The Usefulness of Ultrasound in the Classification of Chronic Liver Disease

Ricardo Ribeiro, Rui Tato Marinho, José Velosa, Fernando Ramalho, J. Miguel Sanches and Jasjit S. Suri

Abstract— **Chronic Liver Disease is a progressive, most of the time asymptomatic, and potentially fatal disease. In this paper, a semi-automatic procedure to stage this disease is proposed based on ultrasound liver images, clinical and laboratorial data. In the core of the algorithm two classifiers are used: a k nearest neighbor and a Support Vector Machine, with different kernels. The classifiers were trained with the proposed multi-modal feature set and the results obtained were compared with the laboratorial and clinical feature set. The results showed that using ultrasound based features, in association with laboratorial and clinical features, improve the classification accuracy. The support vector machine, polynomial kernel, outperformed the others classifiers in every class studied. For the Normal class we achieved 100% accuracy, for the chronic hepatitis with cirrhosis 73.08%, for compensated cirrhosis 59.26% and for decompensated cirrhosis 91.67%.**

I. INTRODUCTION

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

Typically, CLD is established with the presence of hepatitis which can evolve to the end-stage of every CLD - cirrhosis. In the cirrhosis stage there are two phases a compensated one (asymptomatic) followed by the development of liver dysfunction, named decompensated cirrhosis.

Liver biopsy has an important role in the evaluation and staging of CLD. Nevertheless, due to it's invasive nature and the improved accuracy of noninvasive tests, its importance have diminished. In particularly, ultrasound (US) as proven to be an useful diagnostic procedure for CLD.

A review study performed by [1], showed that common features used to assess CLD based on clinical US practice are liver parenchyma echogenicity, texture and liver surface. Depict these features some variability was found in terms of diagnostic accuracy, which implies the subjective nature of US interpretation. So the development of an objective method based on US for CLD staging classification is critical.

Several studies have address this problem, using objective features based on the US images and classification procedures for the study of CLD. The most common features described are based on the first order statistics, co-occurrence matrix, wavelet transform, attenuation and backscattering

parameters and coefficients. In [2], it is referred that spatial resolution plays a central role in the characterization, which is shown in [3] work, where they used the autoregressive model to classify normal liver, hepatitis and cirrhosis.

The main contributions of the present work in the classification of CLD are the use of a pre-processing procedure to extract the textural and anatomical information from the US image; the use of multi-source features - US based features, laboratorial tests and clinical information; the inclusion of the cirrhosis phases in the classification problem; and the evaluation of different classifiers, k-nearest neighbor, support vector machine with polynomial and radial basis function kernels.

The paper will be structured as follows. In the next section a brief description of the pre-processing step, feature extraction and selection algorithms used is given. The experimental results and discussion follow in section III and IV. Finally, conclusions and future work are discussed in section V.

II. MATERIALS AND METHODS

A. Pre-processing algorithm

It is common practice to have the US images pre-processed before doing feature extraction and classification.

In this sense the observed *B-mode* ultrasound image is used to estimate the original *radio-frequency* (RF) envelope image without *log-compression*[4], performed by the ultrasound equipment to reduce the dynamic range of the US signal, and without bright and contrast tunning performed by the medical doctor during the exam. This first step aims at to obtain an estimate of the original envelope of the true RF signal provided by the US probe and make the classification results as independent as possible from the scanner and from the specific acquisition conditions of each image.

In a second step the estimated envelope RF image, $y(i,j)$, is decomposed into two fields; i) the *de-speckled* field, $x(i,j)$, mainly containing the noiseless anatomical information and ii) the *speckle* field, $\eta(i,j)$, containing the textural information[5]. In this type of images, involving coherent radiation, the *speckle* (pseudo)-noise corrupting the image is assumed to be multiplicative in the sense that its variance depends on the underlying noiseless image [6]. Therefore, the observation model assumed in this paper is the following

$$y(i,j) = x(i,j)\eta(i,j), \qquad (1)$$

where $\eta(i,j)$ are considered i.i.d (independent and identically distributed) random variables with *Rayleigh* distribution.

B. Feature Extraction and Selection

Three important US characteristics are used in the perception of CLD: Liver parenchyma echogenicity, its texture and liver surface contour. Based on this information and the knowledge of the must meaningful laboratorial and clinical information, the following features were extracted, according to Table I.

TABLE I

FEATURES EXTRACTED FROM THE US IMAGES, LABORATORIAL AND CLINICAL INFORMATION

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 [7]
	First-order statistics, including the mean (F_2) and variance (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 [8]: 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, $\{a_{0,0}(F_{16}),a_{1,0}(F_{17}),a_{0,1}(F_{18})\}$ (the first order model was adopted because it was confirmed by [3] that in this scope it leads to the minimum error probability).
Laboratorial Information [9]	Total bilirubin (F_{19}) , prothrombin time (F_{20}) , albumin (F_{21}) , creatinine (F_{22}) , aspartate transaminase (F_{23}), alanine transaminase (F_{24}), gamma glutamyl transpeptidase (F_{25}) , glycemia (F_{26}), sodium (F_{27}), urea (F_{28}) and lactate dehydrogenase(F_{29}).
Clinical Information [9]	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 others (6), and the following binary indicators: Tumor (F_{31}), Ascites (F_{32}), presence of free fluid within the peritoneal cavity; encephalopathy (F_{33}), Gastro-Intestinal bleeding (F_{34}) infection (F_{35}), alcoholic habits (F_{36}) and CHILD score (F_{37}).

The liver surface contour is obtain using a snake technique, proposed by [10], from the *de-despeckled* image, which computes one iteration of the energy-minimization of active contour models. To initialize the snake, the operator needs to select four points of the liver surface. From figure 1 we can appreciate the contour steps.

The original feature vector (n=37) was reduced using the sequential forward floating selection. The criterion function used was the overall accuracy of the linear discriminant analysis classification rule. The leave-one-out cross-validation technique was considered for error estimation.

In order to evaluate the usefulness of the US based features in the classification of CLD, we perform the same methodology for the feature subset of only the laboratorial and clinical
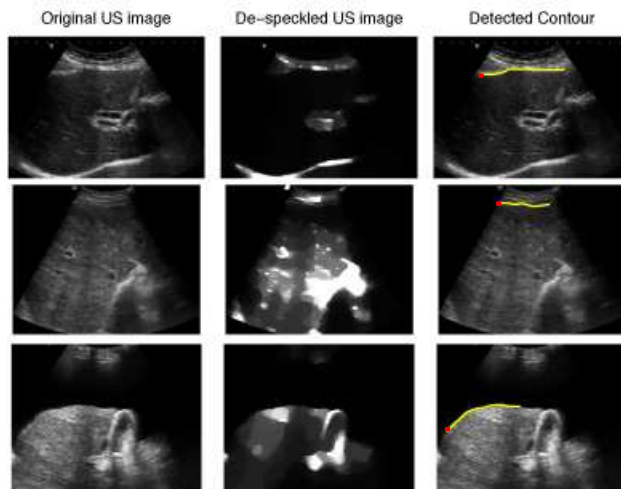

Fig. 1. Method used to detect the anterior liver surface contour. First row corresponds to a normal liver; second row to a compensated cirrhotic liver and the last row to a decompensated cirrhotic liver. The red dot shows the first snake point.

information (n=19) and compare the classification results.

C. Classification

Two different classifiers were implement and tested: i) the SVM and ii) *k-nearest neighbor* (kNN). A short description of each is provided.

The aim of SVM is to find a decision plane that has a maximum distance (margin) from the nearest training pattern [11], [12]. 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 [12]. Then, SVM finds a hyperplane to separate the two classes with the decision boundary set by the support vectors [11]. In this paper, a binary SVM classifier was adopted, using a Gaussian radial-basis kernel function and a polynomial kernel.

The non-parametric kNN classifier is also tested in this paper. 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], [12].

In the implementation step, we train the SVM using a polynomial kernel with a degree varying from 1 to 5, the SVM using a Gaussian radial-basis kernel raging from 0.1 to 5 and the k of the kNN classifier from 1 to 9. The cost, C parameter, was shift between 1, 10, 100 and 500 in each of the other trained parameters for the SVM classifiers.

Due to the fact that this is a multi-classification problem and using the premise that the physicians don't know the CLD stage when they examine the patient, we use one-against-all algorithm to evaluate this classification problem. This approach leads to a highly unbalanced classes, which is addressed by weighting classes according to their size. To avoid overtraining, feature scaling was performed by rescaling the data to have zero mean and unit variance. The classifier selection at each class is done with a ROC analysis approach where the selected classifier, kNN or SVM, is the one that maximizes the sensitivity. Sensitivity is the ability

of the classifier to correctly identify patients who are known to belong to the different disease stages. All classifiers were implemented using a Matlab toolbox for Pattern Recognition, PRTools 4.1 [14].

III. EXPERIMENTAL RESULTS

A total of 115 US liver images from 115 patients, including 26 normal livers (ω_N), 26 chronic hepatitis without cirrhosis (ω_{CHC}), 27 compensated cirrhosis (ω_{CC}) and 36 decompensated cirrhosis (ω_{DC}), 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. A ROI of 128×128 pixels along the medial axis was extracted from each image. No acquisition protocol was used due to the use of the pre-processing algorithm.

TABLE II
FEATURE SELECTION RESULTS USING ONLY THE CLINICAL AND LABORATORIAL FEATURES AND ALL FEATURES.

	All Features	Laboratorial and clinical
ω_N	$F_{30}, F_{37}, F_{24}, F_{36}$	F_{19}, F_{30}
ω_{CHC}	$F_{16}, F_{19}, F_2, F_{24}, F_{25}, F_{33}, F_{26}, F_{28}$	$F_{19}, F_{30}, F_{23}, F_{25}, F_{20}, F_{28}, F_{26}$
ω_{CC}	$F_{15}, F_{29}, F_{10}, F_{17}, F_{11}, F_{20}, F_{19}, F_5$	F_{29}, F_{19}, F_{37}
ω_{DC}	$F_{21}, F_{30}, F_{32}, F_{15}, F_{34}$	$F_{21}, F_{30}, F_{32}, F_{34}, F_{35}$

The selected features for each feature set are shown in Table II. The results showed that US based features improve the performances of the classification rule, with a mean improvement of 4% for each class. In ω_N we improved from 98% to 100%, where in both cases there was only selected laboratorial and clinical features. For ω_{CHC} the complete feature set showed better performance with an accuracy of 84% comparing to 78% obtained with the other set, where the inclusion of US based features is observed. Similar results were obtained for ω_{CC} and ω_{DC}.

From the selection feature procedure, we have obtained an optimal feature set for each class, according to their origin: All features and laboratorial and clinical features (LabClin). As reported earlier we trained three types of classifiers and compared the results in terms of sensitivity for each class.

During the kNN training the results showed that for ω_N the set from All features performed with 100% sensitive for all values of k, whereas the set from the LabClin features only performed with this rate from k=3. Figure 2 illustrate the error rate obtained in the training of the kNN classifier. For the optimal set of All features the best diagnostic yield for ω_{CHC} was only 50.0% with k=7, for ω_{CC}, it was 44.44% sensitive with k=2 and for ω_{DC} the sensitivity obtained was 89.0% with k=1. The results with the best features from the LabClin set showed a sensitivity of 53.85% for ω_{CHC} with k=9, 59.26% for ω_{CC} with k=3 and 5, and only 22.22% for ω_{DC} with k=1, 3, 5, 7 and 9.

In the training step of the SVM with polynomial kernel, the All features optimal set displayed a sensitivity of 100% in ω_N for all the implemented degrees, a sensitivity of 73.08%
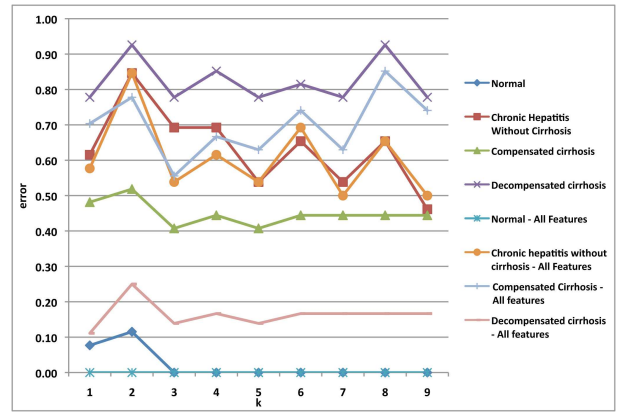

Fig. 2. kNN training results using only the laboratorial and clinical features and for all features.
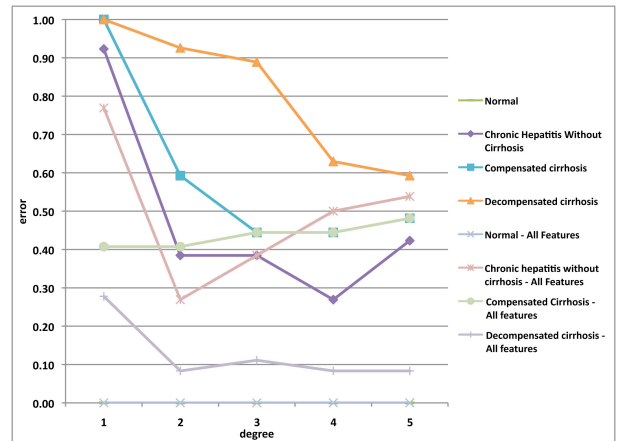

Fig. 3. SVM with polynomial kernel training results using only the laboratorial and clinical features and for all features.

for ω_{CHC} with a degree of 2, 59.26% for ω_{CC} with a degree of 2 and 91.67% for ω_{DC} with a degree of 4. In the case of the selected features with the LabClin set the results demonstrated an sensitivity of 100% for ω_N, 73.08% for ω_{CHC} using a degree of 4, 55.56% for ω_{CC} with a degree of 3 and 4, and for ω_{DC} we achieved a sensitivity of only 40.74% with a degree of 5. The results using the polynomial kernel SVM classifier are resumed in Figure 4.

The implementation of the SVM classifier with the Radial Basis kernel, summarized in Figure 4, exhibit the worst results of the training step, particularly with the LabClin set. For the All feature set it achieved a sensitivity of 100% for ω_N with a radius of 0.1, 0.6 and 1.1, 26.92% for ω_{CHC} with a radius of 1.6, 44.44% for ω_{CC} with a radius of 2.1 and 86.11% for ω_{DC} with a radius of 2.6. In the case of the LabClin set, the results showed a sensitivity of 100% for ω_N (radius=0.6 to 3.6), 42.31% for ω_{CHC} (radius=1.6), 51.85% for ω_{CC} (radius=0.6) and 25.93% for ω_{DC} with a 0.6 radius.

In the SVM implementation algorithm it was found that the optimal cost for both kernel functions was c=10.

IV. DISCUSSION

The proposed algorithm, based on a set of features from US, laboratorial and clinical, for CLD classification has been
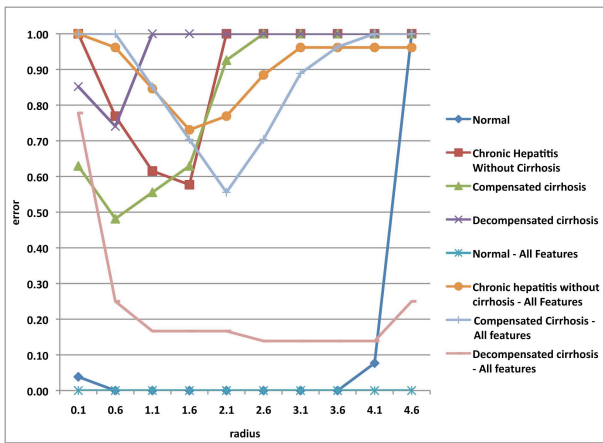
Fig. 4. SVM with radial basis function kernel training results using only the laboratorial and clinical features and for all features.

applied to an experimental database of 115 patients. The classification results are encouraging.

The use of a multi-feature source approach proved to be useful since the classification results outperformed the ones obtained with the the LabClin feature set. From the feature selection results it is noticeable the importance that the US-based features have in the discrimination of advanced CLD stages. Specially for ω_{CHC} and ω_{CC}, which translate the clinical difficulty reported in the classification of these two classes using the traditional approach. The liver contour analysis also manifest an importance in the classification of ω_{CC}. This result is in accordance with [15] that showed the importance of US liver surface nodularity in this disease.

In the normal class, the results showed a perfect accuracy using almost every classifier implementation tested. Out of 26 patients from ω_{CHC}, 19 were classified correctly with the SVM classifier with a polynomial kernel of second degree. In ω_{CC} the results were poorer, showing the best result 59.26% of sensitivity achieved with the same classifier implemented in ω_{CHC}. For ω_{DC}, the best result obtained was achieved with the SVM classifier with polynomial kernel of fourth degree, achieving a sensitivity of 91.67%. From the results attained the SVM with polynomial kernel outperformed the kNN and the SVM with the Radial Basis Kernel classifier.

The final optimized configuration adopted for the classification of CLD, according with the feature and classifiers selection performed in this work, is the following: ω_N - SVM polynomial kernel with degree 2, using clinical [F_{30}, F_{37}, F_{36}] and laboratorial [F_{24}] features; ω_{CHC} - SVM polynomial kernel with degree 2, using US liver texture[F_{16}] and echogenicity [F_2], laboratorial [F_{19}, F_{24}, F_{25}, F_{26}, F_{28}] and clinical [F_{33}] features; ω_{CC} - SVM polynomial kernel with degree 2, using US based features liver texture[F_{15}, F_{10}, F_{17}, F_{11}] and liver surface contour [F_5] and laboratorial [F_{29}, F_{20}, F_{19}] information; and ω_{DC} - SVM polynomial kernel with degree 4, using laboratorial [F_{21}], clinical [F_{30},F_{32}, F_{34}] and US texture [F_{15}].

V. CONCLUSIONS AND FUTURE WORK

The proposed multifeature and multiclassifier 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 paper show that it is possible to identify the different stages of CLD based on US liver images, particularly textural and contour parameters, laboratorial and clinical features. The group with the most severe stage, ω_{DC}, is well identifiable, while patients in lower stages, ω_{CHC} and ω_{CC}, need further analysis.

In future work, the proposed multifeature approach will be expanded to incorporate more textural and morphological features. Moreover future work will also investigate classifier combination techniques as well as other classifier such as Neural Networks.

REFERENCES

[1] R. Allan, K. Thoirs, and M. Phillips, "Accuracy of ultrasound to identify chronic liver disease," *World J Gastroenterol.*, vol. 28, no. 16, p. 35103520, July 2010.

[2] U. Scheipers, H. Ermert, K. Konig, H. Sommerfeld, and T. Senge, "Diagnostics of prostate cancer based on ultrasonic multifeature tissue characterization," *Ultrasonics Symposium, 2004 IEEE*, vol. 3, pp. 2153– 2156, August 2004.

[3] J. 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*, vol. 15, no. 6, pp. 796–801, Dec 1996.

[4] J. Seabra and J. Sanches, "Modeling log-compressed ultrasound images for radio frequency signal recovery," *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, 2008.

[5] J. C. Seabra and J. a. 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*, ser. ISBI'10. IEEE Press, 2010, pp. 284–287.

[6] R. Ribeiro and J. Sanches, "Fatty liver characterization and classification by ultrasound," in *Proceedings of the 4th IbPRIA*. Springer, June 2009, pp. 354–361.

[7] 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*, vol. 30, no. 10, pp. 1321 – 1327, 2004.

[8] K. Maeda, M. Utsu, and P. E. Kihaile, "Quantification of sonographic echogenicity with grey-level histogram width: A clinical tissue characterization," *Ultrasound in Medicine & Biology*, vol. 24, no. 2, pp. 225 – 234, 1998.

[9] S. Sherlock and J.Dooley, *Diseases of the liver and Biliary System*, 11st ed. Blackwell Science Ltd, 2002.

[10] C. Bregler and M. Slaney, *Snakes-A MatLab MEX file to demonstrate snake contour-following*, 1995.

[11] W. Yeh, Y. Jeng, C. Li, P. Lee, and P. Li, "Liver fibrosis grade classification with b-mode ultrasound," *Ultrasound in Medicine & Biology*, vol. 29, pp. 1229–1235, 2003.

[12] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2000.

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

[14] R. Duin, P. Juszczak, P. Paclik, E. Pkalska, D. de Ridder, D. Tax, and S. Verzakov, "PR-Tools4.1, a matlab toolbox for pattern recognition," 2007, http://prtools.org.

[15] S. Gaiani, L. Gramantieri, N. Venturoli, F. Piscaglia, S. Siringo, A. D'Errico, G. Zironi, W. Grigioni, and L. Bolondi, "What is the criterion for differentiating chronic hepatitis from compensated cirrhosis? a prospective study comparing ultrasonography and percutaneous liver biopsy," *Journal of Hepatology*, vol. 27, no. 6, pp. 979 – 985, 1997.