

Research Article

Influence of Expert-Dependent Variability over the Performance of Noninvasive Fibrosis Assessment in Patients with Chronic Hepatitis C by Means of Texture Analysis

Cristian Vicas,¹ Monica Lupsor,² Mihai Socaciu,² Sergiu Nedevschi,¹ and Radu Badea²

¹Computer Science Department, Faculty of Automation and Computer Science, Technical University of Cluj-Napoca, 28 Gheoghe Baritiu Street, 400027 Cluj-Napoca, Romania

²Regional Institute of Gastroenterology and Hepatology, Iuliu Hatieganu University of Medicine and Pharmacy, 19-21 Croitorilor Street, 400162 Cluj-Napoca, Romania

Correspondence should be addressed to Cristian Vicas, cristian.vicas@cs.utcluj.ro

Received 1 August 2011; Accepted 26 September 2011

Academic Editor: Carlo Cattani

Copyright © 2012 Cristian Vicas et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Texture analysis is viewed as a method to enhance the diagnosis power of classical B-mode ultrasound image. The present paper aims to evaluate and eliminate the dependence between the human expert and the performance of such a texture analysis system in predicting the cirrhosis in chronic hepatitis C patients. 125 consecutive chronic hepatitis C patients were included in this study. Ultrasound images were acquired from each patient and four human experts established regions of interest. Textural analysis tool was evaluated. The performance of this approach depends highly on the human expert that establishes the regions of interest ($P < 0.05$). The novel algorithm that automatically establishes regions of interest can be compared with a trained radiologist. In classical form met in the literature, the noninvasive diagnosis through texture analysis has limited utility in clinical practice. The automatic ROI establishment tool is very useful in eliminating the expert-dependent variability.

1. Introduction

Noninvasive detection and staging of liver fibrosis have received more and more attention in scientific literature. One approach involves simple B-mode ultrasound in conjunction with textural analysis. The main assumption of the textural analysis approach is that fibrosis alterations at liver lobule level can induce significant changes in the speckle pattern of the ultrasound image [1]. Even if these alterations are not visible with the naked eye, a texture analysis system can detect and learn these alterations. Textural analysis is viewed as a method to enhance the diagnosis power of B-mode ultrasound by providing the physician with new information. This data can be otherwise inferred only by invasive methods.

The methodology presented in most of the papers [1–9] approaching textural analysis on B-mode ultrasound follows four general steps. First, a physician acquires a liver ultrasound image. Then, on the ultrasound image, another physician (or the same) establishes a rectangular region of interest (ROI). In the third step several textural algorithms

produce a feature vector. This vector is labeled according to biopsy findings. The fourth step implies the training of a classification schema. The resulting classifier can be used to predict fibrosis stages to unknown ultrasound images. In the first two steps there is a human expert that introduces an operator-dependent variability.

This paper addresses the user variability introduced by the second step, the establishment of the ROI. We also evaluate here a novel tool that automatically establishes the regions of interest. This tool was developed by our group and it was successfully applied in eliminating the expert-dependence in noninvasive steatosis quantification [10].

To our knowledge, the expert dependent variability in textural analysis for fibrosis detection was not addressed before. We included almost all the textural algorithms proposed in the literature as means of detecting liver fibrosis stages.

Present study aims to evaluate the dependence between the human expert and the performance of the texture analysis system in predicting cirrhosis in chronic hepatitis C patients.

2. Material and Methods

2.1. Patients. The local Ethical Committee of the University of Medicine and Pharmacy Cluj-Napoca approved this study. The patients provided written informed consent before the beginning of the study, in accordance to the principles of the Declaration of Helsinki (revision of Edinburgh, 2000). We prospectively included in this study 125 patients with hepatitis C infection having fibrosis stage 0 or 4 according to Metavir scoring system. Liver biopsy determined the fibrosis stages. This lot was selected from 1200 patients and was prospectively examined in Third Medical Clinic, Cluj-Napoca, Romania, between May 2007 and August 2009. All patients had positive HCV-RNA and underwent percutaneous liver biopsy (LB), in order to stage and grade their condition.

The exclusion criteria were presence of ascites at clinical or ultrasound examination, coinfection with HBV and/or HIV, other active infectious diseases, and pregnancy.

Alongside the epidemiological data, certain biological parameters were determined on a blood sample taken 12 hours after overnight fasting: alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transferase (GGT), total cholesterol, triglycerides, total bilirubin, and glycemia (Konelab 20i—Thermo Electron Corp., Finland).

2.2. Histopathological Analysis. A liver biopsy specimen was acquired using the TruCut technique with an 1.8 mm (14 G) diameter automatic needle device—Biopty Gun (Bard GMBH, Karlsruhe, Germany). The LB specimens were fixed in formalin and embedded in paraffin. The slides were evaluated by a single expert pathologist unaware of the clinical data. Only biopsy specimens with more than 6 intact portal tracts were eligible for evaluation [11]. The liver fibrosis and necroinflammatory activity were evaluated semiquantitatively according to the Metavir scoring system [12].

Fibrosis was staged on a 0–4 scale as follows: F0—no fibrosis; F1—portal fibrosis without septa; F2—portal fibrosis and few septa; F3—numerous septa without cirrhosis; F4—cirrhosis. The necroinflammatory activity was graded as A0—none; A1—mild; A2—moderate; A3—severe.

In present study, only patients having fibrosis stage 0 or 4 were included.

2.3. Ultrasound Examination. Each patient included in this study underwent an ultrasound examination using a GE Logiq 7 ultrasound machine (General Electric Company, Fairfield, England) with a 5.5 MHz convex phased array probe one day prior to liver biopsy. From each patient there were acquired right lobe ultrasound images with liver tissue without blood vessels or other artifacts with a depth setting of 16 cm using the same preestablished machine protocol. The acquisition protocol was established in such a way that we obtained a maximum amount of information from underlying tissue and in the same time keeping the noise level down. All postprocessing settings were set to minimum.

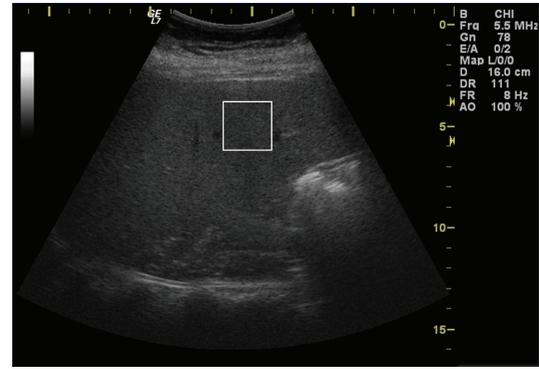


FIGURE 1: Right lobe ultrasound image. White square represents the region of interest.

The frame rate was kept as high as possible in order to avoid movement artifacts. The time gain compensation curve was set to neutral position. Once the device settings were established they were used to examine all the patients. Captured images were saved in DICOM format on the equipment's local hard drive. They were later transferred and processed on a personal computer.

2.4. Regions of Interest for Textural Analysis. The region of interest (ROI) establishment procedures followed the guidelines presented in the literature [1, 13]. The experts were instructed to choose one region of interest for each patient. The ROI had to be placed as close as possible to the vertical axis of the ultrasound image and at 1 cm below the liver capsule. The ROI had to avoid artifacts and anatomical features like blood vessels, liver capsule, shadowing, and so forth. The dimensions of the ROI were 64×64 pixels representing an area of 2.62×2.62 cm. Figure 1 shows an ultrasound image with an ROI. The physician acquired the image from the right live lobe.

In order to evaluate the user variability of the textural system, on the saved images, four experts with different skill level established the ROIs. The first two experts are trained radiologists with experience in gastrointestinal ultrasound investigation. First expert has more than 20 years in ultrasound investigation and the second more than 10 years. The third expert is a radiology intern with 2 years of experience. The fourth expert is a general practitioner trained in ultrasound examination. In addition to these experts, we employed an automated tool for establishing the region of interest. This tool establishes the ROI in a fixed position relative to the geometry of the image. Artifacts are detected using the method proposed in [10, 14]. After the artifacts are detected, we randomly choose a region of interest that has no artifacts. If such a region cannot be set in any of the patient's images, for the respective patient there will be no region of interest established.

The order of the patients and the order of the images for a patient were randomized. With this step we tried to avoid the influence of the image order over the performance detection. Algorithm 1 was used to ensure independent samples, it is graphically depicted in Figure 2.

Input: **patients**—a set with patients and ultrasound images
 Output: **DS**—a list with 25 sets DS_{ij} with regions of interest
 For $i = 1$ to 5
 D_i = Perform a randomization on **patients**
 For $j = 1$ to 5
 Sequentially present to expert j dataset D_i
 DS_{ij} = established ROIs

ALGORITHM 1: ROI establishment.

We computed the center of each region of interest in terms of Cartesian and polar coordinates. For the Cartesian system, the origin is the top left corner and for the polar system, the origin was considered the virtual source of ultrasound waves. Figure 3 sketches these coordinate systems.

2.5. Textural Analysis. In texture analysis there are two main steps [15]. The first step is the computation of several textural attributes that numerically describe the texture (using dedicated algorithms). The second step involves the training and evaluation of a classifier using the previously computed textural features.

Each texture description algorithm has a certain number of parameters that control the feature extraction process. For each algorithm implemented in the present study we used the same proposed set of parameters found in corresponding fibrosis detection papers. These algorithms are first-order statistics [4, 16], gray tone difference matrix [15], gray level co-occurrence matrix [1, 4, 16, 17], multiresolution fractal dimension [1], differential box counting [6, 18], morphological fractal dimension estimators [19], Fourier power spectrum [1, 13], Gabor filters [20], Law's energy measures [1], texture edge co-occurrence matrix [6], phase congruency-based edge detection [21], and texture feature coding matrix [22].

These 12 algorithms processed the entire ROI and computed 234 features per patient. Each feature vector was labeled with the corresponding histopathological finding as healthy or cirrhotic. From 25 sets of regions of interest we generated 25 sets of instances, each set containing one instance per patient.

The classification schema employed here was a logistic model [23–25]. The feature values were normalized in $[0, 1]$ interval prior to classification. Care was taken that the test subset was normalized with the same coefficients as the train set.

Before entering the classification schema, a feature selection process was applied. The relevant features were identified and selected using *correlation-based feature selection* (CFS) algorithm [26]. To avoid *overfitting* phenomena and to ensure that the feature selection step is independent of the underlying data, the following algorithm was applied.

- (1) From each of the 25 sets we selected k instances. These instances were randomly selected in such a way that each class has $k/2$ instances.
- (2) The selected instances were moved into another dataset.

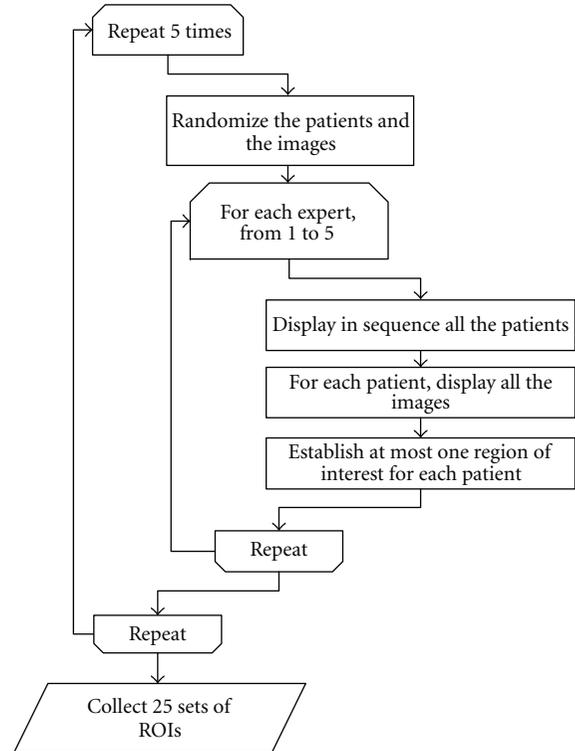


FIGURE 2: Algorithm for establishing the regions of interest. Each of 5 experts established 5 sets of ROIs. The automatic establishment algorithm was treated as a regular expert.

- (3) After 25 iterations we extracted $25 \times k$ instances.
- (4) On this $25 \times k$ dataset we applied the CFS algorithm. We noted the selected features and we processed the original datasets by keeping only relevant features.
- (5) The whole process was iterated 20 times.

For this paper, k was set to 10. The feature selection process is depicted in Figure 4.

The classifier performance estimation was determined using 10-fold stratified CV technique. The performance criterion was area under the curve (AUROC) computed on the collected predictions using Mann-Whitney-Wilcoxon U statistic [27]. In order to better estimate the average performance, the 10-fold CV procedure was iterated 10 times with random fold splitting [28].

The texture analysis system was validated using a set of known textures from Brodatz [29] library. Each image was divided into 100 nonoverlapping regions of interest. Each region has 64×64 pixels area. The textural analysis system was trained to predict the original image from where the region originated. The images were chosen following the guidelines in [15].

2.6. Statistical Analysis. Two-way ANOVA test was used to evaluate the performance variability. The dependent variable was set to be the average AUROC and the independent variables were the expert that established ROIs and the feature set obtained after the feature selection step. Tukey

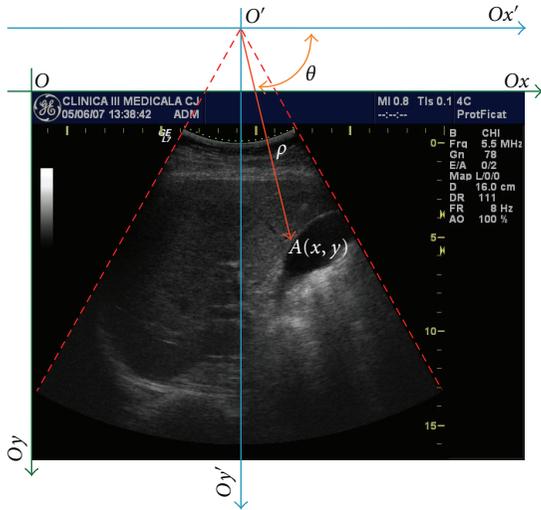


FIGURE 3: Cartesian system, Oxy (green lines) and polar system $O'x'y'$ (blue lines). An ROI center (A) has the $A(x, y)$ Cartesian coordinates and $A(\rho, \theta)$ polar coordinates. The green area represents the position of piezoelectric crystals in the probe and the red lines show the imaging aperture of the convex probe. The depth of the imaging system was set to 16 cm.

post hoc analysis was used to identify the source of variation when the ANOVA test was statistically relevant.

When the assumption of normal distribution with equal variances could not be met we used Kruskal-Wallis one way analysis of variance. The significance threshold was set to $P = 0.01$. In addition to the expert quality we investigated the impact of the ROI position relative to the geometry of the image. We computed the Pearson correlation coefficient between the ROI position and the detection performance for each expert and iteration.

Textural algorithms were implemented in a custom-made software system developed at Technical University of Cluj Napoca, Romania. Classification schema used the LibSVM implementation [30] (public domain, ver. 2.89) integrated in weka framework [23] (public domain, ver. 3.7). Statistical analysis was performed in R (public domain, ver. 2.10).

3. Results

The texture analysis system was validated using three sets of images. First dataset contained regions from D77, D84, D55, D53, and D24 Brodatz [29] textures. Second dataset consisted of D4 and D84 textures. The third set had regions from D5 and D92. The classification accuracy was 98.9 for the first set, 98.4 for the second set, and 97.9 for the third set.

Clinical and biochemical characteristics of the study patients are summarized in Table 1. The median length of the LB samples was 11.38 mm, and the mean number of portal spaces was 11.6. The fibrosis stage distribution in our patients was as follows: F0—51 (40.8%) and F4—74 (59.2%).

Each expert was instructed to select one region of interest for each patient. The process was iterated five times. Expert 1 established in average 121.6 regions (min = 121, max = 122), expert 2—120.8 (115–123), expert 3—122 (122–122), and

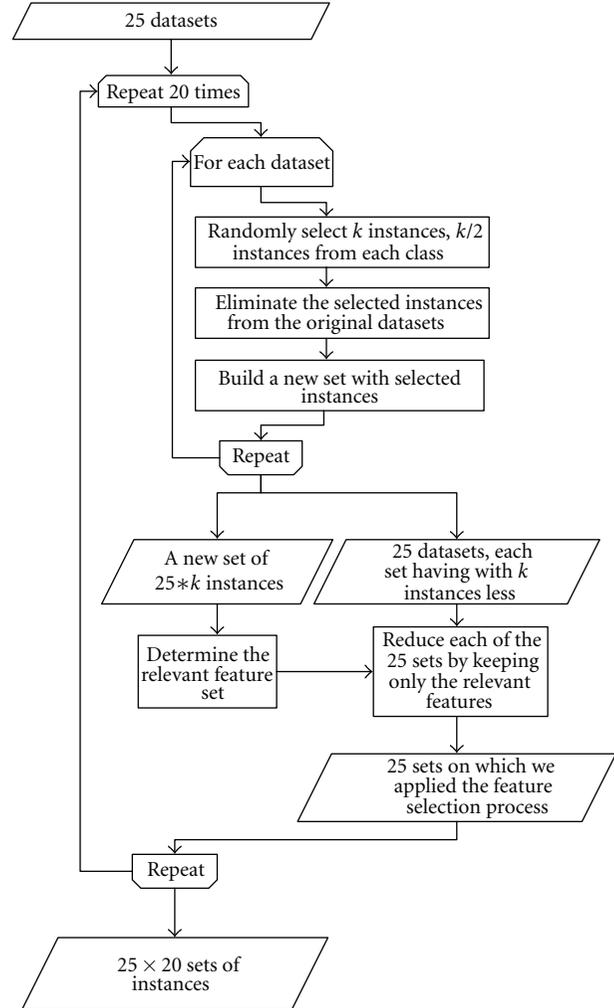


FIGURE 4: Relevant feature selection. To ensure that the selection process is not data dependent, a small number of instances were extracted from each dataset.

expert 4—113 (112–115). The automatic ROI establishment algorithm (expert 5) established 83 images (83–83). There were three patients that had poor quality images and no physician was able to establish an ROI. Two were healthy patients and one was cirrhotic.

We recorded the mean and standard deviation AUROC for each of the experts: expert 1— 0.618 ± 0.059 , expert 2— 0.611 ± 0.085 , expert 3— 0.537 ± 0.062 , expert 4— 0.528 ± 0.075 , and expert 5— 0.611 ± 0.074 .

We investigated the role of feature selection and the user expertise in the performance of the system using two way ANOVA. The only relevant factor was the human expert ($P < 0.0001$) as shown in Figure 5. The other factor, feature selection, was not relevant ($P = 0.8$). In Figure 6 are shown the corresponding box plots.

Post hoc analysis using Tukey method revealed that the differences between experts are significant ($P < 0.001$) with several exceptions, the difference between the expert 1 and 2 and the difference between expert 1 and 5. Note that the expert 5 is the automatic ROI establishment algorithm.

TABLE 1: Characteristics of the study group.

Characteristics of patients	Entire lot	Patients with fibrosis stage 0 Mean \pm SD (interval or %)	Patients with cirrhosis
Number	125 (100%)	51 (40.8%)	74 (59.2%)
Sex (male)	50 (40%)	16 (31.4%)	34 (45.9%)
Age (years)	47.45 \pm 12.13 (22–77)	53.39 \pm 8.93 (33–77)	38.82 \pm 10.97 (22–66)
BMI (kg/m ²)	26.41 \pm 5.15 (18.56–46.48)	28.29 \pm 5.33 (18.83–46.48)	23.9 \pm 3.65 (18.56–33.87)
AST (U/I)	58.54 \pm 47.67 (12–387)	82 \pm 49.57 (23–387)	25.79 \pm 13.47 (12–71)
ALT (U/I)	75.68 \pm 55.66 (8–270)	102.25 \pm 53.94 (21–270)	38.58 \pm 31.87 (8–163)
GGT (U/I)	77.83 \pm 107.77 (13–993)	105.47 \pm 133.33 (27–993)	39.83 \pm 28.13 (13–130)
Total bilirubin (mg/dL)	0.88 \pm 0.64 (0.27–4.27)	1.09 \pm 0.73 (0.4–4.27)	0.59 \pm 0.28 (0.27–1.72)
Alkaline phosphatase (U/I)	263.13 \pm 188.34 (127–1781)	286.98 \pm 215.81 (127–1781)	201.5 \pm 45.61 (142–307)
Glucose (mg/dL)	106.73 \pm 27.75 (72–266)	113.81 \pm 32.78 (72–266)	96.86 \pm 13.72 (72–129)
Cholesterol (mg/dL)	195.29 \pm 45.8 (97–331)	174.22 \pm 36.31 (97–299)	223.83 \pm 41.92 (149–331)
Triglycerides (mg/dL)	124.11 \pm 57.67 (51–349)	123.85 \pm 50.08 (53–316)	124.46 \pm 67.16 (51–349)
Platelet count (10 ⁹ /L)	166.06 \pm 70.32 (42–373)	142.81 \pm 65.35 (42–373)	226.52 \pm 40.94 (151–314)
INR	1.12 \pm 0.2 (0.83–1.84)	1.17 \pm 0.2 (0.89–1.84)	0.99 \pm 0.12 (0.83–1.3)
Right lobe images per patient	12.97 \pm 6.06 (2–33)	13.02 \pm 5.03 (3–24)	12.94 \pm 6.69 (2–33)

Abbreviations: body mass index (BMI), aspartate aminotransferase (AST), alanine aminotransferase (ALT), gamma-glutamyl-transpeptidase (GGT), and international normalized ratio (INR).

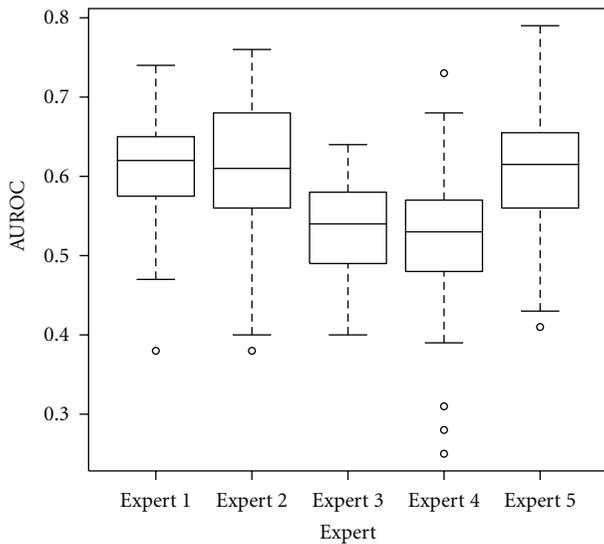


FIGURE 5: Box plot representing the dependency between the estimated performance and the expert that established the regions of interest. The top and the bottom of the boxes are the first and third quartiles, respectively. Thus, the length of the box represents the interquartile range within which 50% of the values were located. The line through the middle of each box represents the median.

In practice, a classifier is trained with data gathered from an expert but it can be used by other physicians. We identified two cases. First case, the expert that trained the classifier uses it in the current practice. In this scenario, the same expert that first established the ROIs establishes the ROIs for the new, unknown images. In the second scenario the expert that establishes the ROIs on the new images is different from the initial expert.

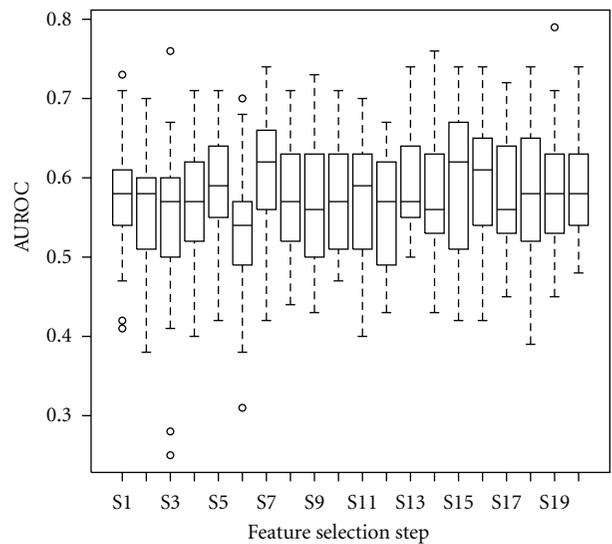


FIGURE 6: Box plot representing the dependency between the estimated performance and the feature selection process. Each label on the horizontal axis represents a separate feature selection step. The top and the bottom of the boxes are the first and third quartiles, respectively. Thus, the length of the box thus represents the interquartile range within which 50% of the values were located. The line through the middle of each box represents the median.

The first scenario was simulated here by training a classifier with each dataset from each expert. Resulting classifier was evaluated using the other datasets from the same expert obtained at different ROI establishment step. Kruskal-Wallis test revealed that there is a significant variation due to the human expert ($P < 0.001$), as seen in Figure 7.

Again, most experienced experts provided best performance. During this test we ignored the results from expert 5.

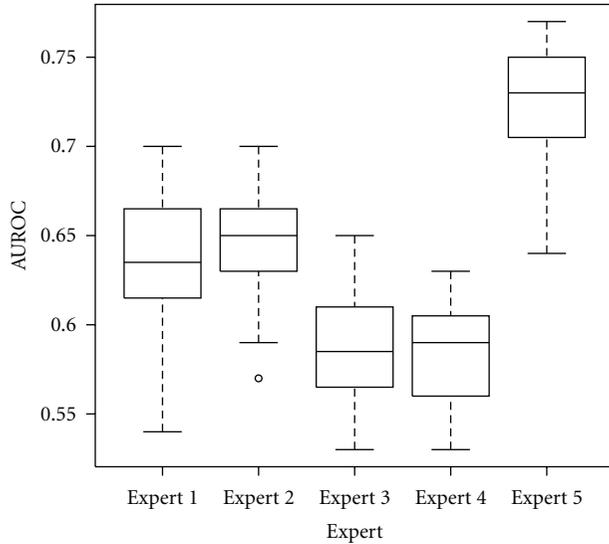


FIGURE 7: Box plot representing the estimated performance obtained when the same expert trains and uses the texture analysis tool in clinical practice. The top and the bottom of the boxes are the first and third quartiles, respectively. Thus, the length of the box thus represents the interquartile range within which 50% of the values were located. The line through the middle of each box represents the median.

Because this expert establishes the regions in the same position, deciding only to accept or reject an image, we noted that there is a significant subset of images that are always selected by this algorithm in all 5 iterations. This subset positively biases the performance evaluation in the case of expert 5, because one will find identical samples in the train and test set. The same analysis applied on the human experts revealed that few images were common between the ROI establishment iterations.

In the second scenario, the expert who uses the non-invasive tool is different from the expert that provided the training data for the system. We trained the classifier with the data collected from one expert and then test it with the data collected from the other experts. Kruskal-Wallis analysis revealed an interesting fact; there is no significant variance due to experts ($P = 0.0506$) as shown in Figure 8. In both scenarios the analysis did not revealed significant variance due to the feature selection step.

In the following we investigated the impact of the ROI position relative to the image geometry. The center coordinates of the ROIs were converted to polar space. The center of the polar space was set to be the virtual source of ultrasounds. For each ROI the angle θ and the vector length, ρ , were computed. For each expert and iteration we computed the mean angle and length. A linear regression was performed between these coordinates and the mean performance of the expert i during iteration j . We computed the Pearson correlation coefficient and its relevance. In Figures 9 and 10 are the shown the results.

The correlation coefficients were -0.44 (between ρ and AUROC) and -0.48 (between θ and AUROC). This correlation is not statistically significant for the chosen

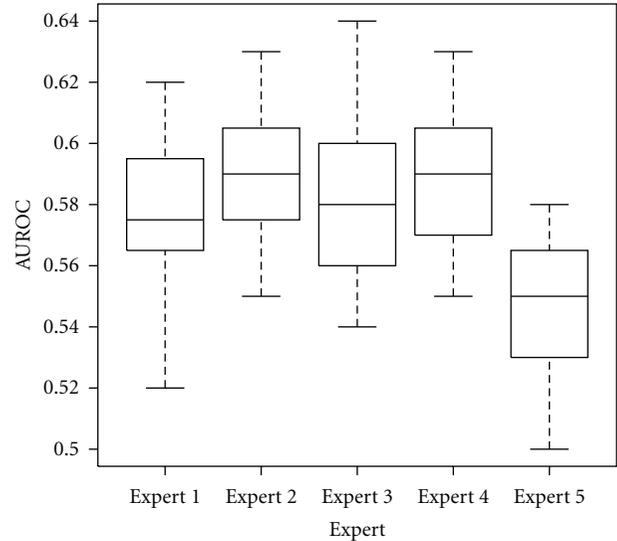


FIGURE 8: Box plot representing the estimated performance obtained when the texture analysis system is trained with datasets provided by one expert and used with ROIs established by a different expert. The top and the bottom of the boxes are the first and third quartiles, respectively. Thus, the length of the box thus represents the interquartile range within which 50% of the values were located. The line through the middle of each box represents the median.

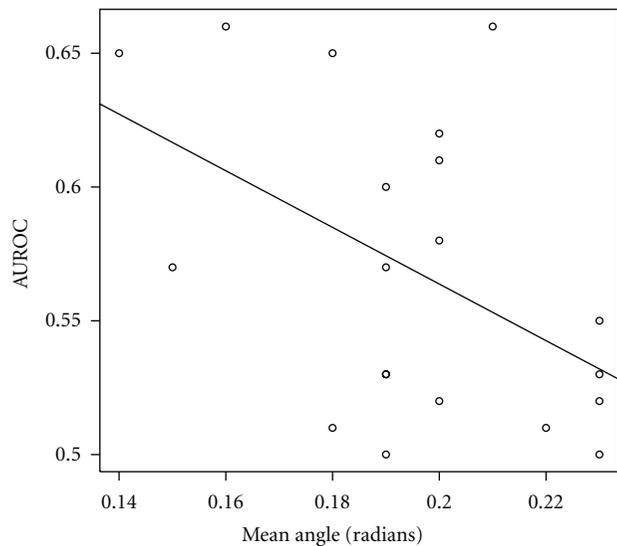


FIGURE 9: The dependence of the estimated performance in relation to the θ coordinate of the ROI's center.

threshold. However, it is possible that a link exists between the ROI position and the classification performance because the results became relevant for a higher threshold ($P < 0.05$).

We also compared the mean positions of the ROIs when expressed in simple Cartesian coordinates. For each ROI the center coordinates were computed relative to the top left corner of the image. One-way ANOVA showed that the O_x (horizontal) coordinate is not relevant but for the O_y (vertical) coordinate, higher performances were obtained for the regions that were established closer to the upper part of

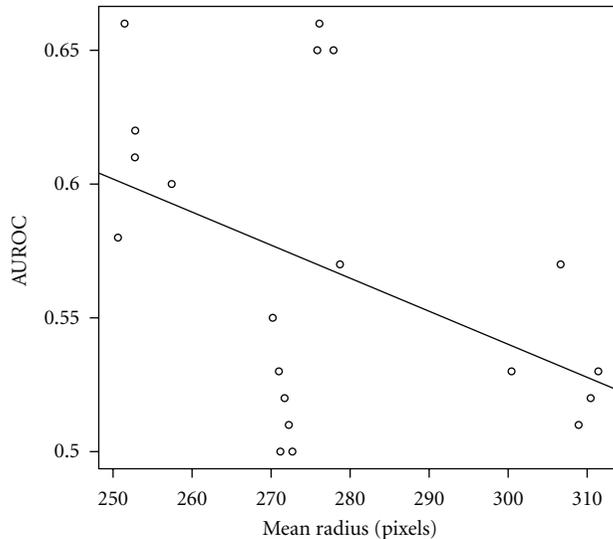


FIGURE 10: The dependence of the estimated performance in relation to the ρ coordinate of the ROI's center.

the image. Figure 11 shows the BOXPLOT graph. Again the expert 5 was ignored because the automatic algorithm always established the regions in the same position.

4. Discussions

Liver biopsy is an imperfect golden standard in fibrosis staging. It is an invasive procedure and even if the method allows direct examination of the liver tissue there is a certain variability due to the reduced tissue volume and due to the fact that a human expert qualitatively evaluates the biopsy [31–33].

There are numerous research directions involving non-invasive fibrosis staging and noninvasive diagnosis of liver diseases in general [34, 35]. Papers [8, 21, 22, 36, 37] studying texture analysis as a noninvasive staging tool reported high performances in cirrhosis detection [36] and even in fibrosis staging [8]. In these papers there are variations in terms of studied pathology and classification evaluation methodology. We believe that these factors might have positively biased the results reported by other authors.

Present study aims to evaluate the dependence between the human expert and the performance of such a texture analysis system in predicting the cirrhosis in chronic hepatitis C patients. In the same time the present paper brings the following contributions to the noninvasive fibrosis detection field: it includes only patients with chronic hepatitis C, excluding other pathologies; it integrates almost all textural algorithms met in fibrosis detection and it proposes a more rigorous performance evaluation methodology that gives results closer to the real performance of a classifier.

In present study we included only patients with chronic hepatitis C etiology. Other papers that address the noninvasive detection of cirrhosis include patients having different pathologies like fatty infiltration [16]. Another important highlight of this paper is the volume of patients. There are few papers that study more than 100 patients but not all the

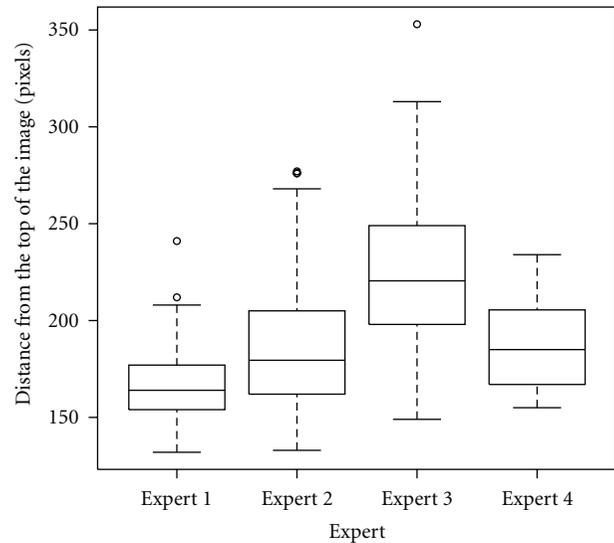


FIGURE 11: Box plot representing the O_y coordinate distribution. The top and the bottom of the boxes are the first and third quartiles, respectively. Thus, the length of the box thus represents the interquartile range within which 50% of the values were located. The line through the middle of each box represents the median.

patients included in these studies have chronic hepatitis C, or the etiology for cirrhosis is not specified [4, 16, 38].

Performance estimation algorithm proposed in this paper ensures that each time the classifier is tested the test data are new and unseen at the training or feature selection phase. The metaparameter sets are evaluated on unseen data to ensure that we do not select a classifier instance that overfits the training data. The cross-validation loop ensures that even this search procedure does not overfit the data. The 10-time repetition of the evaluation phase ensures a better estimation of the mean performance. No other papers employed repeated performance estimation on their classification schemas. When performing one iteration the data might get partitioned in such a way that by accident the performance estimation is very high. For example, in some iterations the performance reached levels as high as 0.79. Of course, the mean performance estimated over 10 runs is smaller. The same phenomenon of increased variance can be noted when the performance measure is computed on each test fold and not on the entire prediction vector. In 2-fold CV a “lucky” splitting might give a very high performance reading.

In present paper, the CV predictions are collected and the performance is measured on a vector that has the same dimensions as the initial dataset.

Textural feature selection is performed on an independent dataset. This dataset is obtained by randomly sampling the original datasets. It is important to note that each instance that is included in the feature selection dataset is excluded from the original dataset. As a result, the feature selection process has less chances of overfitting.

The particular set of features does not influence the detection rates. The subset of features selected at each step has a high variability. High ranking features cover large

spectra of algorithms, from statistical algorithms to multiresolution analysis. This indicates that the specific algorithm used to numerically describe the texture has its importance but there are fewer chances that new textural algorithms will make a great impact over cirrhosis detection and fibrosis staging.

The design of the experiment, where each expert establishes 5 sets of ROIs ensures that the samples are independent and normally distributed. Each set of patients have different randomizations in order to minimize the effect of patient/image succession over the experiment. Moreover, for each patient the order of the images is altered. It is important to note that the order of images is the same for all the experts. Expert x viewed the patients and images in the same order as the expert y when establishing ROI for the same dataset z .

The main finding of this paper is that the performance of the studied software diagnosis tool depends on the expert that employs this tool. In the results section we have shown that there is a significant performance variation between experts. The results presented here showed that more experienced experts tend to capture the same aspects of the ultrasound image, aspects that are consistent with the histological findings. If this tool is trained and employed by an experienced physician it might give some extra information about the underlying pathology.

The results from the second scenario, when the expert that uses the texture analysis tool is different from the expert that provided the data for training, revealed the fact that there is little use for texture analysis tool in screening processes.

The classical methodology has a severe drawback. It requires a human expert to establish a representative area where the texture will be analyzed. Replacing the human expert with a computerized solution improves the usefulness of such a software analysis tool. The results shows that such a tool can have a performance similar to a highly trained expert. This result is another important contribution of this paper to the noninvasive diagnosis field.

5. Conclusions

Texture analysis can enhance the diagnosis power of the B-mode ultrasound image. The performance of this approach depends highly on the human expert that establishes the regions of interest. In classical form met in the literature noninvasive diagnosis through texture analysis has limited utility in clinical practice. Further work in this domain has to be focused in finding another noninvasive descriptors for fibrosis.

Acknowledgments

The authors would like to thank Vlad Popovici from Swiss Institute of Bioinformatics for the priceless help in designing the experiments presented here. All authors contributed equal parts to this paper.

References

- [1] C. M. Wu, Y. C. Chen, and K. S. Hsieh, "Texture features for classification of ultrasonic liver images," *IEEE Transactions on Medical Imaging*, vol. 11, no. 2, pp. 141–152, 1992.
- [2] K. J. W. Taylor, C. A. Riely, and L. Hammers, "Quantitative US attenuation in normal liver and in patients with diffuse liver disease: importance of fat," *Radiology*, vol. 160, no. 1, pp. 65–71, 1986.
- [3] B. J. Oosterveld, J. M. Thijssen, P. C. Hartman, R. L. Romijn, and G. J. E. Rosenbusch, "Ultrasound attenuation and texture analysis of diffuse liver disease: methods and preliminary results," *Physics in Medicine and Biology*, vol. 36, no. 8, pp. 1039–1064, 1991.
- [4] Y. M. Kadah, A. A. Farag, J. M. Zurada, A. M. Badawi, and A. B. M. Youssef, "Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images," *IEEE Transactions on Medical Imaging*, vol. 15, no. 4, pp. 466–478, 1996.
- [5] D. Gaitini, Y. Baruch, E. Ghersin et al., "Feasibility study of ultrasonic fatty liver biopsy: texture vs. attenuation and backscatter," *Ultrasound in Medicine and Biology*, vol. 30, no. 10, pp. 1321–1327, 2004.
- [6] G. T. Cao, P. F. Shi, and B. Hu, "Liver fibrosis identification based on ultrasound images captured under varied imaging protocols," *Journal of Zhejiang University*, vol. 6, no. 11, pp. 1107–1114, 2005.
- [7] D. Gaitini, M. Lederman, Y. Baruch et al., "Computerised analysis of liver texture with correlation to needle biopsy," *Ultraschall in der Medizin*, vol. 26, no. 3, pp. 197–202, 2005.
- [8] H. Yamada, M. Ebara, T. Yamaguchi et al., "A pilot approach for quantitative assessment of liver fibrosis using ultrasound: preliminary results in 79 cases," *Journal of Hepatology*, vol. 44, no. 1, pp. 68–75, 2006.
- [9] J. W. Jeong, S. Lee, J. W. Lee, D. S. Yoo, and S. Kim, "The echotextural characteristics for the diagnosis of the liver cirrhosis using the sonographic images," in *Proceedings of the 29th Annual International Conference of Engineering in Medicine and Biology Society (EMBS '07)*, pp. 1343–1345, August 2007.
- [10] C. Vicas, S. Nedevschi, M. Lupsor, and R. Badea, "Automatic detection of liver capsule using Gabor filters. Applications in steatosis quantification," in *Proceedings of the IEEE 5th International Conference on Intelligent Computer Communication and Processing (ICCP '09)*, pp. 133–140, August 2009.
- [11] P. Bedossa and T. Poynard, "An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group," *Hepatology*, vol. 24, no. 2, pp. 289–293, 1996.
- [12] P. Bedossa, P. Bioulac-Sage, P. Callard et al., "Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C," *Hepatology*, vol. 20, no. 1 I, pp. 15–20, 1994.
- [13] C. Abe, C. E. Kahn Jr., K. Doi, and S. Katsuragawa, "Computer-aided detection of diffuse liver disease in ultrasound images," *Investigative Radiology*, vol. 27, no. 1, pp. 71–77, 1992.
- [14] C. Vicas, M. Lupsor, R. Badea, and S. Nedevschi, "Detection of anatomical structures on ultrasound liver images using Gabor filters," in *Proceedings of the 17th IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR '10)*, pp. 249–253, Romnia, Cluj Napoca, May 2010.
- [15] A. Materka and M. Strzelecki, *Texture Analysis Methods: A Review*, Technical University of Lodz, Institute of Electronics, 1998.

- [16] A. M. Badawi, A. S. Derbala, and A. B. M. Youssef, "Fuzzy logic algorithm for quantitative tissue characterization of diffuse liver diseases from ultrasound images," *International Journal of Medical Informatics*, vol. 55, no. 2, pp. 135–147, 1999.
- [17] W. C. Yeh, S. W. Huang, and P. C. Li, "Liver fibrosis grade classification with B-mode ultrasound," *Ultrasound in Medicine and Biology*, vol. 29, no. 9, pp. 1229–1235, 2003.
- [18] W. L. Lee, K. S. Hsieh, Y. C. Chen, and Y. C. Chen, "A study of ultrasonic liver images classification with artificial neural networks based on fractal geometry and multiresolution analysis," *Biomedical Engineering*, vol. 16, no. 2, pp. 59–67, 2004.
- [19] Y. Xia, D. G. Feng, and R. Zhao, "Morphology-based multifractal estimation for texture segmentation," *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 614–623, 2006.
- [20] A. Ahmadian, A. Mostafa, M. D. Abolhassani, and Y. Salimpour, "A texture classification method for diffused liver diseases using Gabor wavelets," in *Proceedings of the 27th Annual International Conference of the Engineering in Medicine and Biology Society (EMBS '05)*, pp. 1567–1570, September 2005.
- [21] G. Cao, P. Shi, and B. Hu, "Ultrasonic liver characterization using phase congruency," in *Proceedings of the 27th Annual International Conference of the Engineering in Medicine and Biology Society (EMBS '05)*, pp. 6356–6359, September 2005.
- [22] M. H. Horng, Y. N. Sun, and X. Z. Lin, "Texture feature coding method for classification of liver sonography," *Computerized Medical Imaging and Graphics*, vol. 26, no. 1, pp. 33–42, 2002.
- [23] I. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, 2005.
- [24] J. Friedman, T. Hastie, and R. Tibshirani, *The Elements of Statistical Learning*, Springer Series in Statistics, Springer, 2001.
- [25] N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," *Machine Learning*, vol. 59, no. 1-2, pp. 161–205, 2005.
- [26] M. Hall, *Correlation-Based Feature Selection of Discrete and Numeric Class Machine Learning*, University of Waikato, 1998.
- [27] M. Pepe, *The Statistical Evaluation of Medical Tests for Classification and Prediction*, Oxford University Press, Oxford, UK, 2003.
- [28] V. Popovici, "Assessment of classification models for medical applications," in *Proceedings of the Workshop on Computers in Medical Diagnoses*, Cluj-Napoca, Romania, 2009.
- [29] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, Dover, New York, NY, USA, 1966.
- [30] LIBSVM: a library for support vector machines, 2001, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [31] A. Regev, M. Berho, L. J. Jeffers et al., "Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection," *American Journal of Gastroenterology*, vol. 97, no. 10, pp. 2614–2618, 2002.
- [32] P. Bedossa, D. Dargere, and V. Paradis, "Sampling variability of liver fibrosis in chronic hepatitis C," *Hepatology*, vol. 38, no. 6, pp. 1449–1457, 2003.
- [33] M. Guido and M. Ruge, "Liver biopsy sampling in chronic viral hepatitis," *Seminars in Liver Disease*, vol. 24, no. 1, pp. 89–97, 2004.
- [34] I. N. Guha and W. M. Rosenberg, "Noninvasive assessment of liver fibrosis: serum markers, imaging, and other modalities," *Clinics in Liver Disease*, vol. 12, no. 4, pp. 883–900, 2008.
- [35] S. Bonekamp, I. Kamel, S. Solga, and J. Clark, "Can imaging modalities diagnose and stage hepatic fibrosis and cirrhosis accurately?" *Journal of Hepatology*, vol. 50, no. 1, pp. 17–35, 2009.
- [36] G. T. Cao, P. F. Shi, and B. Hu, "Liver fibrosis identification based on ultrasound images," in *Proceedings of the 27th Annual International Conference of the Engineering in Medicine and Biology Society (EMBS '05)*, vol. 1–7, pp. 6317–6320, September 2005.
- [37] A. Mojsilovic, M. Popovic, S. Markovic, and M. Krstic, "Characterization of visually similar diffuse diseases from b-scan liver images using nonseparable wavelet transform," *IEEE Transactions on Medical Imaging*, vol. 17, no. 4, pp. 541–549, 1998.
- [38] A. Szebeni, G. Tolvaj, and A. Zalatnai, "Correlation of ultrasound attenuation and histopathological parameters of the liver in chronic diffuse liver diseases," *European Journal of Gastroenterology and Hepatology*, vol. 18, no. 1, pp. 37–42, 2006.



Hindawi
Submit your manuscripts at
<http://www.hindawi.com>

