Classification and Staging of Chronic Liver Disease based on Ultrasound, Laboratorial and Clinical Data

Ricardo Ribeiro, Rui T. Marinho, Jasjit S. Suri, and J. Miguel Sanches

Abstract Chronic Liver Disease is a progressive disease, most of the time asymptomatic, and potentially fatal. In this chapter, an automatic procedure to stage the disease is proposed based on ultrasound liver images, clinical and laboratorial data. A new hierarchical classification and feature selection approach, inspired in the current diagnosis procedure used in the clinical practice, here called *Clinical Based Classifier*, is described. The classification procedure follows the well established strategy of liver disease *differential diagnosis*. The decisions are taken with different classifiers by using different features optimized to the particular task for which they were designed. It is shown that the *Clinical Based Classifier* method outperforms the traditional *one against all* method because it take into account the natural evolution of the hepatic disease. Different specific features are used to detect and classify different stages of the liver disease as it happens in the classical diagnosis performed by the medical doctors.
The proposed method uses multi-modal features, extracted from ultrasound images, laboratorial and clinical data, that are known to be more appropriated according the disease stage we want to detect. Therefore, a battery of classifiers and features are optimized and used in a hierarchical approach in order to increase the accuracy of the classifier.

Ricardo Ribeiro
Institute for Systems and Robotics and Department of Bioengineering from Instituto Superior Técnico / Technical University of Lisbon, Portugal; Escola Superior de Tecnologia da Saúde de Lisboa, Portugal e-mail: ricardo.ribeiro@estesl.ipl.pt
Rui T. Marinho
Liver Unit, Department of Gastroenterology and Hepatology, Hospital de Santa Maria, Medical School of Lisbon, Portugal. e-mail: rui.marinho@mail.telepac.pt
Jasjit S. Suri
Biomedical Technologies Inc., Denver, CO, USA; Idaho State University (Aff.), ID, USA
J. Miguel Sanches
Institute for Systems and Robotics and Department of Bioengineering from Instituto Superior Técnico / Technical University of Lisbon, Portugal e-mail: jmrs@ist.utl.pt

For the normal class we achieved 100% accuracy, for the chronic hepatitis 69.2%, for compensated cirrhosis 81.48% and for decompensated cirrhosis 91.7%.

1 Introduction

Chronic liver disease (CLD) is a significant cause of morbidity and mortality in developed countries and commonly is caused by viral hepatitis and alcohol abuse [1].

The initial stages of CLD are usually asymptomatic such as *steatosis* or *hepatitis*. *Hepatitis* is the inflammation of the liver, resulting in liver cell damage and destruction[1]. It is caused by hepatitis viruses, which can have several types, or by other factors, e.g. alcohol. Moreover the natural evolution of the disease may lead to *cirrhosis* or even *hepatocellular carcinoma*, which are more severe pathological conditions, with high morbidity and mortality. *Cirrhosis* is a chronic disease that is characterized anatomically by widespread nodules in the liver combined with fibrosis [2]. It is possible to distinguish two phases in cirrhosis, a stable form, called *compensated cirrhosis*, and a more dangerous form that could lead to widespread failure of the liver, called *decompensated cirrhosis* [3].

Liver biopsy has been the preferred tool in the evaluation and staging of the CLD. However, its invasive nature and the development of other more accurate non-invasive alternatives have lead to a decreasing on its usage for assess the CLD. Among these alternatives, CLD staging based on ultrasound (US) data has proven to be a promising and safer alternative to biopsy.

In the review study presented in [1] it is shown that echogenicity, texture characterization and surface morphology of the liver parenchyma are effective features to diagnose the CLD. However, the evaluation of these features is normally affected by the subjective assessment of the human operator. This factor may lead to significant errors in the diagnosis and staging of CLD, since US liver images can show great variability, as shown in Fig. 1. Therefore, new objective feature extraction and classification methodologies in a *Computer Assisted Diagnosis* framework are needed.

Several studies presented in the literature use objective features, extracted from US images, and propose classification procedures to assess CLD [4]. Some of the most common features are based on the first order statistics, co-occurrence matrix, wavelet transform, attenuation and backscattering parameters and coefficients. A brief description of some of these studies is given next.

In [5], an experimental study was performed aiming at to discriminate the liver fibrosis from ultrasound images. They computed fractal features, entropy measures and co-occurrence information from ultrasound images to characterize the liver parenchyma form a textural point of view and the classification results showed an overall accuracy (OA) of 85.2% by using a Fisher linear classifier. Other important work described in [6] shows the ability of the Wavelet coefficients, also computed from US images, to characterize the diffuse disease of the liver. Their goal was to discriminate normal, steatotic and cirrhotic conditions. An OA of 90% is obtained and comparison results by using other classes of features, such as co-occurrence information, Fourier descriptors and fractal measures, show that the wavelet based classifier outperforms the classifiers based on the other features, 87%, 82% and 69%, respectively.

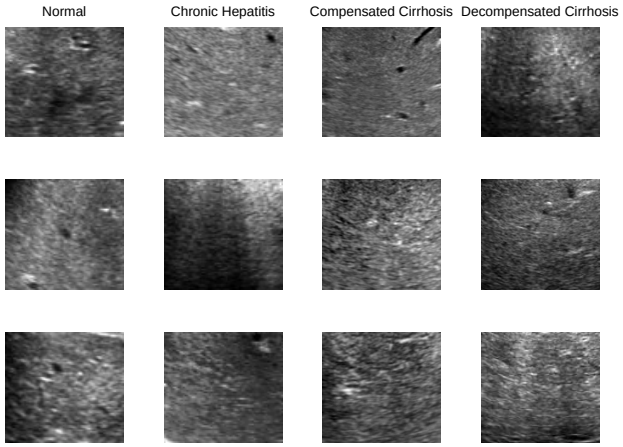

Fig. 1 Ultrasound images variability in the different stages of Chronic Liver Disease

[7] categorized patient in normal (72), fatty liver (66) and chronic liver disease (64), in order to evaluated the usefulness of standard deviation to measure the homogeneity of hepatic parenchyma from based on Ultrasound images. They observe significant differences ($p < 0.0001$) between the chronic liver disease group and the normal and fatty liver groups. They also concluded that higher average standard deviation values are related to wide distribution of intensity values within the ROI, as reported for CLD group, which explains the characteristic appearance of heterogeneous echo texture in CLD groups, such as chronic hepatitis and liver cirrhosis. Depict the good results obtained, the authors suggest careful in the use of this feature, since it is highly dependent on the ROI location.

Two main contributions for the CLD assessment are presented in this work; i) multi-modal features, extracted form US images, laboratorial and clinical data and ii) a new classification procedure inspired in the clinical practice, here called *Clinical Based Classifier* (CBC).

The discriminative power of the automatic classifier can be greatly increased if the natural evolution and staging of the disease is taken into account.

The remainder of this chapter is organized as follows: Section 2 introduces the pre-processing algorithm used, explaining the feature extraction and selection procedures, as well as the classifiers and the dataset used in this work. In Section 3 the results are presented showing the feature selection results and the classification results for each of the used classifier. The discussion of the results are presented in Section 4 and conclusions are presented in Section5.

2 Methods

The CBC aims at discriminate normal and three main pathologies in the CLD scope; i) *Chronic Hepatitis*, ii) *Compensated Cirrhosis* and iii) *Decompensated Cirrhosis*.

The diagnosis of these pathologies is performed in the today clinical practice based on several sources of medical data such as US liver parenchyma images, laboratorial exams and clinical indicators recommended in well established and accepted medical guidelines [3]. The diagnosis, however, is obtained by integrating all information based mainly on subjective criteria of the medical doctor.

The CBC is a quantitative and highly automatic procedure that gives the medical doctor objective and accurate information to help in the liver diagnosis process.

The CBC approach is composed by three main components; i) Features computation from multi-modal sources, ii) design and training of a specific suitable classification strategy that takes into account the CLD specificities and iii) diagnosis and validation of the method.

The main novelty of the method proposed in this chapter is a hierarchical classifier that mimics the structural approach of *differential diagnosis* followed in the clinical practice to identify the different stages of the CLD [8]. Instead of trying to classify a given liver in one stage of the disease from a set of possible stages by using a *multi-class* classifier, e.g. *k-Nearest Neighbor* (kNN) or *Support Vector Machine* (SVM), the hierarchical approach, represented in Fig.2, is used. In this strategy several partial binary decisions are taken according the natural evolution of the disease. In each step, a decision is taken by different binary classifiers trained, tunned and optimized specifically for that task.

The first classification step (CS) discriminates *normal* versus *pathology* liver. If the liver is classified as pathologic in this first step, discrimination of *chronic hepatitis without cirrhosis* and *cirrhosis* is attempted. In the last step *compensated cirrhosis* and *decompensated cirrhosis* are discriminated. The decompensated cirrhosis is assumed as the end-stage of every chronic liver disease before *hepatocellular carcinoma*.

The CBC (see Fig.2) design and optimization is performed at two levels: i) features and ii) classifier type and parametrization selection used specifically in each CS. This means that at each CS of our hierarchical approach the classifier type and features can be different.

The feature selection procedure is formulated as an optimization task with the sensitivity maximization criterion [9,10]. The set of features at each CS is tunned for the specificities of the corresponding CLD stage prior the classifier type selection. This is done by the *sequential forward floating selection* (SFFS) [11] method with the linear discriminant analysis (LDA) criterion. The leave-one-out cross-validation technique is used for error estimation purposes.

The classifier selection at each CS is done with ROC analysis where the selected classifier, kNN or SVM [12], is the one that jointly maximizes the *true positive rate* (TPR) and the *true negative rate* (TNR).

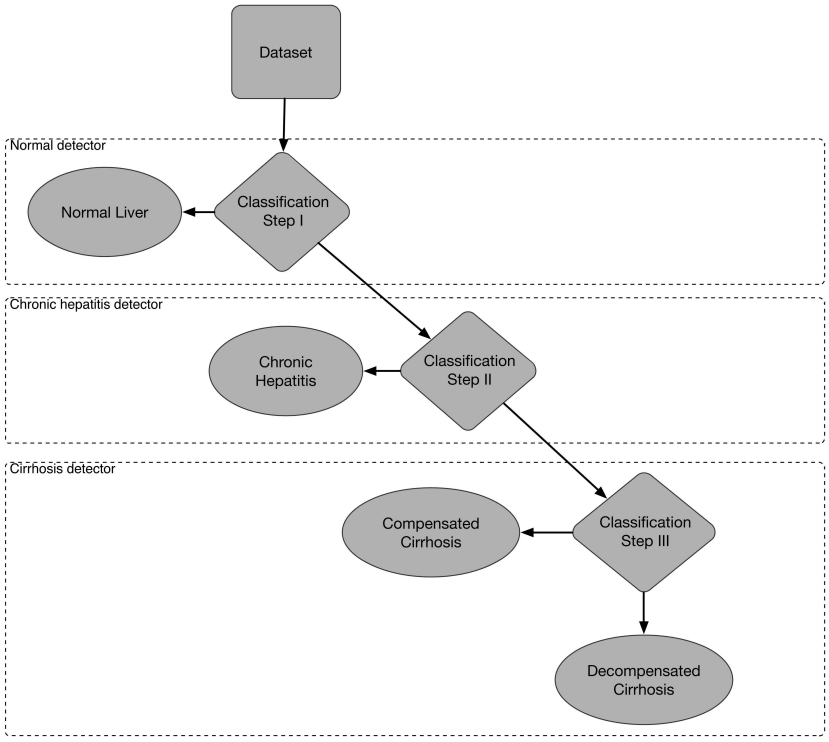

Fig. 2 Design of the CBC decomposition strategy for CLD classification.

The non-parametric kNN classifier is tested in this study. It classifies a test sample to a class according to the majority of the training neighbors in the feature space by using the minimum Euclidean distance criterion [13,14]. The algorithm for the *nearest neighbor rule* is summarized as follows [12]; Given an unknown feature vector x and a distance measure, then:

- Out of the N training vectors, identify the k nearest neighbors, regardless of class label.
- Out of these k samples, identify the number os vectors, k_i, that belong to class ω_i, i=1, 2, ..., M.
- Assign x to the class ω_i with the maximum number k_i of samples.

The other classifier tested in is the SVM classifier. Its aim is to find a decision plane that has a maximum distance (margin) from the nearest training pattern[14,15]. Given the training data $\{(x_i, \omega_i)|\omega_i = 1\ or\ -1,\ i = 1,...,N\}$ for a two-class classification (where x_i is the input feature; ω_i is the class label and N is the number of training sample), the SVM maps the features to a higher-dimensional space [14]. Then, SVM finds a hyperplane to separate the two classes with the decision boundary set by the support vectors[15].

The general form of the decision function $g(x)$ for the SVM is [15]:

$$g(x) = \sum_{i=1}^{N} \alpha_i y_i z_i^T z + b \tag{1}$$

where the α_i and b are selected during the training process, constrained by $\sum \alpha_i y_i = 0$ and $0 \leq \alpha_i \leq c$, where c is a user-defined penalty term, regulating the general performance of the SVM. Under certain conditions, the computationally intensive mapping process can be reduced with an appropriate kernel function K such that the decision function g(x) becomes,

$$g(x) = \sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b \tag{2}$$

In this paper, the kernels adopted are the polynomial,

$$K(x_i, x) = (x_i^T x + 1)^d \tag{3}$$

and the Gaussian radial-basis function,

$$K(x_i, x) = e^{-r|x_i - x|^2}, \tag{4}$$

where d is the degree of the polynomial kernel and r is the radius coefficient of the Gaussian radial basis function kernel [15].

The performance of the CBC is assessed by comparing the corresponding classification results with the common multi-class decomposition strategy - *one against all*. The *one against all* (OAA) strategy consists on building one classification procedure per class, trained to distinguish the samples in a single a class from the samples in all remaining classes.

The features extracted from US data are some of the most important and discriminative ones to the diagnosis. Therefore, in the next section a detailed description of them and the way they are computed is described.

2.1 Ultrasound Image pre-processing and features extraction

The US images are corrupted by a type of multiplicative noise, called *speckle*, that is usually considered undesirable to interpret the morphological information about the anatomy of the organs under analysis. However, the *speckle* pattern contains useful information about the tissues that can be used in the medical diagnosis [16].

Here, the method described in [16] is used. The B-mode US images, acquired by the commonest US scanner available at most of medical facilities, is pre-processed and decomposed into two fields: i) *De-speckle* and ii) *Speckle*.

The pre-processing stage is used to estimate the RF US image that is not usually available in the US scanners but that is needed for the decomposition procedure. The decomposition is performed in two steps; i) Denoising, where the estimated RF US

image, $y_{i,j}$, is filtered to obtain the *De-speckle* field $x_{i,j}$ and ii) *Speckle* extraction, $\eta_{i,j}$, obtained from the RF US noisy image, $y_{i,j}$, and form the *De-speckle* field, $x_{i,j}$, obtained in the previous step, under the adoption of the *Multiplicative White Rayleigh Model* [16],

$$\eta_{i,j} = \frac{y_{i,j}}{x_{i,j}}. \tag{5}$$

The *Denoising* algorithm is formulated in a Bayesian framework were the pixels of the noisy RF US estimated image are assumed to be Rayleigh distributed [16]. Here, the *De-speckle* field is used to extract morphological features, such as liver contour regularity and attenuation coefficient with dept, and the *Speckle* one, containing the noise pattern, is used to extract textural features from the parenchyma of the liver. Figure 3 illustrates an example of the decomposition methodology.

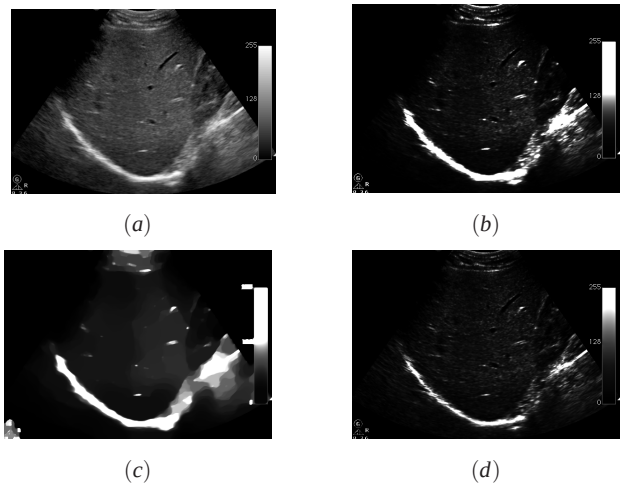

Fig. 3 Decomposition procedure of US liver parenchyma. a) Observed *B-mode* US image. Estimated b) envelope RF image, c) *de-despeckle* and d) *speckle* image fields.

Next a small description of each class of features is provided.

Acoustic attenuation coefficient

Acoustic attenuation along depth has been extensively used in the literature and in clinical practice as an indicator of the CLD. However, the true correlation between this indicator and the hepatic disease is still controversial as referred in [17], namely with respect to *cirrhosis* where some author report strong correlation between attenuation with depth while others do not find any significant relation. Other studies suggested that fibrosis can also produce large attenuation values but it was also reported fibrotic cases with normal attenuation.

In [18] the attenuation/backscatter is computed from the observed B-Mode US image instead of the more common RF US image. This is for sake of simplicity

and availability, since in most of the scanners the RF signal is not available. In their method a ROI corresponding to the biopsy site is selected and a polygonal line, describing the biopsy needle trajectory along the depth direction, is registered. The average gray-level values for each point along a polygonal line is computed by averaging 7 horizontal pixels and the attenuation coefficient is obtained by linear regression. They show that the attenuation clearly discriminates pure-fatty from normal livers where the area under the curve (AUC) is equal to 1.00 [18]. Another interesting result indicates that the presence of inflammation or fibrosis, even in cases of severe stetosis, leads to a significant decreasing on the discriminative power of the attenuation slope to distinguish healthy from steatotic livers.

Therefore, the attenuation coefficient, alone, does not has discriminative power for CLD diagnosis. However, together with other features in a more general classification framework, as the one presented here, it can provide valuable information for the performance of the classifier.

The attenuation coefficient, in $dBMHz^{-1}cm^{-1}$, can be obtained as follows [17],

$$\alpha(f) = \frac{1}{2l_{liver}} 10 \log_{10} \left(\frac{|S_p(f)|}{|S_{p|f}(f)|} \right) \tag{6}$$

where f is the frequency in MHz, l_{liver} is the thickness of the sample, $|S_p(f)|$ is the power spectrum without specimen, and $|S_{p|f}(f)|$ is the power spectrum with specimen. This generates an approximately linear attenuation curve inside the frequency band whose slope, obtained linear regression, is function of frequency. The global value is computed by integrating in frequency.

In this work the attenuation coefficient α is estimated by using the method proposed in [18] that is basically the following linear regression problem

$$\hat{m} = \arg\min_{m} \mathscr{J}(X) \tag{7}$$

where X is the $N \times M$ *De-speckle* field. The objective function is

$$\mathscr{J} = \sum_{l=0}^{N} (\alpha l + b - f(l))^2 \tag{8}$$

where $f(l) = \frac{1}{M} \sum_{c=1}^{M} x_{l,c}$ is the average value intensity of each *De-speckle* image line, in the assumption, that depth increases along each column

First-order statistics

First order statistics, computed from the histogram, $h(i)$, of the estimated gray-scale RF US image, Y, where pixel spatial correlation is not taken into account, are some of the most important features for the echo texture analysis of the liver parenchyma [7,13]. The first order characteristics used in this work are the mean and standard deviation

$$\mu = \frac{1}{NM}\sum_{i}^{256} h(i), \tag{9}$$

$$\sigma = \sqrt{\left(\frac{1}{NM-1}\sum_{i=1}^{256}(h(i)-\mu)^2\right)}, \tag{10}$$

Co-occurrence

The elements of the *Co-ocurrence* tensor, $Co = \{c_{i,j}(\Delta_l, \Delta_c)\}$, describe the gray level spatial inter-relationship in the image [13]. More precisely, element $c_{i,j}(\Delta_l, \Delta_c)$ represents the joint probability of the pixel intensities i and j in relative spatial position of (Δ_l, Δ_c) [15] and can be computed as follows

$$c_{i,j}(\Delta_l, \Delta_c) = \sum_{l=1}^{N}\sum_{c=1}^{M}\begin{cases} 1 \; if(\eta_{l,c} = i) \wedge (\eta_{l+\Delta l, c+\Delta c} = j) \\ 0 \; otherwise \end{cases} \tag{11}$$

In [15] it is reported an OA of 90% using features extracted from the *Co-occurrence* tensor in the detection of the *cirrhosis* condition. They show, in a high-frequency ($25MHz$) US study described in [19], that the values of the features obtained from the *Co-occurrence* tensor are able to discriminate cirrhotic, steatotic and healthy livers. However, the classification accuracy decreased when discrimination between different grades of steatosis and fibrosis are attenmpted.

The following statistical features, computed from the *Co-occurrence* tensor for $(\Delta_l, \Delta_c) \in \{(6,0)\}$, as suggested by [6], are used in this work:

- *Contrast*: measure the local variations in the co-occurrence matrix,

$$\sum_{i,j} |i-j|^2 c(i,j) \tag{12}$$

- *Correlation*: measure the joint probability occurrence of the specified pixel pairs,

$$\sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)c(i,j)}{\sigma_i \sigma_j} \tag{13}$$

- *Energy*: also known as the angular second moment [13],

$$\sum_{i,j} c(i,j)^2 \tag{14}$$

- *Homogeneity*: measures the closeness of the distribution of elements in the matrix to the matrix diagonal,

$$\sum_{i,j} \frac{c(i,j)}{(1+|i-j|)} \tag{15}$$

The Wavelet Transform

The *Wavelet Transform* is used here to perform a multi-scale analysis of the *Speckle* field, containing the noise pattern of the estimated RF US image, for liver parenchyma textural characterization purposes. The decomposition is performed according the Fig.4 where a sequence of low-pass, (G), and high-pass, (H), filtering operations followed by down-sampling the results, $\downarrow 2$, generates a pyramidal representation of the original image with decreasing resolution comprising a lower resolution low-pass component (approximation component), (LL), and three high-pass components (detailed components) along the horizontal, (HL), vertical, (LH), and diagonal directions, (HH), according Fig.5. An example of a multi-scale wavelet transform analysis using the *Speckle* field of an US liver image is provided in Fig. 6.

The high-pass components, (H) contain the detailed information of the image at different resolution scales along three directions while the low-pass versions (L) contain the approximation component.

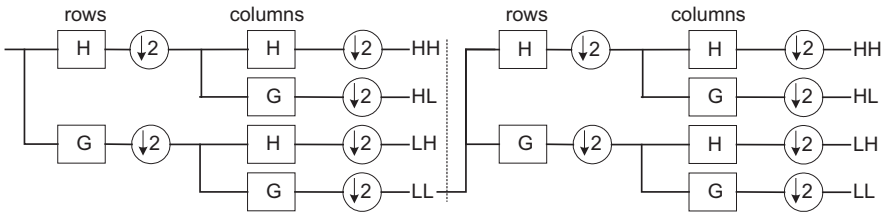

Fig. 4 Discrete image wavelet decomposition scheme (from [20])

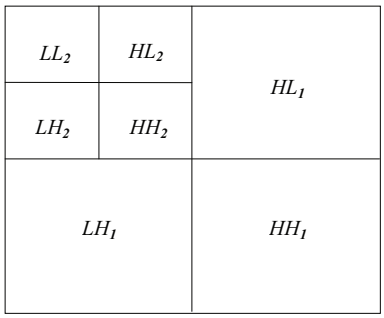

Fig. 5 Wavelet pyramidal decomposition of images.

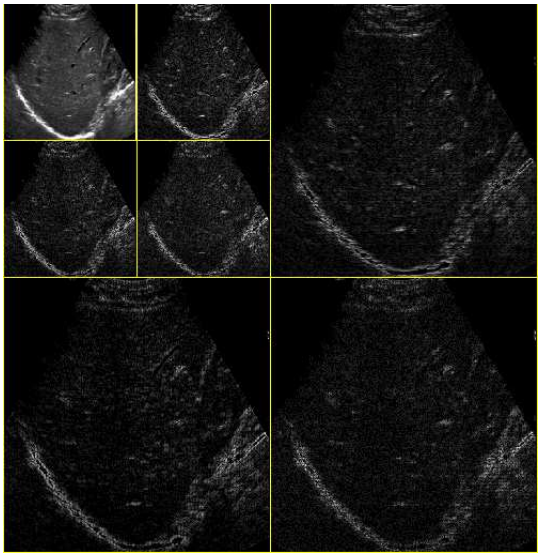

Fig. 6 Wavelet pyramidal decomposition example with the *Speckle* field of an US liver image.

Tissue characterization based on wavelet multi-resolution analysis as been performed in several works [6,15,19,21]. This approach is effective in the morphological characterization of the image from the approximation fields and at the same time in a textural characterization at several resolution scales from the detailed fields. For instance, in [6], the application of non-separable wavelet transform features for liver tissue characterization is investigated. The estimation of a texture quality is performed with the four-level *quincunx* decomposition which makes it possible to obtain feature vectors with maximal length of five elements [6] and an overall classification accuracy of 90% and a specificity of 92%. Sensitivity in the detection of cirrhosis and steatosis is 92% and 97%, respectively. Similar results are also reported in [15,19].

Autoregressive model

The autoregressive (AR) model approach has been used since long time with success in several applications of engineering where identification and characterization of systems and processes is needed [22]. In the canonical definition of a 1D p-order AR model each sample is modeled as a linear combination of the previous p samples with unknown coefficients, a_k [23]

$$x(n) = \sum_{i=1}^{p} a_i x(n-i) + r(n) \tag{16}$$

where the residue signal, $r(n)$, is assumed to be white and zero mean normal distributed. For image applications the following 2D formulation of the (p,q)-order AR model is used [24]

$$x(n,m) = \sum_{i=0}^{p} \sum_{j=0}^{q} a_{ij} x(m-i, n-j) + r(n,m) \tag{17}$$

where $x(n,m)$ is the ^{n,m}th pixel of the image and $a_{0,0} = 0$.

There are many algorithms to compute AR parameters; Levinson Burg, least-squares, gradient based, lattice filter and Kalman filter [25]. In this work, we will use the most popular, the least-squares algorithm [25]

The order of the model, (p,q), controls the error associated with the AR signal approximation [26]. Small orders ignore the main and long term statistical properties of the original signal while larger ones may lead to over-fitting effects [26,27]. Therefore, choosing the order of the model becomes a key problem and there are several methods to do it [23–27]. Here the first order model was adopted because it was confirmed by [28] that in this scope it leads to the minimum error probability.

Liver surface contour

Beside the textural features used to discriminate the pathologies of the liver, as referred before, US images may also be used to compute morphological features correlated with some of these pathologies. The importance of the US image in the assessment liver surface nodularity, portal vein mean flow velocity and the enlargement of the caudate lobe is stressed in [29]. Particularly, liver surface nodularity as been documented as a reliable sign in the detection of liver cirrhosis [3,30,31].

An accuracy of more than 70% as been reported by [3] and in [30] the authors showed that the observed liver contour irregularities directly correlated with the gross apperance of the cirrhotic liver as seen at laparoscopy. Liver surface nodularity in US can be well appreciated when ascites (presence of free fluid within the abdominal cavity) is present or when a high-frequency transducer (7.5 - 12 MHz) is used [32]. In [3], where a low-frequency transducer (3.5 -5 MHz), also refer that liver surface is an important parameter associated with the histopathological diagnosis of liver cirrhosis.

Despite the consensual correlation of the liver surface morphology with cirrhosis, the effectiveness of the different diagnosis methods used in the clinical practice is very limited because the analysis is most of the time subjective, non-reproducible and operator-dependent [33].

Here, a semi-automatic objective method for the liver surface characterization is proposed. The liver surface contour is segmented from the *De-speckle* component of the US image by using a snake technique proposed by [34], as shown in Fig. 7.

The irregularity of the contour is measured by means of the root mean square of the different angles, produced by the points that characterize the contour and the variation of the points of the contour in the *y* axis. In this approach the contour first

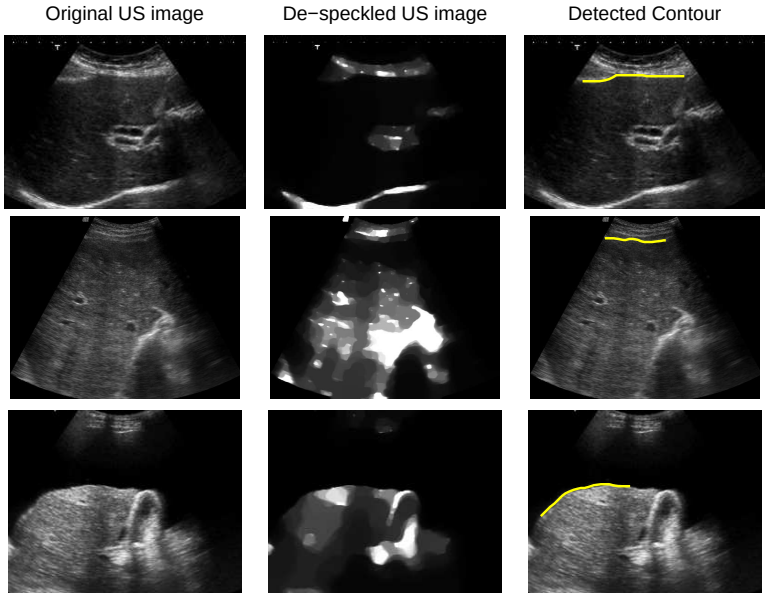

Fig. 7 Method used to detect the liver surface contour. First row corresponds to a normal liver; second row to a compensated cirrhotic liver and the last row to a decompensated cirrhotic liver.

point is assumed as the reference point. The first order statistics (mean and variance) of the referred variables are also extracted.

Biochemical and Clinical features

Besides image based features, several other clinical data and biochemical tests are useful for evaluating and managing patients with hepatic dysfunction.

These features are selected according to their purported clinical and pathophysiological role in CLD [8]. The clinical and pathophysiological characteristics of CLD can be grouped in terms of hepatic insufficiency, portal hypertension, hyperdynamic circulation, liver inflammation, necrosis and fibrosis, as well as etiologic factors [8,29,31]. Hepatic insufficiency is suggested by the Child-Pugh score, albumin, total bilirubin, encephalopathy and prothrombin time [8]. Portal Hypertension is usually accessed by the presence of ascites, varices and gastro-intestinal bleeding, among others. Creatinine and sodium are variables used for the study of hyperdynamic circulation [8]. Liver inflammation, necrosis, fibrosis and histology can be evaluated, according to [8], based on aspartate transaminase (AST), gamma glutamyl transpeptidase (gGT), lactate dehydrogenase (LDH) and alanine transaminase (ALT). Also the etiologic factors are taking into account, namely the alcohol abstinence, alcoholic ethiology, hepatitis B and C and other cause for the CLD [29].

The features extracted from the US images, the laboratorial and clinical information that were considered for the feature selection procedure are listed in Table 1.

Source	Feature
Liver echogenicity (de-speckled field)	**Acoustic attenuation coefficient** (F_1), measured by the slope coefficient of the linear regression of intensities along the depth/lines [18]
	First-order statistics, including the mean (F_2) and standard deviation (F_3) of the pixel intensities;
Liver surface contour (de-speckled field)	**Root mean square** of the **angles** produced by the points that characterize the contour (F_4), where the first point was assumed as the reference point.
	Root mean square of the **coordinates** of the contour points in the y axis (F_5)
	Mean (F_6) and **variance** (F_7) of the **angles**
	Variance of the y axis **coordinates** at each point (F_8).
	Correlation coefficient of the y axis coordinates (F_9).
Liver Texture (speckle field)	**Co-occurrence matrix**, which enables to derive [35]: the contrast (F_{10}) , correlation (F_{11}) that measures the joint probability occurrence of specific pixel pairs, energy (F_{12}) of the image (obtained by summing of squared elements of the image) and homogeneity (F_{13}) which quantifies the closeness of the distribution of matrix elements to its diagonal.
	Wavelet energies, measured by the vertical (F_{14}) and horizontal (F_{15}) detail energies of the first Haar wavelet decomposition.
	Autoregressive (AR) coefficients of a first order 2D model, $\{a_{0,0}(F_{16}), a_{1,0}(F_{17}), a_{0,1}(F_{18})\}$.
Laboratorial Information [29]	Total bilirubin (F_{19}) , prothrombin time (F_{20}) , albumin (F_{21}) , creatinine (F_{22}), AST (F_{23}), ALT (F_{24}), gGT (F_{25}) , glycemia (F_{26}), sodium (F_{27}), urea (F_{28}) and lactate dehydrogenase(F_{29}).
Clinical Information [29]	Cause of disease (F_{30}), which include none (0), alcohol (1), hepatitis B (2), hepatitis C (3), alcoholic hepatitis B (4) and C (5) and All (6), and the following binary indicators: tumor (F_{31}), ascites (F_{32}), encephalopathy (F_{33}), Gastro-Intestinal bleeding (F_{34}) infection (F_{35}), alcoholic habits (F_{36}) and Child-Pugh score (F_{37}).

Table 1 Features extracted from the US images, Laboratorial and Clinical information

2.2 Data set

Patients with an established diagnosis of chronic liver disease, in the various stages of the disease, were included for this study. A total of 115 US liver images from 115 patients, including 26 normal livers (*Normal*), 26 chronic hepatitis without cirrhosis (*Chronic Hepatities*), 27 compensated cirrhosis (*Compensated Cirrhosis*) and 36 decompensated cirrhosis (*Decompensated Cirrhosis*), were involved in the experiments.

The patients were selected from the Gastroenterology Department of the Santa Maria Hospital, in Lisbon, with known diagnosis based on liver biopsy results. Their

clinical histories, laboratory tests and US images were obtained in the same day. The study protocol was approved by the Ethics Committee of the referred Hospital, it was explained to the the patients and informed consent was obtained in each case, according to the principles of the Declaration of Helsinki.

All patients underwent hepatic US exam, performed by an expertize using a Philips ©CX50 scanner with a broadband curved array transducer with a frequency range of 2.0 to 5.0MHz. All US images were stored in DICOM format. A ROI of 128×128 pixels along the medial axis was extracted from each image and pre-processed by using the algorithm described in Section2.1, as shown in Fig. 8.

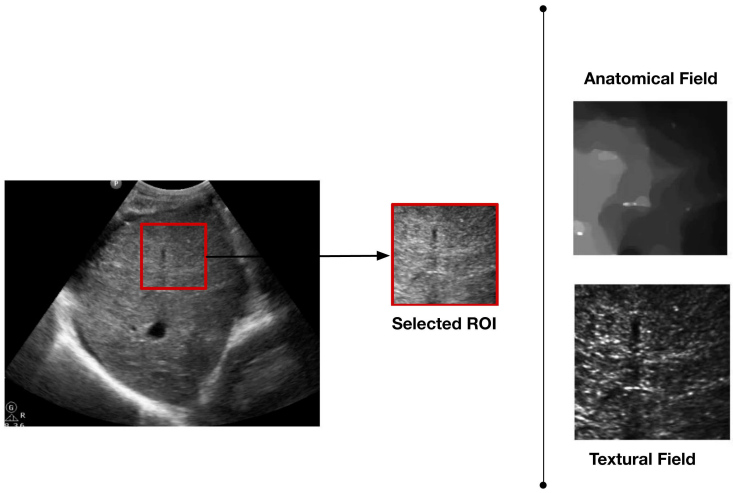

Fig. 8 Generic design for the selection of a ROI from the Ultrasound images and posterior pre-processing result

3 Results

In this section the results with data from real patients, obtained with the Matlab toolbox for Pattern Recognition, PRTools 4.1 [36], are presented to validate the proposed CBC method.

3.1 Feature Extraction and Selection

The features extracted from the different medical modalities computed in this work are all used along the classification pathway of the CBC. However, at each step,

only a fraction of the entire set is used, so a feature selection procedure is needed. Consequently, an optimum sub-set of features from the whole set is selected by maximizing the accuracy result of the LDA classifier. This optimization procedure is used for both classification strategies, OAA and CBC. Notice that at this stage we are interested in choosing the best features and not the best classifiers, that the reason only the LDA classifiers is used [9]. For comparison and performance analysis, classification results at each CS without feature selection is also computed and the results are shown in Table 2.

	Feat. Set	**Normal**	**Chronic Hepatitis**	**Comp. Cirrhosis**	**Decomp. Cirrhosis**
	US	3.9 (67.8)	23.1 (75.7)	25.9 (75.7)	44.4 (67.8)
No FS	LabClin	92.3 (91.3)	11.5 (72.1)	40.7 (74.8)	75.0 (86.0)
	All	88.5 (90.4)	30.8 (70.4)	37.0 (69.6)	72.2 (86.2)
	$OAA_{LabClin}$	100.0 (98.0)	62.0 (78.3)	48.0 (77.0)	81 (89.6)
FS	OAA	**100.0 (100.0)**	**68.0 (80.4)**	56.0 (78.0)	86 (92.0)
	CBC	**100.0 (100.0)**	65.4 (75.3)	**80.5 (82.5)**	**90.7 (82.5)**

Table 2 Sensitivity and overall performance, in parenthesis, percentage results for the one against all strategy with the US-based feature set (US), with laboratorial and clinical feature set (LabClin) and All features (All) without the feature selection procedure (No FS); and the results using the feature selection (FS) procedure for the one against all strategy with laboratorial and clinical feature set ($OAA_{LabClin}$), the with the complete feature set (OAA) and Clinical Based Classifier (CBC)

Table 2 shows the results obtained with only image, only clinical based features and with the complete features set without and with FS procedure. It is observed that by using only US-based features without feature selection, unacceptable results for clinical practice are obtained with sensitivities of 3.9%, 23.1%, 25.9% and 44.4% for the *Normal*, *Chronic Hepatitis*, *Compensated Cirrhosis* and *Decompensated Cirrhosis* class, respectively. Classification only with clinical based features, and again without feature selection, which is the usual approach in the daily clinical practice (LabClin), leads also to unsatisfactory results, in particular, for *Chronic Hepatitis* and *Compensated Cirrhosis* with sensitivities of 11.5% and 40.7%, respectively. The combination of both sets of features leads to marginal improvement for the *Chronic Hepaptitis* class but a decreasing for all other classes.

Feature selection seems to be a key operation to a good performance of the classification process as it can be observed from Table 2. In the second part of the Table the classification results are presented with FS for three different classificaion strategies; i) *One-Against-All* (OAA) only with laboratorial and clinical based features ($OAA_{LabClin}$), ii) OAA and iii) CBC.

The classifier inspired on *differential diagnosis* (CBC) with both US and LabClin features, selected with the proposed FS scheme, outperforms the other configurations with exception of the *Chronic Hepatitis* where the OAA strategy with FS is the better (68.0(80.4)%), as observed in Table 2.

The specific features for each CS on the CBC and OAA are listed in Table 3.

	OAA			CBC		
	US	Lab	Clinical	US	Lab	Clinical
N	-	-	24, 30, 36, 37	-	-	24, 30, 36, 37
CH	2, 16	19, 24, 25, 26, 28	33	16, 17	19, 24, 26, 28	37
CC	5,10,11,15,17	19,20,29	-	10,11,14,16,17	-	-
DC	15	21	30, 32, 34	10, 11, 14, 16, 17	-	-

Table 3 Set of features selected for the one against all strategy (OAA) and Clinical Based Classifier (CBC). **N**=Normal; **CHC**=Chronic Hepatitis; **CC**=Compensated Cirrhosis; and **DC**=Decompensated Cirrhosis.

The optimum features for the *Normal* detector in both strategies are the same, belonging entirely to the LabClin set; clinical $[F_{30}, F_{37}, F_{36}]$ and laboratorial $[F_{24}]$, which means that the traditional approach in clinical practice to detect illness based on clinical and laboratorial indicators seems to be appropriated.

For the *Chronic Hepatitis* the OAA strategy selects 8 features, mainly from the laboratorial, $[F_{19}, F_{24}, F_{25}, F_{26}, F_{28}]$ and clinical, $[F_{33}]$, sets. However, it also selects features from the US images; liver texture $[F_{16}]$ and liver echogenicity $[F_2]$. In the case of the CBC strategy for the *Chronic Hepatitis* detector, different optimum features are selected; 4 from the laboratorial set, $[F_{19}, F_{24}, F_{26}, F_{28}]$, two from the US set, texture $[F_{17}, F_{16}]$, and one from the clinical set, $[F_{37}]$.

The optimum set of features in the case of *Compensated Cirrhosis*, with the OAA strategy, are based on US images; texture, $[F_{15}, F_{10}, F_{17}, F_{11}]$, and surface contour, $[F_5]$, and on laboratorial information, $[F_{29}, F_{20}, F_{19}]$. For the *Decompensated Cirrhosis* this same strategy makes use of 5 features preponderantly from clinical data.

The discrimination of *Compensated* and *Decompensated Cirrhosis*, is obtained in the CBC strategy, by using only 5 US-based textural features, $[F_{16}, F_{17}, F_{14}, F_{10}, F_{11}]$.

These are the optimum features used in each CS. Next the results of a similiar procedure by using both strategies, OAA and CBC, for classifier selection are presented.

3.2 Classifier selection

For each CS one out of two classifiers, kNN or SVM, is selected according the classification accuracy optimality criterion for that step. Different parameterizations of each classifier are tested and the optimum configuration and classifier is select for each CS.

kNN classifier

Nine different neighborhood configurations, corresponding to $k = 1..9$, were tested in each CS with both optimal sets of features obtained in the previous section.

	N detector			CHC detector			CC detector			DC detector		
	TNR	TPR (N)	OA	TNR	TPR (CHC)	OA	TNR	TPR (CC)	OA	TNR	TPR (DC)	OA
k=1	**100**	**100**	**100**	87.64	42.31	77.39	87.5	29.63	73.91	92.41	88.89	91.3
k=2	100	100	100	93.26	15.38	75.65	95.45	22.22	78.26	96.2	75	89.57
k=3	100	100	100	87.64	46.15	78.26	90.91	44.44	80	**94.94**	**86.11**	**92.17**
k=4	100	100	100	91.01	38.46	79.13	95.45	33.33	80.87	96.2	83.33	92.17
k=5	100	100	100	88.76	46.15	79.13	**93.18**	**37.04**	**80**	92.41	86.11	90.43
k=6	100	100	100	93.26	30.77	79.13	94.32	25.93	78.26	93.67	83.33	90.43
k=7	97.75	100	98.26	**91.0**	**50**	**81.74**	92.05	37.04	79.13	91.14	83.33	88.7
k=8	97.75	100	98.26	94.38	34.62	80.87	95.45	14.81	76.52	93.67	83.33	90.43
k=9	97.75	100	98.26	88.76	50	80	94.32	25.93	78.26	92.41	83.33	89.579

Table 4 TNR, TPR and OA percentage results of the kNN classifier for the evaluation set of the OAA strategy for each class. **N**=Normal; **CHC**=Chronic Hepatitis; **CC**=Compensated Cirrhosis; and **DC**=Decompensated Cirrhosis.

	CS_1 (Normal detector)			CS_2 (CHC detector)			CS_3 (Cirrhosis detector)		
	TPR (N)	TNR	OA	TPR (CHC)	TNR	OA	TPR (CC)	TNR (DC)	OA
k=1	**100**	**100**	**100**	42.30	87.30	74.16	0	100	77.78
k=2	100	100	100	73.08	68.25	69.66	22.22	100	74.60
k=3	100	100	100	61.54	82.54	76.40	70.37	94.44	84.13
k=4	100	100	100	69.23	71.43	70.78	81.48	88.89	84.13
k=5	100	100	100	57.69	80.95	74.16	**80.78**	**91.67**	**85.71**
k=6	100	100	100	69.23	74.60	73.03	77.78	91.67	84.13
k=7	100	97.75	98.26	65.38	77.78	74.16	74.07	94.44	84.13
k=8	100	97.75	98.26	**73.08**	**74.60**	**74.16**	74.07	94.44	84.13
k=9	100	97.75	98.26	53.85	77.78	70.79	74.07	94.44	82.54

Table 5 TPR, TNR and OA percentage results of the kNN classifier for the evaluation set of the CBC strategy. **N**=Normal; **CHC**= Chronic Hepatitis; **CC**=Compensated Cirrhosis; and **DC**=Decompensated Cirrhosis.

Table 4 illustrates the results obtained for the OAA strategy with the kNN classifier. These results show a sensitivity to discriminate the *Normal* class of 100% for all configurations. The best sensitivities for *Chronic Hepatitis*, *Compensated Cirrhosis* and *Decompensated Cirrhosis* are 50.0%(k=7), 44.44% (k=2) and 86.11% (k=2,5), respectively.

The equivalent results obtained by the CBC strategy are shown in Table 5. As in the previous strategy, the performance of the normal detector is optimum, which means, maxima TPR, TNR and OA, for $k = 1, 2, 3, 4, 5, 6$.

In the case of the chronic hepatitis detector the best OA (74.16%) is for k=8, corresponding to a TPR of 73.08% and FPR of 74.6% (cirrhosis detection). In the last CS, the cirrhosis detector, the results reveal an OA of 85.71% for k=5. The TPR is 80.78% (*Compensated Cirrhosis*) and the TNR is 91.67% (*Decompensated Cirrhosis*).

SVM classifier with polynomial kernel

The same analysis performed previously for the kNN classifier will be now presented for the SVM one.

Several parameterizations of the polynomial kernel were tested. The parameter are the cost, $c = 1, 10, 100, 500$, and the degree, $d = 1, 2, 3, 4, 5$. Here, only the best results obtained for $c = 1$ are presented.

degree	N detector			CHC detector			CC detector			DC detector		
	TNR	TPR (N)	OA	TNR	TPR (CHC)	OA	TNR	TPR (CC)	OA	TNR	TPR (DC)	OA
1	96.63	100	97.39	89.89	23.08	74.78	81.82	59.26	76.52	91.14	72.22	85.22
2	98.88	100	99.13	**87.64**	**73.08**	**84.35**	**85.23**	**59.26**	**79.13**	93.67	91.67	93.04
3	98.88	100	99.13	85.39	61.54	80	79.55	55.56	73.91	92.41	88.89	91.3
4	**100**	**100**	**100**	89.89	50	80.87	80.68	55.56	74.78	**94.94**	**91.67**	**93.91**
5	98.88	100	99.13	88.76	46.15	79.13	80.68	51.85	73.91	89.87	91.67	90.43

Table 6 TNR, TPR and OA percentage results of the SVM classifier with polynomial kernel for the evaluation set of the OAA strategy. **N**=Normal; **CHC**= Chronic Hepatitis; **CC**=Compensated Cirrhosis; and **DC**=Decompensated Cirrhosis.

The results obtained with the polynomial kernel by using the OAA strategy are listed in Table 6. An ideal classification result is obtained in the normal detector for $d = 4$. For chronic hepatitis and cirrhosis detectors the best result, for $d = 2$, are: sensitivity 73.08% and 59.26% for *Chronic Hepatitis* and *Compensated Cirrhosis*, respectively. The optimum configuration in the *Decompensated Cirrhosis* case is for $d = 4$, corresponding to a senitivity of 91.67%.

degree	CS_1 (Normal detector)			CS_2 (CHC detector)			CS_3 (Cirrhosis detector)		
	TPR (N)	TNR	OA	TPR (CHC)	TNR	OA	TPR (CC)	TNR (DC)	OA
1	96.63	100	97.39	**69.23**	**84.13**	**79.76**	**77.78**	**91.67**	**85.71**
2	98.87	100	99.13	65.38	79.37	75.28	70.37	86.11	79.36
3	98.87	100	99.13	69.23	82.54	78.65	66.67	80.56	74.60
4	**100**	**100**	**100**	61.54	84.13	77.53	62.96	83.33	74.60
5	98.87	100	99.13	65.38	82.54	77.53	66.67	83.33	76.19

Table 7 TPR, TNR and OA percentage results of the SVM classifier with polynomial kernel for the evaluation of the CBC strategy. **N**=Normal; **CHC**= Chronic Hepatitis; **CC**=Compensated Cirrhosis; and **DC**=Decompensated Cirrhosis.

Table 7 shows the results obtained by using the proposed CBC approach. The normal detector for $d = 4$ is ideal under all figures of merit (OA, TPR, TNR = 100%). For the remaining CS the optimal kernel degrees are $d = 1$. In the chronic hepatitis detector the best OA is 79.76%, with a sensitivity of 69.23%. For the *Compensated Cirrhosis* the sensitivity is 77.78% while for the *Decompensated Cirrhosis* is 91.67%.

SVM classifier with Gaussian radial-basis kernel

To optimize the SVM with Gaussian radial-basis kernel, for the different classification tasks, the radius parameter (r) was ranged from 0.1 to 5 with steps of 0.5. The

c parameter was also tested in different values, $c = 1, 10, 100, 500$, for each radius value. The best results are obtained for $c = 10$.

radius	N detector			CHC detector			CC detector			DC detector		
	TNR	TPR (N)	OA	TNR	TPR (CHC)	OA	TNR	TPR (CC)	OA	TNR	TPR (DC)	OA
0.1	**100**	**100**	**100**	100	0	77.39	100	0	76.52	98.73	22.2	74.78
0.6	100	100	100	95.51	3.85	74.78	100	0	76.52	96.2	75	89.57
1.1	100	100	100	93.26	15.38	75.65	97.73	14.81	78.26	94.94	83.33	91.3
1.6	96.63	100	97.39	**93.26**	**26.92**	**78.26**	96.59	29.63	80.87	91.14	83.33	88.7
2.1	95.51	100	96.52	92.13	23.08	76.52	**96.59**	**44.44**	**84.35**	91.14	83.33	88.7
2.6	96.63	100	97.39	92.13	11.54	73.91	95.45	29.63	80	**91.14**	**86.11**	**89.57**
3.1	95.51	100	96.52	96.63	3.85	75.65	94.32	11.11	74.78	91.14	86.11	89.57
3.6	95.51	100	96.52	98.88	3.85	77.39	93.18	3.7	72.17	91.14	86.11	89.57
4.1	95.51	100	96.52	98.88	3.85	77.39	96.59	0	73.91	91.14	86.11	89.57
4.6	95.51	100	96.52	98.88	3.85	77.39	97.73	0	74.78	91.14	75	86.09

Table 8 TNR, TPR and OA percentage results of the SVM classifier with radial basis kernel for the evaluation set of the OAA strategy. **N**=Normal; **CHC**= Chronic Hepatitis; **CC**=Compensated Cirrhosis; and **DC**=Decompensated Cirrhosis.

The SVM classification results with Gaussian radial-basis kernels for the OAA strategy are summarized in Table 8. For this strategy, this classifier exhibit the worst results for all detectors when compared with the other tested classifiers . The sensitivity is 100% for *Normal* class, with the first three radius values, a sensitivity of 26.92% for *Chronic Hepatitis* class, with $r = 1.6$, 44.44% for *Compensated Cirrhosis*, with a $r = 2.1$, and 86.11% for *Decompensated Cirrhosis* class, with $r = 2.6$.

radius	CS_1 (Normal detector)			CS_2 (CHC detector)			CS_3 (Cirrhosis detector)		
	TPR (N)	TNR	OA	TPR (CHC)	TNR	OA	TPR (CC)	TNR (DC)	OA
0.1	**100**	**100**	**100**	0	100	70.79	0	100	57.14
0.6	100	100	100	23.08	96.83	75.28	22.22	100	66.67
1.1	100	100	100	**38.46**	**95.24**	**78.65**	70.37	94.44	84.12
1.6	96.63	100	97.39	34.62	95.24	77.53	81.48	88.89	85.71
2.1	95.50	100	96.52	34.62	93.65	76.40	**81.48**	**91.67**	**87.30**
2.6	96.63	100	97.39	23.08	93.65	73.03	77.78	91.67	85.71
3.1	95.50	100	96.52	19.23	95.24	73.03	74.07	94.44	85.71
3.6	95.50	100	96.52	11.54	96.83	71.91	74.07	94.44	85.71
4.1	95.50	100	96.52	11.54	100	74.16	74.07	94.44	85.71
4.6	95.50	100	96.52	11.54	100	74.16	74.07	91.67	84.13

Table 9 TPR, TNR and OA percentage results of the SVM classifier with radial basis kernel for the evaluation set of the CBC strategy. **N**=Normal; **CHC**= Chronic Hepatitis; **CC**=Compensated Cirrhosis; and **DC**=Decompensated Cirrhosis.

With the CBC strategy the results are better than using the OAA one as shown in Table 9.As in the previous case, the normal detector presents OA, TPR and TNR of 100% with $r = 0.1, 0.6, 1.1$. The wort results is for the *Chronic Hepatitis* where the sensitivity is 38.46% and the OA is 78.65% ($r = 1.1$). For the last CS of the CBC approach, the optimal radius is $r = 2.1$ resulting in an OA of 87.3%. The individual performances showed a sensitivity of 81.48% and 91.67% for *Compensated Cirrhosis* and *Decompensated Cirrhosis*, respectively.

4 Discussion

The proposed CBC algorithm for diagnosis and staging of the CLD has been applied to an experimental database of 115 patients, with well established diagnosis, according the guidelines accepted in the gastroenterology medical community.

The multi-modal feature set proved to be an appropriate approach for this classification problem. An algorithm was designed and implemented to select the optimal set of features for each CS. The performance of the FS algorithm was evaluated and the results compared with the ones obtained without FS procedure. This procedure, based on the SFFS with LDA criterion, leads to a clearly improvement of the classification accuracy and sensitivity in all classes/detectors.

The results reveal an important issue: each modality feature set has, at each CS, different discriminative power. In the first CS, laboratorial and clinical features are more discriminative than the US based features. As already stated before, this fact confirms the appropriateness of the traditional diagnosis method used in clinical practice. For more severe stages of CLD, the FS procedure automatically attributes more and more importance to the US image based features. US textural features, obtained from the first order AR model coefficients and the wavelet coefficients obtained form the first Haar wavelet decomposition, are particular relevant.

This fact corroborates the difficulty reported in the literature [3] in discriminating advanced stages of the CLD only from laboratorial and clinical information. This is particularly visible in the last CS, where only US based features are automatically selected.

The US textural features selected here are in accordance with the results of several other studies [6,15,21,27]. In these studies, wavelet transform [6,15,21] and AR coefficients [27] based features, from US images, are presented as having high discriminative power in the assessment of CLD stages.

Liver contour US based features were extract from a semi-automatic post-processing algorithm. Apparently, this set of features is not very important because they are only selected in one of the strategies, the OAA one. In the proposed CBC they are not selected at all. This class of features was included in this study because several works in the literature [3,30,31] refer its potential usefulness for CLD diagnosis. However, our results also reinforce the conclusions on [33], where it is stated that the liver surface morphology analysis performed based on US data is subjective and non-reliable.

Another important conclusion is that US echogenicity is not suitable for discrimination in CLD. This confirms previous results in the literature [17,18] where it was shown the low discriminative power of the features $[F_1, F_2, F_3]$.

The main novelty of this work is the proposal of a new classification and staging strategy for CLD where the natural evolution of the disease is taken into account. The well established protocol of *differential diagnosis*, used in the medical community, also inspired the design of the CBC. This approach outperforms the OAA strategy.

For the *Normal* class the results showed a perfect accuracy in discriminating the normal from the diseased patients, for all classifiers in both decompositions. It is

possible to refer that we improved the reported results described in the literature [6,13,15]. This is due to the proposed multi-modal feature and feature selection approach, since the system only selected clinical and laboratorial data. The disease clinical knowledge proved to be crucial for this CS.

Liver fibrosis is related, in part, with *Chronic Hepatitis* patients without cirrhosis. In the study preformed by [15] it is referred the difficulty in the classification of liver fibrosis based on US images. Their study revealed an OA of 72%, with a sensitivity of 60% for patients with fibrosis grade of 3 (*Chronic Hepatitis* class) and 88.6% for patients with fibrosis grade of 4 (*Cirrhosis* class). In the present study, we were able to improve the detection of the *Chronic hepatitis* class by using the SVM with a polynomial kernel of degree $d = 1$, when compared with the results presented in [15]. Our results are an OA of 79.8% with a sensitivity of 69.2% and 84.1% for *Chronic Hepatitis* and *Cirrhosis* class, respectively.

Concerning the cirrhosis detector, with the CBC strategy, [3] reported that the accuracy in detecting signs of compensated cirrhosis has not been well investigated.

Concerning the cirrhosis detection, our results show how important the US-based features are in CBC. The best results were obtained with the Guassian radial-basis SVM classifier ($r = 2.1$), where an OA of 87.3% with a sensitivity of 81.48% for the *Compensated cirrhosis* class and 91.7% for the *Decompensated cirrhosis* class were achieved. In [3] a sensitivity of 82.2% for the *Compensated cirrhosis* class was reported.

The final optimized configuration adopted for the CBC classifier, according all tests of features and classifiers selection performed in this work, is the following (see Fig.9): CS_1, Normal detector, is a kNN with ($k = 1$); CS_2, Chronic hepatitis detector, is a polynomial kernel SVM with ($d = 1$); and CS_3, Cirrhosis detector, is a Guassian radial-basis kernel SVM with ($r = 2.1$).

5 Conclusions and Future Work

Many scientific and medical problems require different concepts or classes to be distinguished. The goal is to extract sufficient discriminatory information to assign an object reliably, each in the medical field is of major importance. In this sense this study was develop with the premise that the multi-class classification problem should address the normal behavior of chronic liver disease. To solve this investigation question, we develop the *Clinical Based Classifier* strategy. The proposed multi-feature and multi-classifier system, based on a pre-processing US image decomposition proved to be a useful approach to the CLD classfication problem.

The results presented in this study showed that it is possible to identify the different stages of chronic liver disease based on US liver images, particularly textural features, as well as laboratorial and clinical features. This is achievable by decomposing the data based on the different stages of the disease : The group with the most severe stage, *Decompensated cirrhosis*, is well discriminate, while patients in

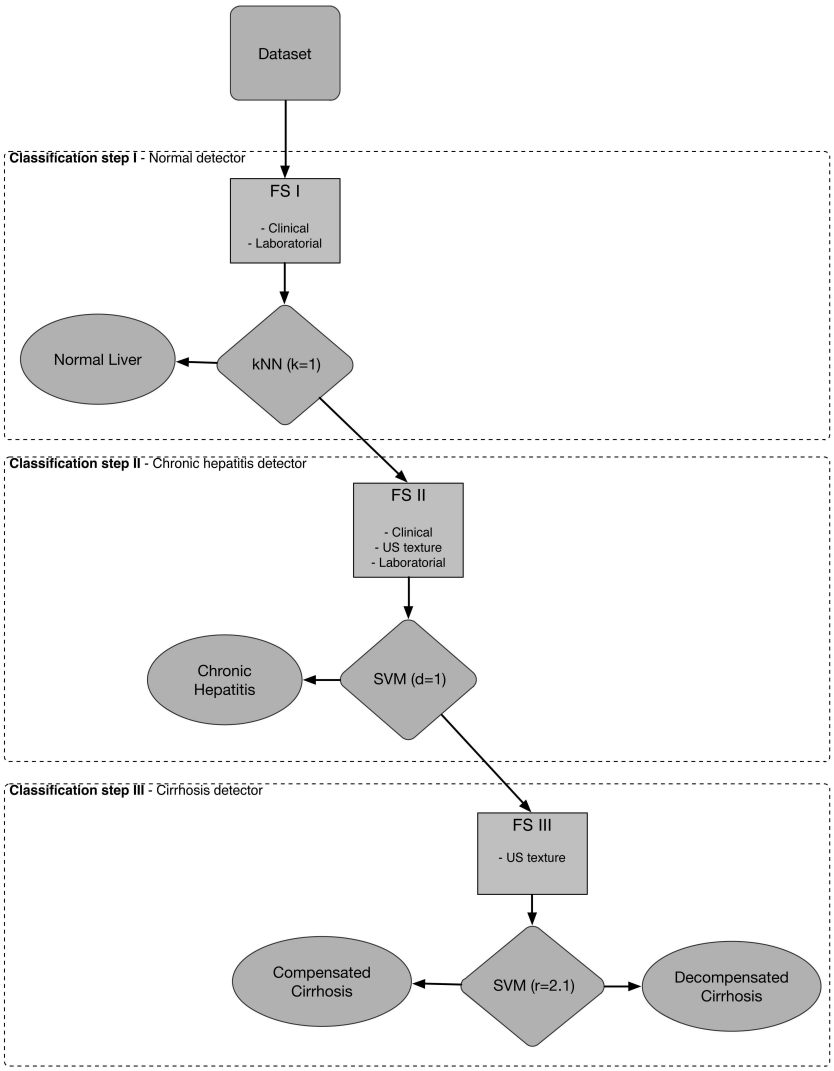

Fig. 9 Proposed algorithm for the classification of Chronic Liver Disease based on US images, clinical and laboratorial features

lower stages, *Chronic Hepatitis* and *Compensated cirrhosis*, need further analysis, in order to improve even more the classification results.

In future work, the proposed multi-feature approach will be expanded to incorporate more textural and morphological features. Moreover future work will also investigate classifier combination techniques as well as other features selection algorithms.

References

1. Richard Allan, Kerry Thoirs, and Maureen Phillips. Accuracy of ultrasound to identify chronic liver disease. *World J Gastroenterol*, 28(16):3510–3520, July 2010.
2. İhsan Ömür Bucak and Semra Baki. Diagnosis of liver disease by using cmac neural network approach. *Expert Syst. Appl.*, 37:6157–6164, September 2010.
3. Stefano Gaiani, Laura Gramantieri, Nicola Venturoli, Fabio Piscaglia, Sebastiano Siringo, Antonia D'Errico, Gianni Zironi, Walter Grigioni, and Luigi Bolondi. What is the criterion for differentiating chronic hepatitis from compensated cirrhosis? a prospective study comparing ultrasonography and percutaneous liver biopsy. *Journal of Hepatology*, 27(6):979 – 985, 1997.
4. SiRui Zhou and Jiuqing Wan. A survey of algorithms for the analysis of diffused liver disease from b-mode ultrasound images. In *Electronic Measurement Instruments, 2009. ICEMI '09. 9th International Conference on*, pages 2–576 –2–582, aug. 2009.
5. Guitao Cao, Pengfei Shi, and Bing Hu. Liver fibrosis identification based on ultrasound images captured under varied imaging protocols. *J Zhejiang Univ Sci B*, 6(11):1107–14, 2005.
6. A. Mojsilovic, S. Markovic, and M. Popovic. Characterization of visually similar diffuse diseases from b-scan liver images with the nonseparable wavelet transform. *Image Processing, International Conference on*, 3:547, 1997.
7. Chang Hee Lee, Jae Woong Choi, Kyeong Ah Kim, Tae Seok Seo, Jong Mee Lee, and Cheol Min Park. Usefulness of standard deviation on the histogram of ultrasound as a quantitative value for hepatic parenchymal echo texture; preliminary study. *Ultrasound in Medicine & Biology*, 32(12):1817 – 1826, 2006.
8. Gennaro D'Amico, Guadalupe Garcia-Tsao, and Luigi Pagliaro. Natural history and prognostic indicators of survival in cirrhosis: A systematic review of 118 studies. *Journal of Hepatology*, 44(1):217 – 231, 2006.
9. Anil Jain and Douglas Zongker. Feature selection: Evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:153–158, 1997.
10. Chao Sima, Sanju Attoor, Ulisses BragNeto, James Lowey, Edward Suh, and Edward R. Dougherty. Impact of error estimation on feature selection. *Pattern Recognition*, 38(12):2472–2482, 2005.
11. P. Pudil, J. Novovicová, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15(11):1119–1125, 1994.
12. Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern Recognition, Fourth Edition*. Academic Press, 4th edition, 2008.
13. Y. Kadah, 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*, 15:466–478, 1996.
14. Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2000.
15. W. Yeh, Y. Jeng, C. Li, P. Lee, and P. Li. Liver fibrosis grade classification with b-mode ultrasound. *Ultrasound in Medicine & Biology*, 29:1229–1235, 2003.
16. José C. Seabra and João M. Sanches. On estimating de-speckled and speckle components from b-mode ultrasound images. In *Proceedings of the 2010 IEEE international conference on Biomedical imaging: from nano to Macro*, ISBI'10, pages 284–287. IEEE Press, 2010.
17. M. Meziri, W. Pereira, A. Abdelwahab, C. Degott, and P. Laugier. In vitro chronic hepatic disease characterization with a multiparametric ultrasonic approach. *Ultrasonics*, 43(5):305 – 313, 2005.
18. D. Gaitini, Y. Baruch, E. Ghersin, E. Veitsman, H. Kerner, B. Shalem, G. Yaniv, C. Sarfaty, and H. Azhari. Feasibility study of ultrasonic fatty liver biopsy: Texture vs. attenuation and backscatter. *Ultrasound in Medicine & Biology*, 30(10):1321 – 1327, 2004.
19. W. Yeh, Y. Jeng, C. Li, P. Lee, and P. Li. Liver steatosis classification using high-frequency ultrasound. *Ultrasound in Medicine & Biology*, 31(5):599 – 605, 2005.

20. Shutao Li and John Shawe-Taylor. Comparison and fusion of multiresolution features for texture classification. *Pattern Recognition Letters*, 26(5):633 – 638, 2005.
21. Jiuqing Wan and SiRui Zhou. Features extraction based on wavelet packet transform for b-mode ultrasound liver images. In *Image and Signal Processing (CISP), 2010, 3rd International Congress on*, volume 2, pages 949 –955, Oct. 2010.
22. Reijo Takalo, Heli Hytti, and Heimo Ihalainen. Tutorial on Univariate Autoregressive Spectral Analysis. *Journal of Clinical Monitoring and Computing*, 19(6):401–410, December 2005.
23. K. Wear, R. Wagner, and B. Garra. A comparison of autoregressive spectral estimation algorithms andorder determination methods in ultrasonic tissue characterization. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 42(4):709–716, July 1995.
24. P. Stathaki and A. Constantinides. Robust autoregressive modelling through higher order spectralestimation techniques with applications to mammography. *Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, 1:189–193, November 1993.
25. J. Girault, F. Ossant, A. Ouahabi, D. Kouame, and F. Patat. Time-varying autoregressive spectral estimation for ultrasoundattenuation in tissue characterization. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 45(3):650–659, May 1998.
26. N. Farnoud, M. Kolios, and S. Krishnan. Ultrasound backscatter signal characterization and classification using autoregressive modeling and machine learning algorithms. In *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, volume 3, pages 2861–2864 Vol.3, Sept. 2003.
27. T. Wang, J. Saniie, and X. Jin. Analysis of low-order autoregressive models for ultrasonic grainsignal characterization. *IEEE transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 38(2):116–124, March 1991.
28. J.S. Bleck, U. Ranft, M. Gebel, H. Hecker, M. Westhoff-Bleck, C. Thiesemann, S. Wagner, and M. Manns. Random field models in the textural analysis of ultrasonic images of the liver. *Medical Imaging, IEEE Transactions on*, 15(6):796–801, Dec 1996.
29. S. Sherlock and J.Dooley. *Diseases of the liver and Biliary System*. Blackwell Science Ltd, 11 edition, 2002.
30. V Simonovsky. The diagnosis of cirrhosis by high resolution ultrasound of the liver surface. *Br J Radiol*, 72(853):29–34, 1999.
31. Annalisa Berzigotti, Juan G. Abraldes, Puneeta Tandon, Eva Erice, Rosa Gilabert, Juan Carlos García-Pagan, and Jaime Bosch. Ultrasonographic evaluation of liver surface and transient elastography in clinically doubtful cirrhosis. *Journal of Hepatology*, 52(6):846 – 853, 2010.
32. V. Droga and D. Rubens. *Ultrasound Secrets*. Hanley and Belfus, 2004.
33. J A Ladenheim, D G Luba, F Yao, P B Gregory, R B Jeffrey, and G Garcia. Limitations of liver surface US in the diagnosis of cirrhosis. *Radiology*, 185(1):21–23, 1992.
34. Chris Bregler and Malcolm Slaney. *Snakes-A MatLab MEX file to demonstrate snake contour-following*, 1995.
35. Kazuo Maeda, Masaji Utsu, and Paul E. Kihaile. Quantification of sonographic echogenicity with grey-level histogram width: A clinical tissue characterization. *Ultrasound in Medicine & Biology*, 24(2):225 – 234, 1998.
36. R.P.W. Duin, P. Juszczak, P. Paclik, E. Pkalska, D. de Ridder, D.M.J. Tax, and S. Verzakov. PR-Tools4.1, a matlab toolbox for pattern recognition, 2007. http://prtools.org.