Medical Informatics Europe '97
C. Pappas et al. (Eds.)
IOS Press, 1997

Computer Image Analysis of Ultrasound Images for Discriminating and Grading Liver Parenchyma Disease Employing a Hierarchical Decision Tree Scheme and the Multilayer Perceptron Neural Network Classifier

D. Cavouras[1], I. Kandarakis[1], I. Theotokas[3], E. Kanellopoulos[1], D. Triantis[2], I.Behrakis[3], E. K. Manesis[4], I. Vafiadi-Zoumpouli[5], P. Zoumpoulis[3]

Departments of [1]Medical Instrumentation Technology and [2]Electronics, Technological Educational Institution of Athens, Ag. Spyridonos Street, Aigaleo GR-122 10, Athens, Greece
[3]Department of Ultrasound, Evgenidio Hospital of Athens University, Papadiamantopoulou 20, 11528 Athens, Greece
[4]Academic Department of Medicine, Hippokration General Hospital, Athens, Greece
[5]First Department of Propaedeutic Medicine, "Laiko" General Hospital, Athens,Greece

Abstract. Differential diagnosis of liver parenchyma disease and grading of the hepatic disease on ultrasound is a common radiological problem that influences patient management. The aim of this study was to apply image analysis methods on ultrasound images for discriminating liver cirrhosis from fatty liver infiltration and for grading hepatic disease, which is important in the management of the patients. Ultrasound images of histologically confirmed 18 livers with cirrhosis, 37 livers with fatty infiltration, and 24 normal livers of healthy volunteers were selected and were digitized for further computer processing. Twenty two textural features were calculated from small matrix samples selected from the ultrasound image matrix of the liver parenchyma. These features were used in the design a three level hierarchical decision tree classification scheme, employing the multilayer perceptron neural network classifier at each hierarchical tree level. At the first tree level, classification accuracy for distinguishing normal from abnormal livers was 93.7%, at the second level the accuracy for discriminating cirrhosis from fatty infiltration was 90.9%, and at the third level the accuracy in distinguishing between low and high grading liver cirrhosis or fatty infiltration was 94.1% and 84.9% respectively. The proposed computer software system may be of value to the radiologists in assessing liver parenchyma disease.

1. Introduction

Clinical diagnosis in ultrasound is much dependent on the skill of the user and diagnostic accuracy may differ accordingly. To increase diagnostic accuracy computer aided tissue characterization methods have been developed and applied in diagnostic ultrasound [1]. Differentiating between liver cirrhosis and fatty liver infiltration on ultrasound is difficult since pictorial information in both liver diseases is similar.

In this work ultrasound images of normal livers and of livers with cirrhosis or fatty liver infiltration were processed employing image analysis methods. The purpose was to analyze the internal structure of the liver tissue to obtain information concerning its nature. This textural information and a neural network multilayer perceptron (MLP) classifier were used in the design of a software system for characterizing the type of liver tissue and for the grading of the cirrhosis or of the fatty liver infiltration. Grading is important in the management of patients treated for hepatitis and, to our knowledge, there are no studies of liver parenchyma grading by ultrasound image analysis methods.

2. Materials and Methods

The study comprised 55 ultrasonic liver images of 55 patients with histologically confirmed parenchymal disease and 24 ultrasonic images of healthy volunteers. In the group with liver disease there were 37 patients with fatty liver infiltration and 18 patients with cirrhosis due to chronic hepatitis. All ultrasound examinations were performed on an Acuson 128XP/10 scanner by the same experienced sonographers (P.Z. and I.T.) using similar scanning conditions to provide for quality and uniformity in patient management and data acquisition. In each examination at least one longitudinal scan of the liver through the area of interest was obtained. Each liver image was digitized by connecting the video output of the ultrasound scanner to a Screen Machine II frame grabber using 512x512x8 image resolution. One 10x10 or 20x20 pixels image-matrix sample from the area of interest was selected from each image. Data processing was performed on a Pentium 133MHz computer.

The computer software for classifying the sample-matrix of the liver into normal, cirrhosis, and fatty infiltration and for grading the disease was designed as a three level hierarchical decision tree (Fig. 1). At the first and second levels two-class discrimination tests were performed; normal-abnormal (1^{st} level), cirrhosis-fatty infiltration (2^{nd} level). At the third level two-class discrimination tests were performed concerning disease grading (low or high) for either the cirrhosis or the fatty infiltration. At each level, the multilayer perceptron classifier [2,3] and 22 textural features from each liver sample-matrix were employed; 4 features were obtained from the histogram of the sample-matrix, 14 were calculated from the co-occurrence matrix [3,4] and 4 were run length features [5]. At each level of the hierarchical tree the discriminatory ability of each one of the 22 textural features was tested employing the Student's t-test; only the best discriminating features ($p<0.001$) were selected and were employed in the design of the neural network classifier. At each tree level, the performance of the MLP classifier was tested by the leave-one-out method [3], and for all possible combinations (i.e., 2, 3, and 4 feature combinations) of the selected features in order to determine the highest classification accuracy with the minimum number of features. Classification accuracy at each hierarchical tree level was also tested in relation to the number of hidden layer nodes and hidden layers of the MLP classifier.

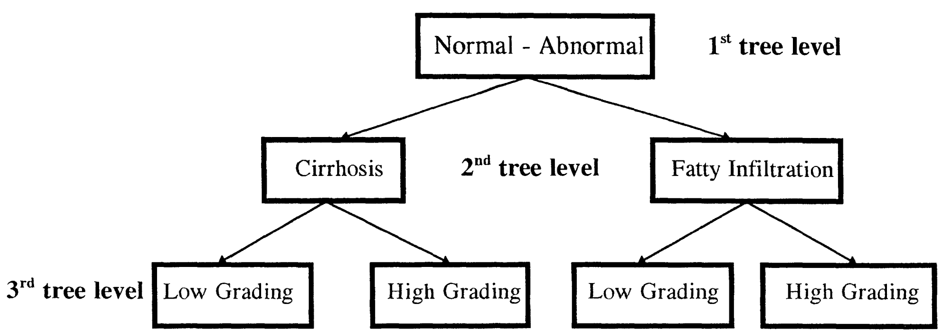

Figure 1. Hierarchical decision tree scheme using the MLP classifier at each tree level.

3. Results

Table 1: MLP-classifier discrimination accuracy between Normal-Abnormal livers

MLP:	Normal	Abnormal	Sums	Success
histology				
Normal	19	5	24	79 %
Abnormal	0	55	55	100%
Sums	19	60	79	
		Total Success		93.7 %

Table 2: MLP-classifier discrimination accuracy between fatty liver infiltration and liver cirrhosis

MLP:	Fatty	Cirrhosis	Sums	Success
histology				
Fatty	33	4	37	89.2%
Cirrhosis	1	17	18	94.4%
Sums	25	31	55	
		Total Success		90.9%

At the first level of the hierarchical decision tree, best classification accuracy in discriminating normal and abnormal livers (Table 1) was 93.7% employing 2 textural features, namely mean value and angular second moment. In Table 1, the sum of the terms in the first column (Normal) gives the number of livers classified as normal and the second column sum (Abnormal) gives the number of livers classified as abnormal. The row sums Normal and Abnormal give the actual numbers of the normal livers and of the histollogically verified abnormal livers (cirrhosis or fatty infiltration) respectively. Thus, the MLP classifier discriminated correctly 79% of the normal and 100% of the diseased livers. The overall accuracy is then calculated as the percentage of correctly classified by the MLP normal and abnormal livers, i.e. the sum of the diagonal terms in Table 1.

At the second level of the hierarchical tree, the discrimination test on the 55 livers, which were correctly classified as abnormal at the 1st level, gave classification accuracy of 90.9% (Table 2) in correctly discriminating cirrhosis from fatty infiltration, employing 3 textural features, namely mean value, variance, and auto-correlation. Thus, the overall accuracy of the software system in correctly classifying cirrhosis is 94.4% (i.e., 17 cirrhotic livers out of 18 were classified correctly) and fatty liver infiltration is 89.2% (33/37).

At the third level of the hierarchical tree, the classification accuracy in correctly grading cirrhosis or fatty liver infiltration is shown in Tables 3 and 4 respectively. Only correctly classified data at the 2nd level were used. In the case of fatty liver infiltration the total classification success in distinguishing between high and low grading was 84.6% employing three textural features, the mean value, the angular second moment, and the contrast. In the case of cirrhosis (Table 4), low to high grading discrimination was 94.44% employing the following 4 textural features: variance, entropy, sum-entropy, difference entropy.

Table 3: MLP-classifier discrimination accuracy in fatty liver infiltration grading

MLP:	Fatty-high	Fatty-low	Sums	Success
histology				
Fatty-high	11	2	13	84.6%
Fatty-low	3	17	20	85 %
Sums	14	19	33	
		Total Success	84.9 %	

Table 4: MLP-classifier discrimination accuracy in liver cirrhosis grading

MLP:	High	Low	Sums	Success
histology				
High	6	1	7	85.7 %
Low	0	10	10	100.0%
Sums	6	11	17	
		Total Success	94.1 %	

4. Discussion

Determination of liver parenchyma disease is a crucial parameter in patient management. In assessing the type of hepatic disease the patient's history, clinical findings, laboratory tests and ultrasonic findings have to be taken into account. In several cases however, diagnosis is not ensured and patients are subjected to biopsy. Additional information, therefore, could be of assistance in reducing invasive techniques and in patient management. Some of the textural information used in the present work has no obvious relationship to the diagnostic criteria commonly employed by the sonographer in interpreting liver echograms. That textural information is difficult to perceive visually but it gives a mathematical description of additional hepatic textural properties. Thus, textural features sum-entropy and difference-entropy used in distinguishing between low and high cirrhosis grading at the 3rd level of the hierarchical tree, although they quantify image structural information, they cannot be attributed to specific textural characteristics [4]. On the contrary, at the first level of the hierarchical tree the textural features employed were the mean value, signifying liver parenchyma echogenicity, and the angular second moment, describing parenchyma homogeneity [4]. The number of textural features employed at each tree level was the minimum that provided highest classification accuracy. Increasing the number of features did not improve the classification accuracy, probably because of feature intercorrelation.

At each tree level, highest classification accuracy was achieved employing 2 MLP hidden layers but the number of the MLP hidden layer nodes differed among the hierarchical tree levels from 4 to 10. Use of greater numbers of layer nodes at each tree level did not improve classification accuracy while system design complexity increased. Also, in the design of the MLP classifier a maximum of 400 passes over the training set of data was allowed for convergence, since the MLP classifier training procedure proved extremely time consuming.

According to our results, at the first hierarchical tree level the MLP classifier characterized correctly all abnormal livers while it misclassified as abnormal 5 out of 24 normal livers. The performance of the MLP classifier at this level may be evaluated [6] by its sensitivity, which is the percentage of abnormal livers correctly classified (100% in this case), and by its specificity, which is the percentage of normal livers correctly classified (79%). It is important to note that the 93.7% classification accuracy achieved by the MLP classifier is higher than that of other studies using conventional statistical classifiers [1]. High percentages were also reached at the second level between cirrhosis and fatty infiltration (90.0%) with the MLP classifier being more successful in distinguishing cirrhotic livers (94.4%). This may be due to the fact that cirrhosis is a distinct stage of a process caused by hepatitis B or C or alcoholism; liver tissue changes by infiltrating fatty cells, then connective tissue is created finally resulting in cirrhosis. Promising results were also achieved at the third level of the hierarchical tree where hepatic disease grading was evaluated.

5. References

[1] U. Raeth, D. Schlaps, B. Limberg, I. Zuna, A. Lorenz, G. van Kaick, W.J. Lorenz, and B. Kommerell, Diagnostic accuracy of computerized B-scan texture analysis and conventional ultrasonography in diffuse parenchymal and malignant liver disease, *J. Clin. Ultrasound* **13** (1985) 87-99.

[2] A. Khotanzad and J. Liu, Classification of invariant image representations using a neural network, *IEEE Trans. Acoust, Speech, Signal Processing* **38** (1990) 1028-1038.

[3] D. Cavouras and P. Prassopoulos, Computer image analysis of Brain CT images for discriminating hypodense cerebral lesions in children, *Med. Inform.* **19** (1994) 13-20.

[4] R. M. Haralick, K. Shanmugam, and I. Dinstein, Textural features for image analysis, *IEEE Trans. Syst., Man., Cybern,* **SMC-3** (1973) 610-621.

[5] M. M. Galloway, Texture analysis using gray level run lengths, *Comput. Graphics and Image Processing,* **4** (1975) 172-179.

[6] S.G. Morris and J.L. Rose, Application of pattern recognition techniques to breast cancer detection: ultrasonic analysis of 100 pathologically confirmed tissue areas, *Ultrasonic Imaging*, **4** (1982) 378-396.

Address for correspondence:
Prof. D. Cavouras,
37-39 Esperidon Street,
Kallithea 17671, Athens, Greece
e-mail: cavouras @ hol.gr
fax: +301-9594-558 (home), (+301) 5910 975 (work).