtosis for the classification of normal and liver affected by fatty liver disease (abnormal cases). The combination of image texture, higher order spectra, and discrete wavelet transform-based features that were extracted from the liver ultrasound images was used for training the classifier. Among the extracted features, three highly discriminatory significant features alone were used to train and build two supervised learning-based classifiers. Using only three features, the DT classifier presented a high accuracy of 93.3%. The sensitivity and specificity were 88.9% and 100%, respectively. It can be seen that significant performance measures have been obtained using a considerably large dataset. Since the technique is fully automated and highly user friendly, it can be easily used in clinical practice. We believe that with the inclusion of more representative features, it should be possible to improve the current accuracy of the technique. In future, we intend to evaluate the proposed technique using a larger dataset containing images from different patients acquired by different operators and containing images belonging to various pathologies.

[a]Electronic mail: aru@np.edu.sg
[b]Author to whom correspondence should be addressed. Electronic mail: vinitha.sree@gmail.com
[c]Electronic mail: ricardo.s.t.ribeiro@gmail.com
[d]Electronic mail: ganapathy.krishnamurthi@gmail.com
[e]Electronic mail: rui.marinho@mail.telepac.pt
[f]Electronic mail: jmrs@isr.ist.utl.pt
[g]Electronic mail: jsuri@comcast.net

[1]S. Sherlock and J. Dooley, *Diseases of the Liver and Biliary System* (Blackwell Science, Malden, 2002).
[2]V. Droga and D. Rubens, *Ultrasound Secrets* (Hanley and Be[illegible], 2004). Q4
[3]J. E. Lavine and J. B. Schwimmer, "Nonalcoholic fatty liver disease in the pediatric population," Clin. Liver Dis. **8**(3), 549–558 (2004).
[4]G. C. Farrell and C. Z. Larter, "Nonalcoholic fatty liver disease: From steatosis to cirrhosis," Hepatology **43**(2), S99–S112 (2006).
[5]D. Joy, V. R. Thava, and B. B. Scott, "Diagnosis of fatty liver disease: Is biopsy necessary?," Eur. J. Gastroenterol. Hepatol. **15**(5), 539–543 (2003).

[6] R[illegible]heung, "Complications of Liver Biopsy. Gastrointestinal Emergencies," in *Gastrointestinal Emergencies*, edited by T. C. K. Tham, J. S. A. Collins, and R. Soetikno (Blackwell, ■, 2009), pp. 72–79.

[7] V. Ratziu, F. Charlotte, A. Heurtier, S. Gombert, P. Giral, E. Bruckert, A. Grimaldi, F. Capron, T. Poynard, and LIDO Study Group, "Sampling variability of liver biopsy in nonalcoholic fatty liver disease," Gastroenterology **128**(7), 1898–1906 (2005).

[8] S. Saadeh, Z. M. Younossi, E. M. Remer, T. Gramlich, J. P. Ong, M. Hurley, K. D. Mullen, J. N. Cooper, and M. J. Sheridan, "The utility of radiological imaging in nonalcoholic fatty liver disease," Gastroenterology **123**(3), 745–750 (2002).

[9] S[illegible] Quinn and B. B. Gosink, "Characteristic sonographic signs of hepatic fatty infiltration," AJR, Am. J. Roentgenol. **145**(4), 753–755 (1985).

[10] K. J. Foster, A. H. Griffith, K. Dewbury, C. P. Price, and R. Wright, "Liver disease in patients with diabetes mellitus," Postgrad. Med. J. **56**(661), 767–772 (1980).

[11] Y. Yajima, K. Ohta, T. Narui, R. Abe, H. Suzuki, and M. Ohtsuki, "Ultrasonographical diagnosis of fatty liver: Significance of the liver-kidney contrast," Tohoku. J. Exp. Med. **139**(1), 43–50 (1983).

[12] S. H. Saverymuttu, A. E. Joseph, and J. D. Maxwell, "Ultrasound scanning in the detection of hepatic fibrosis and steatosis," Br. Med. J. (Clin. Res Ed) **292**(6512), 13–15 (1986).

[13] U. L. Mathiesen, L. E. Franzen, H. Aselius, M. Resjö, L. Jacobsson, U. Foberg, A. Frydén, and G. Bodemar, "Increased liver echogenicity at ultrasound examination reflects degree of steatosis but not of fibrosis in asymptomatic patients with mild/moderate abnormalities of liver transaminases," Dig. Liver Dis. **34**(7), 516–522 (2002).

[14] M. H. Mendler, P. Bouillet, A. Le Sidaner, E. Lavoine, F. Labrousse, D. Sautereau, and B. Pillegand, "Dual-energy CT in the diagnosis and quantification of fatty liver: Limited clinical value in comparison to ultrasound scan and single-energy CT, with special reference to iron overload," J. Hepatol. **28**(5), 785–794 (1998).

[15] S. R. Mehta, E. L. Thomas, J. D. Bell, D. G. Johnston, and S. D. Taylor-Robinson, "Non-invasive means of measuring hepatic fat content," World J. Gastroenterol. **14**(22), 3476–3483 (2008).

[16] A. Qayyum, "MR spectroscopy of the liver: Principles and clinical applications," Radiographics **29**, 1653–1664 (2009).

[17] Y. M. Kadah, A. A. Farag, J. M. Zurada, A. M. Badawi, and A. M. Youssef, "Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images," IEEE Trans. Med. Imaging **15**(4), 466–478 (1996).

[18] C. Nikias and A. Petropulu, *Higher-Order Spectral Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1997).

[19] K. C. Chua, V. Chandran, U. R. Acharya, and C. M. Lim, "Analysis of epileptic EEG signals using higher order spectra," J. Med. Eng. Technol. **33**(1), 42–50 (2009).

[20] A. [illegible] and A. Katsevich, *The Radon Transform and Local Tomography* (CRC, ■, 1996).

[21] K. C. Chua, V. Chandran, R. Acharya, and C. M. Lim, "Automatic identificatio[illegible]pilepsy by HOS and power spectrum parameters using EEG signals: [illegible]parative study," in *Conference Proceedings of the IEEE on Engineering in Medicine and Biology Society*, Vancouver, British Columbia (■, ■, 2008), pp. 3824–3827.

[22] K. C. Chu[illegible]Chandran, U. R. Acharya, and C. Lim, "Application of higher orde[illegible]pectra to identify epileptic EEG," J. Med. Syst. **35**, 1–9 (2010).

[23] O. Faust, U. R. Acharya, C. M. Lim, and B. H. C. Sputh, "Automatic identification of epileptic and background EEG signals using frequency domain parameters," Int. J. Neural Syst. **20**(2), 159–176 (2010).

[24] K. C. Chua, V. Chandran, U. R. Acharya, and C. M. Lim, "Cardiac state diagnosis using higher order spectra of heart rate variability," J. Med. Eng. Technol. **32**(27), 145–155 (2008).

[25] M. Mirm[illegible]K. Xie, and J. S. Suri, *Handbook of Texture Analysis* (Imperial Colle[illegible]ss, ■, 2009).

[26] J.-H. Tan, E. Ng, U. R. Acharya, and C. Chee, "Study of normal ocular thermogram using textural parameters," Infrared Phys. Technol. **53**(2), 120–126 (2010).

[27] M. M[illegible]loway, "Texture analysis using grey level run lengths," NASA STI/[illegible] Technical Report No. N 75, 1974.

[28] I. M. Kapetanovic, S. Rosenfeld, and G. Izmirlian, "Overview of commonly used bioinformatics methods and their applications," Ann. N.Y. Acad. Sci. **1020**, 10–21 (2004).

[29] D. T. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining* (Wiley Interscience, New Jersey, 2004), pp. 107–126.

[30] T. J. Ross, *Fuzzy Logic with Engineering Applications* (Wiley, West Sussex, 2004).

[31] M. Sugeno, *Industrial Applications of Fuzzy Control* (Elsevier Science, North-Holland, 1985).

[32] C. E. Metz, "Basic principles of ROC analysis," Semin. Nucl. Med. **8**(4), 283–298 (1978).

[33] M. H. Z[illegible]and G. Campbell, "Receiver-operating characteristic (ROC) plots: A[illegible]amental evaluation tool in clinical medicine," Clin. Chem. **39**(4), 561–577 (1993).

[34] E. Kyriacou, S. Pavlopoulos, G. Konnis, D. Koutsouris, P. Zoumpoulis, and I. Th[illegible]s, "Computer assisted characterization of diffused liver disease using image texture analysis techniques on B-scan images," in *Proceedings of the IEEE Nuclear Science Symposium* (■, ■, 1997), Vol. 2, pp. 1479–1483.

[35] E. Kyriacou, S. Pavlopoulos, D. Koutsouris, P. Zoumpoulis, and L. The[illegible], "Computer assisted characterization of liver tissue using image [illegible] analysis techniques on B-scan images," in *Proceedings of the 19th International Conference of IEEE EMBS* (■, ■, 1997), Vol. 2, pp. 806–809.

[36] S. Pavlopoulos, E. Kyriacou, D. Koutsouris, K. Blekas, A. Stafylopatis, and P. Zoumpoulis, "Fuzzy neural network-based texture analysis of ultrasonic images," IEEE Eng. Med. Biol. Mag. **19**(1), 39–47 (2000).

[37] A. M. Badawi, A. S. Derbala, and A. M. Youssef, "Fuzzy logic algorithm for quantitative tissue characterization of diffuse liver diseases from ultrasound images," Int. J. Med. Inf. **55**(2), 135–147 (1999).

[38] J. Wan and S. Zhou, "[illegible]atures extraction based on wavelet packet transform for B-mode ultras[illegible]iver images," in *the 3rd International Congress Image Signal Proceedings (CISP)* (■, ■, 2010), Vol. 2, pp. 949–955.

[39] W. L. L[illegible]Y. C. Chen, and K. S. Hsieh, "Ultrasonic liver tissues classification by [illegible] feature vector based on M-Band wavelet transform," IEEE Trans. Med. Imaging **22**(3), 382–392 (2003).

[40] R. Ribeiro and J. Sanches, "Fatty liver characterization and classification by ul[illegible]nd," in *Proceedings of the 4th Iberian Conference on Pattern Recognition And Image Analysis*, Lecture Notes in Computer Science Vol. 5524 (■, ■, 2009), pp. 354–361.

[41] W. C. Yeh, Y. M. Jeng, C. H. Li, P. H. Lee, and P. C. Li, "Liver steatosis classification using high-frequency ultrasound," Ultrasound Med Biol. **31**(5), 599–605 (2005).

[42] S. Mukherjee, A. Chakravorty, K. Ghosh, M. Roy, A. Adhikari, and S. Mazumda[illegible]"C[illegible]rroborating the subjective classification of ultrasound images of n[illegible]and fatty human livers by the radiologist through texture analysis and SOM," in *Proceedings of the International Conference on Advanced Computing and Communications* (■, ■, 2007), pp. 197–202.

[43] S. G. Mougiakakou, I. K. Valavanis, K. S. Nikita, A. Nikita, and D. Kelekis, "Characterizati[illegible]CT liver lesions based on texture features and a multiple neural netw[illegible]assification scheme," in *Proceedings of the 25th International Conference of IEEE EMBS* (■, ■, 2003), Vol. 2, pp. 1287–1290.

[44] S. G. Mougiakakou, I. K. Valavanis, N. A. Mouravliansky, A. Nikita, and K. S. Nikita, "DIAGNOSIS: A telematics-enabled system for medical image archiving, management, and diagnosis assistance," IEEE Trans. Instrum. Meas. **58**(7), 2113–2120 (2009).

[45] R. H. [illegible]n intelligent model for liver disease diagnosis," Artif. Intell. Med. 4[illegible]53–62 (2005).

[46] R. H. Lin and C. L. Chuang, "A hybrid diagnosis model for determining the types of the liver disease," Comput. Biol. Med. **40**(7), 665–670 (2010).