Fibrosis detection from ultrasound imaging. The influence of necro-inflammatory activity and steatosis over the detection rates.

Cristian Vicas, Sergiu Nedevschi,
Technical University of Cluj-Napoca
{cristian.vicas,sergiu.nedevschi}@cs.utcluj.ro

Monica Lupsor, Radu Badea
3'rd Medical Clinic Cluj Napoca
{mmlupsor, rbadea2003}@yahoo.com

Abstract

Diagnosing liver fibrosis using non invasive procedures is challenging because the visual aspects in US imaging between healthy and fibrosis liver are very much alike. In this paper texture analysis (texture feature computation and texture classification) are employed in order to increase the diagnosis value of US examination. An overview on fibrosis detection is necessary in order to determine and evaluate the best approach. Biopsy and METAVIR score are used to assess the liver pathology. The influence of steatosis and necro-inflamatory activity over the fibrosis detection is also investigated.

Four feature selection methods based on gain ratio, chi squared statistic, correlation and symmetrical uncertainty are evaluated. The results show that fibrosis, steatosis and activity alter the US image texture and implicitly the texture features.

It seems that the best approach in liver fibrosis identification is to build imagistic models for each fibrosis grade.

1. Introduction

The fibrosis is the scarring response formed in the chronic injury of any cause. It is a dynamic process, with a possibility of reversibility. For the moment, the golden standard in evaluating fibrosis is the liver biopsy. Using the liver biopsy one can establish with certainty the diagnosis, one can assess the severity of necroinflamation and fibrosis and one can distinguish the simultaneous liver diseases. On the other hand, it is an invasive procedure, with possible side-effects. [1],[2]

An alternative examination procedure is ultrasound imaging. In fibrosis the visual aspects of healthy/affected liver are very similar so, the diagnosis value of the ultrasound imaging is relatively low in these diseases. Image processing techniques are used to study ultrasound images and to improve the diagnosis value of ultrasound in diffuse liver diseases.

Features are computed over image textures and the images are classified based on the computed feature vectors.

One fractal dimension based algorithm (local multifractal morphological exponents) is used for the first time in fibrosis level detection field.

This paper is structured in the following manner: in chapter 2 are presented shortly the feature computation algorithms and the support vector machine classifier, in chapter 3 are presented some experimental results and chapter 4 contains discussions and conclusions.

2. Texture features and classification algorithm

Eight feature computation algorithms are employed: histogram statistics, grey tone difference matrix, grey level co-occurrence matrix, multifractal differential box counting, morphological multifractal exponents, multi resolution fractal dimension, Law's energy measures and wavelet transform.

Most of these algorithms are common in texture analysis field so they will be presented briefly.

2.1. First and second order statistics

Histogram statistics.[3]

The shape of the gray level histogram can provide clues about the image. Central moments are derived from the histogram in order to characterize the texture. These are mean, variance, skewness, kurtosis, energy and entropy.

Grey-Tone difference matrix [4]. A Grey-Tone Difference Matrix a column vector containing G elements where G is the total number of gray levels. Its entries are computed measuring the difference between the intensity level of a pixel and the average intensity computed over a square window centered at the pixel.

Five different features were derived from the GTDM, to quantitatively describe such perceptual texture properties as [4]:coarseness, contrast, busyness, complexity and texture strength.

Grey-Level co-occurence matrix [4]. The algorithm builds a co-occurrence matrix and computes various statistics over this matrix [5].

We compute features like angular second moment, contrast, inverse difference, entropy and correlation.

2.2. Fractal based features

Multifractal Differential Box Counting. There are several methods available to estimate the multifractal dimension of an image. One of the commonly used methods [6] proposed by Chaudhuri and Sarkar [7] is based on the differential box-counting (DBC) algorithm. A 2D gray scale image is treated as a 3D surface [8]. The area of this surface is estimated at different scales. These measures are obtained by means of counting the minimum number of boxes of different sizes, needed to cover the surface.

In order to describe the distribution of different subfractals, a measure $\mu_\varepsilon(i,j)$ is defined on the box grid [9] Equation 1:

$$\mu_\varepsilon = \frac{n_\varepsilon(i,j)}{\widetilde{N}_\varepsilon}$$

Equation 1 μ_ε measure. N_ε is the total number of boxes and n_ε is the number of boxes in *(i,j)* grid.

The partition and estimation are performed for different scales, and the multifractal dimension of order q can be estimated using Equation 2

$$D_q = \frac{1}{1-q}\lim_{\varepsilon\to 0}\frac{\ln\left[\sum_{i,j}\mu_\varepsilon(i,j)^q\right]}{\ln\left(\frac{1}{r}\right)}, q \neq 1$$

Equation 2. Multifractal measure. r=box scaling factor.

Morphological Multifractal Exponents. This feature extraction algorithm is based on the same principle of mapping the 2D image into a 3D surface. The fractal dimension is computed using a variable size structured element (SE) [9] in order to eliminate details whose scale is less than the size of SE.

Similar to the definition of the multifractal measure used in [8] a local natural measure $\mu_\varepsilon(i,j)$ is defined in a window of size *WxW*:

$$\mu_\varepsilon(i,j) = \frac{|f_\varepsilon(i,j) - f(i,j)|}{\sum_{i,j}^{W}|f_\varepsilon(i,j) - f(i,j)|}$$

Equation 3 The μ_ε measure. $f(i,j)$ is the gray level of the *(i,j)* pixel and $f_\varepsilon(i,j)$ is the gray level of the pixel from dilated image

The measure of order q at scale ε can be computed using Equation 4:

$$I(q,\varepsilon) = \alpha\sum_{i,j}^{W}\mu_\varepsilon(i,j)^q$$

Equation 4 The measure of order *q*

Subsequently, a set of multifractal texture descriptors, namely the local morphological multifractal exponents (LMME) is shown in Equation 5:

$$L_q = \frac{1}{q}\lim_{\varepsilon\to 0}\frac{\ln(I(q,\varepsilon))}{\ln(\frac{1}{\varepsilon})}, q \neq 0$$

Equation 5 Local morphological multifractal exponents

Multi resolution fractal dimension.[10]. Multi resolution texture analysis allows us to separately examine texture elements at different scales.

The pyramidal approach is one of the most popular ways for multi resolution image analysis. Based upon the concept of pyramidal data analysis, we can define the image at resolution level *i* by averaging the corresponding pixel values from image at resolution *i+1*.

The fractal dimension can be estimated by using least square linear regression to determine the slope of of the curve *id(k)* versus *k* in log-log scale. The *id(k)* is defined in Equation 6.

$$id(k) = \frac{\sum_{x=0}^{M-1}\sum_{y=0}^{M-k-1}|I(x,y)-I(x,y+k)| + \sum_{x=0}^{M-k-1}\sum_{y=0}^{M-1}|I(x,y)-I(x+k,y)|}{2M(M-k-1)}$$

Equation 6 The id(k) fractal coefficient

In order to obtain a multi resolution analysis one has to compute the fractal dimension at several scales.

2.3. Laws' texture energy measures

Laws' texture energy measures [10], [11] are derived from three simple vectors of length 3, L3=(1,2,1) E3=(-1,0,1) S3=(-1,2,-1). If these vectors

are convolved with themselves or with each other, we obtain five vectors of length 5.

If we multiply the column vectors of length 5 by row vectors of the same length, we obtain Laws' 5 x 5 masks. To use these masks to describe texture in an image, we convolve them with the image and use statistics (e.g., energy) of the results as texture properties.

2.4. Wavelet transform

An image can be considered a function that varies in time. Various parts of the frequency spectra can be analyzed. Wavelet transform retains both time and spatial data. Wavelet transform divides the image in smaller sub-regions. This paper uses discrete wavelet transform. [12]

The features computed are root mean square variation, entropy and first moment of power spectrum.

Here the wavelet transform is applied to scale down the image four times.

2.5. Classification Algorithm

Support Vector Machines (SVM) or kernel machines are a family of learning methods that can represent complex, nonlinear functions [13] [14].

The basic idea for this algorithm is to map a lower dimension feature space where the classes are not linearly separable into a higher dimensional feature space where these classes are linearly separable.

Cross Validation is employed to assess the classification accuracy. From 10 fold cross validation and 5 runs the kappa statistic is derived [15]. This coefficient measures the agreement between predicted and observed categorizations of a dataset, while correcting for agreement that occurs by chance (Equation 7). The maximum value for kappa is 1 corresponding to a perfect classifier. The value 0 corresponds to a random classifier.

$$\kappa = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)},$$

Equation 7 Computing the kappa statistic

where Pr(a) is the relative observed agreement among raters, and Pr(e) is the probability that agreement is due to chance.

In order to compare two kappa values the Student t test is used with $p<0.05$.

2.6 Feature selection algorithms [14]:

2.6.1. Gain Ratio is derived by taking into account the number and size of subsets into which an attribute splits the dataset, disregarding any information about the class.

2.6.2. Chi-squared statistic is used to assess the correlation between each feature and the corresponding class. The features are sorted descending and only the relevant features are kept ($p<0.05$). In previous studies [16] it has been emphasized that the mean value for some features varies significantly from one fibrosis class to another.

2.6.3. CFS evaluation Evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them. It iteratively adds attributes with the highest correlation with the class as long as there is not already an attribute in the subset that has a higher correlation with the attribute in question

2.6.4 Symmetrical uncertainty evaluates the correlation between two attributes using the joint entropy. The smallest subset of attributes that correlates with the class is selected.

3. Experimental Results

3.1. Patients and US images

Over 500 patients were investigated using US examination and liver biopsy. For each patient a number of US images were acquired from left lobe at 8 cm depth. The US machine parameters were set in such a way that the quantity of the information received from the tissue was maximized. The post processing settings were set to off. The fibrosis grade was assessed using METAVIR gradation [16]. Patients were divided into four groups. Fibrosis 1 through 3 and cirrhotic patients. The last group wasn't biopsied because cirrhosis was assessed by other means.

The fifth group consists of clinically healthy patients. This group wasn't biopsied as well.

On the acquired US images a region of interest of 64x64 pixels was manually set. The ROI's have to be as close as possible to the vertical middle line of the image. Also, the ROI's have to be set in the focal region of the US machine. The focal region is manually set by the physician during acquisition in order to focus the US beams into the liver tissue. The ROI's have to be free of artifacts.

A special designed framework was used to establish the regions of interest and compute texture attributes.

All 8 presented algorithms were implemented in the framework, tested and validated. Using different parameter sets for each algorithm 1100 features were computed from each ROI.

The number of patients and US images for each group is shown in Table 1:

Table 1 The number of patients and US images used in present study

Group	No of patients	No of US images
Healthy	22	55
Fibrosis 1	70	267
Fibrosis 2	82	303
Fibrosis 3	35	116
Cirrhosis	29	77

3.2. Results

3.2.1. Differentiating between healthy and fibrosis patients implies comparisons between the healthy lot and the patients with various fibrosis grades. In Table 2 are shown the results.

Table 2. Comparison between healthy patients and various fibrosis grades

Comparison case	Kappa
Healthy – Fibrosis 1	0.14
Healthy – Fibrosis 2	0.1
Healthy – Fibrosis 3	0.25
Healthy – Cirrhosis	0.24

One can notice that higher fibrosis grades can be distinguished better than the lower ones.

In Table 3 is shown the influence of steatosis grade over normal – fibrosis discrimination.

Table 3. The influence of steatosis(fatty) grade over the normal – fibrosis grade comparison regardless of activity grade

Comparison case	Kappa
Healthy – Fibrosis 1 Fatty 0	0.13
Healthy – Fibrosis 1 Fatty 1	0.41
Healthy – Fibrosis 2 Fatty 0	0.1
Healthy – Fibrosis 2 Fatty 1	0.57
Healthy – Fibrosis 2 Fatty 2	0.73
Healthy – Fibrosis 3 Fatty 0	0.37
Healthy – Fibrosis 3 Fatty 1	0.18

A pattern is observed, when the fibrosis is associated with the steatosis the detection rates increase significantly. There is an exception at fibrosis grade 3 where the presence of steatosis lowers the detection rates. This exception is most likely caused by an experiment error, although further investigation is necessary in order to confirm this rule. As one can notice not all of the fibrosis/fatty combinations are present in the table because of the reduced sample volume.

From these results one might conclude that the presence of the steatosis increases the detection rates, although is more likely that the classifier detects the associated steatosis not the fibrosis grade. Further studies must be developed in order to investigate these results.

In Table 4 is shown the influence of the activity over the same normal – fibrosis grade comparison.

Table 4 The influence of activity over normal – fibrosis comparison regardless of steatosis grade

Comparison case	Kappa
Healthy – Fibrosis 1 Activity 0	0.27
Healthy – Fibrosis 1 Activity 1	0.3
Healthy – Fibrosis 1 Activity 2	0.3
Healthy – Fibrosis 2 Activity 1	0.29
Healthy – Fibrosis 2 Activity 2	0.12
Healthy – Fibrosis 2 Activity 3	0.24
Healthy – Fibrosis 3 Activity 1	0.28
Healthy – Fibrosis 3 Activity 2	0.41

An interesting observation is that when we train the classifier to distinguish between healthy/fibrosis 1 it performs significantly worse than in the situation presented above (where we separate the lot according to activity grade). Similar behavior is noticed at fibrosis 2 and 3 with few exceptions regarding fibrosis 1 Activity 2 and fibrosis 3 Activity 1 where the discrimination rates remain the same. (The differences don't have statistical significance when comparing to corresponding Healthy – Fibrosis comparison from Table 2)

The activity grade has a major impact on detecting the fibrosis grade. The detection rates increase only when we consider each activity grade separately but decreases when we combine all activity grades into one lot. In steatosis comparison case (Table 3) one can notice that the discrimination rates increase with the considered steatosis grade. The activity doesn't behave in this manner so we can conclude that indeed the fibrosis is detected in this comparison case (Table 4) but training classifiers separately for each activity grade increases the detection accuracy.

3.2.2 Differentiating between lower fibrosis grades and cirrhotic patients is important. In Table 5 is shown the detection rates between cirrhotic patients and various fibrosis grades.

Table 5 Fibrosis grades – cirrhotic lot comparison. The effect of activity and fatty.

Comparison case	Kappa
Cirrhosis – Fibrosis 1	0.19
Cirrhosis – Fibrosis 2	0.13
Cirrhosis – Fibrosis 3	0.15
Cirrhosis – Fibrosis 1 Activity 0	0.18
Cirrhosis – Fibrosis 1 Activity 1	0.32
Cirrhosis – Fibrosis 1 Activity 2	0.32
Cirrhosis – Fibrosis 2 Activity 1	0.01
Cirrhosis – Fibrosis 2 Activity 2	0.21
Cirrhosis – Fibrosis 2 Activity 3	-0.07
Cirrhosis – Fibrosis 3 Activity 2	0.19
Cirrhosis – Fibrosis 3 Activity 3	0.4
Cirrhosis – Fibrosis 1 Fatty 0	0.14
Cirrhosis – Fibrosis 1 Fatty 1	0.36
Cirrhosis – Fibrosis 2 Fatty 0	0.15
Cirrhosis – Fibrosis 2 Fatty 1	0.27
Cirrhosis – Fibrosis 2 Fatty 2	0.4
Cirrhosis – Fibrosis 3 Fatty 0	0.08
Cirrhosis – Fibrosis 3 Fatty 1	0.15

Cirrhosis liver is differentiated better from healthy or lower fibrosis grade than from increased fibrosis grade.

Steatosis has the same effect as in normal – fibrosis comparison cases. Lots that contain greater steatosis grades are detected better.

Activity has a different behavior. In general a higher activity score increases the discrimination accuracy. Exception is in fibrosis 2 case where the results are worse. Unfortunately the fibrosis 2 activity 0 and fibrosis 3 activity 0 cases were not present in the comparisons.

When we differentiate fibrosis patients from healthy patients the activity grade doesn't have a major impact. (The improvements appear when we divide the lot according to activity grade). The situation changes when discriminating cirrhotic patients where the presence of activity (grades greater than 0) increases the discrimination rates. Careful examination is needed here in order to rule out the possibility that activity is detected instead of fibrosis.

Comparing the fibrosis grade 1 with fibrosis grade 2 we obtain the results in Table 6

Table 6 Comparison between fibrosis grade 1 and fibrosis grade 2. The effect of activity and steatosis

Comparison case	Kappa
Fibrosis 1 – Fibrosis 2	0.11
Fibrosis 1 Fatty 0 – Fibrosis 2 Fatty 0	0.05
Fibrosis 1 Fatty 1 – Fibrosis 2 Fatty 0	0.32
Fibrosis 1 Fatty 1 – Fibrosis 2 Fatty 1	0.04
Fibrosis 1 Activity 1 – Fibrosis 2 Activity 1	0.09
Fibrosis 1 Activity 2 – Fibrosis 2 Activity 2	0.03

Fibrosis grades 1 and 2 have a medical importance because they determine the treatment approach.

When the same steatosis grade is present (0 or 1) the discrimination between these grades is very poor. The same results are found when we keep constant the activity grade. In case of different fatty grade the differentiation improves significantly. Again, this indicates that we detect different steatosis grades not different fibrosis grades. Unfortunately the reduced number of patients from these lots keeps us from deducing any rule about steatosis/activity influence over fibrosis 1 – fibrosis 2 comparison case.

3.2.3. Detecting steatosis grade and determining how fibrosis affects the detection rates is an important medical issue. In Table 7 are shown the comparisons between steatosis grades.

Table 7. Detection rates of steatosis grades.

Comparison case	Kappa
Fatty 0 – Fatty 1	0.19
Fibrosis 1 Fatty 0 – Fibrosis 1 Fatty 1	0.26
Fibrosis 2 Fatty 0 – Fibrosis 2 Fatty 1	0.28
Fibrosis 3 Fatty 0 – Fibrosis 3 Fatty 1	0.23
Fatty 0 – Fatty 2	0.45
Fibrosis 2 Fatty 0 – Fibrosis 2 Fatty 2	0.59
Fatty 1 – Fatty 2	0.32

The discrimination between steatosis grades is better as the steatosis grade difference increases (better results on 0 – 2 than between 0 – 1 grades)

The steatosis grade 0 from grade 1 is discriminated in the same proportion regardless of the associated fibrosis grade. Unfortunately there weren't enough patients to make this assessment for steatosis grade 0 – steatosis grade 2 comparison.

3.2.4. Feature selection algorithms. Each comparison case was used to train and test 5 SMO algorithms with the same parameters. The only difference is that we used the four presented feature selection algorithms. One algorithm used all 1100 features. Student *t* test

was employed in order to determine which is the algorithm fitted to classify the comparison case. In the end a ranking for these algorithms was constructed. The results show that the SMO algorithm trained with full feature set gives us the better results in most of the cases.

This result is somehow biased by the fact that there are comparisons that study the fibrosis grade, or the steatosis grades etc. Each comparison case might give us different results in terms of algorithm ranking.

Because this study was designed to give us an overview over the fibrosis grades detection, no separate algorithm ranking was employed. In the future we will address specific comparison cases (i.e. differentiating fibrosis grade 1 from fibrosis grade 2) and therefore the feature selection algorithms will be evaluated again.

4. Discussions and conclusions

Liver biopsy identifies three components: fibrosis, inflammatory/necrosis (quantified as activity) and steatosis (fatty). These three components affect the ultrasound interference pattern generated by the tissue. Present study reveals that using texture features one cannot clearly separate the quantity of fibrosis/steatosis/necrosis present into the liver tissue.

It appears that fibrosis and steatosis produce the largest texture alterations.

The association of fibrosis with steatosis and activity is not equally distributed among the patients. For example our patient lots do not contain a single patient with fibrosis 3 associated with activity grade 0 or 1. Fibrosis 1 combined with activity 3 is present at only one patient. These distributions might bias the comparisons where we investigate only one parameter (i.e. fibrosis regardless of steatosis or activity grade) because in fact the patient lots don't contain all the possible combinations of the steatosis/activity in the same proportion. As a result is possible that, for example, when we try to discriminate between fibrosis 1 and fibrosis 2 grade in fact we will detect the steatosis grade 1 from steatosis grade 0 because steatosis grade 1 is predominant in fibrosis 1 lot and steatosis grade 0 is predominant in fibrosis grade 2 lot.

Epidemiological studies must be consulted in order to find the proportions in which various combinations of fibrosis/steatosis/activity are found.

In all comparison cases we can note a significantly increased detection rate when different steatosis grades are present. Because of the reduced number of patients involved in this study we couldn't perform all the comparisons where only one parameter varies and the rest remain constant. From the few ones that were available we concluded that the steatosis is better detected than the fibrosis.

One way to approach the fibrosis identification is to build imagistic models for each fibrosis grade. This requires that we have a large number of patients who have pure fibrosis (i.e. activity and fatty will have grade 0) for each fibrosis grade. Along with the fact that pure fibrosis patients are scarce there is still the problem with healthy and cirrhotic groups. These patients cannot be biopsied so the activity/fatty grade 0 cannot be safely assumed. Also, the steatosis grade in cirrhotic patients cannot be assessed.

Once such models are built, we hope that it might be possible to train classifiers that will correctly detect not only pure fibrosis but also fibrosis in the presence of activity or steatosis. Feature computation algorithms could be developed in such a way that they will not be affected by the presence of steatosis.

Another approach is to develop classifiers that will distinguish between most frequent fibrosis – activity – steatosis associations. The behavior of these classifiers in the presence of other kind of associations is hard to be estimated at this point. Although this might be an easier way to detect fibrosis, the cost of misclassification could be greater than in the previous approach.

Feature selection algorithms must be evaluated with respect to certain comparison case. After this step is accomplished, one can address the problem of which features are relevant for detecting fibrosis grade, for detecting steatosis, etc.

A meta-classifier might be the solution in fibrosis grade detection, where individual classifiers are trained for specific tasks and one "meta" classifier is used to select which classifier will give the answer for current input.

Building imagistic models for each fibrosis grade and for each steatosis grade might be the best way to develop this meta-classifier.

More patients have to be included in the study in order to be able to assess all the grade combinations and get a better view on how the fibrosis and steatosis interact.

5. Acknowledgements

This work was sponsorised by the Romanian National Authority for Scientific Research as part of the CEEX program FINALISM 94/31.07.2006 and SIDEF 71/17.07.2006

6. References

1 S. Saadeh, G. Cammell, W.D. Carey, Z. Younossi, D. Barnes, K. Easley. "The role of liver biopsy in chronic hepatitis C." in *Hepatology*, vol. 33(1), pp. 196-200, 2000.

2 T. Poynard, V. Ratziu, P. Bedossa. "Appropriateness of liver biopsy." *Can J Gastroenterol*, vol. 14(6), pp. 543-548, 2000.

3 M. Amadasun and R. King, "Textural Features Corresponding to Textural Properties", *IEEE Transactions on System, Man Cybernetics*, vol. 19, no. 5, pp. 1264-1274, 1989.

4 A. Materka, M. Strzelecki "Texture analysis methods - A review.", in *COST B11 Report,*Brusseles, June, 1998.

5 A. Ukovich, G. Impoco, G. Ramponi "A Tool Based on The Co-Occurrence Matrix to Measure the Performance of Dynamic Range Reduction Algorithms", *IEEE IST 2005 International Workshop on Imaging Systems and Techniques Niagara Falls*, 13-14 May 2005

6 C. Gui-Tao, S. Peng-fei, H. Bing, "Liver fibrosis identification based on ultrasound images captured under varied imaging protocols" *Journal of Zhejiang University SCIENCE* vol. 6B(11), pp. 1107-1114, 2005.

7 N. Sarkar, B.B. Chaudhuri, "Multifractal and generalized dimensions of gray-tone digital images," *Signal Process.*, vol. 42, pp. 181–190, Mar. 1995.

8 G. Du, T.S. Yeo, "A novel multifractal estimation method and its application to remote image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 4, pp. 980–982, Jul. 2002.

9 Y. Xia, D. Feng, R. Zhao "Morphology-Based Multifractal Estimation for Texture Segmentation" *Ieee Transactions On Image Processing*, vol. 15, no. 3, march 2006.

10 W.C. Yeh, H. Sheng-Wen, "Liver fibrosis grade classification with b-mode ultrasound", *Ultrasound in Med. & Biol.*, vol. 29, no. 9, pp. 1229–1235, 2003.

11 K. I. Laws, "Texture energy measures," in *Proc. Image Understanding Workshop***,** pp. 47-51, 1979

12 H. Yoshida, D. Casalino, "Wavelet-packet-based texture analysis for differentiation between benign and malignant liver tumours in ultrasound images", *Phys. Med. Biol*., vol. 48, pp. 3735-3753, 2003.

13 S. Russell, P. Norvig "Artificial Intelligence: A Modern Approach," *Prentice Hall Series in Artificial Intelligence*, Englewood Cliffs, New Jersey, 1995.

14 H. Ian., W. Frank "Data Mining: Practical machine learning tools and techniques", 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

15 J. Cohen, "A coefficient of agreement for nominal scales." *Educational and Psychological Measurement,* vol. 20, pp 37–46.

16 M. Lupşor, S. Nedevschi, C. Vicas, R. Badea "Estimating the fibrosis stage in the human liver tissue using image processing methods on ultrasonographic images. Preliminary results." *Proceedings of the Euro-Mediterranean Medical Informatics and Telemedicine. 3rd International Conference.* pp. 311-318, 2007.