The Role of the Textural Microstructure Co-occurrence Matrices in the Automatic Detection of the Cirrhosis Severity Grades from Ultrasound Images

Delia Mitrea*, Sergiu Nedevschi*, Mihail Abrudean*, Monica Lupsor-Platon, Radu Badea****

**Technical University of Cluj-Napoca, Faculty of Automation and Computer Science, Baritiu Str., No. 26-28, 400027, Cluj-Napoca, Romania (Tel: 0040-740-084147; e-mail: delia.mitrea@cs.utcluj.ro).*
***I. Hatieganu University of Medicine and Pharmacy, Department of Medical Imaging, V. Babes Str., No. 8, Cluj-Napoca, Romania (e-mail:monica.lupsor@umfcluj.ro)*

Abstract: Cirrhosis is a lethal disease that can also precede liver cancer. The golden standard for diagnosing this affection, the biopsy, is invasive and dangerous. The noninvasive detection of the cirrhosis severity grades is a major challenge in these conditions. The objective of our research is to automatically discover the cirrhosis evolution stages from ultrasound images, through unsupervised classification techniques, based on textural features. In this work, the role of the textural microstructure co-occurrence matrices in the detection of the cirrhosis severity grades was studied. The textural features were provided at the input of unsupervised classification methods. Through specific techniques, the relevant textural features for class separation and their specific values for each severity grade were determined. The method was validated through cluster visualizations and supervised classification, the resulted accuracy being above 90%. The effect of the Principal Component Analysis (PCA) technique upon the discovery process of the cirrhosis grades was also studied.

Keywords: ultrasound image analysis, textural microstructure co-occurrence matrices (TMCM), unsupervised classification, cirrhosis severity grades, computer-aided diagnosis.

1. INTRODUCTION

Cirrhosis represents a diffuse liver disease, which marks the beginning of hepatic restructuring. The main characteristics of this affection are the association of fibrosis, regeneration nodules and necrosis, leading to important changes in the structure of the hepatic parenchyma. Cirrhosis constitutes the end stage of multiple liver diseases and sometimes it leads to death by itself. The regeneration nodules, specific to cirrhosis, can also transform into dysplastic nodules, evolving into hepatocellular carcinoma (HCC), the most frequent form of liver cancer.

The B-mode ultrasonography is a non-invasive examination method, being safe, inexpensive, repeatable and easy to apply. Alternative methods for patient examination, such as the Computer Tomography (CT), Magnetic Resonance Imaging (MRI), the elastographic techniques (e.g. Transient Elastography, Acoustic Radiation Force Imaging, etc) or even Contrast Enhanced Ultrasonography (CEUS) are usually irradiating and/or expensive (Carr, 2010). We aim to build safe computer assisted diagnosis systems, which can be widely implemented, considering also those medical units that are not endowed with sophisticated, last generation equipment. This method is supposed to be applied repeatedly, in order to monitor the disease evolution with the best accuracy. In our research, the information derived from B-mode ultrasound images, processed through advanced texture analysis methods was considered, in order to assess its ability for characterizing the cirrhosis severity grades. This represents an important step towards a computerized, non-invasive, virtual biopsy system.

Within ultrasound images, cirrhosis is featured mainly by the tissue homogeneity decrease, due to the fact that the nodules continuously appear and evolve. These nodules could be hypo-echogenic, or even unapparent. Other changes that could occur are: increased volume, in the case of toxic cirrhosis; decreased volume, in the case of viral cirrhosis; shape and contour modification; vessel structure modifications. An eloquent example of a cirrhotic liver, together with the rectangular region of interest (ROI) selected on it is illustrated in the next figure (Fig. 1).

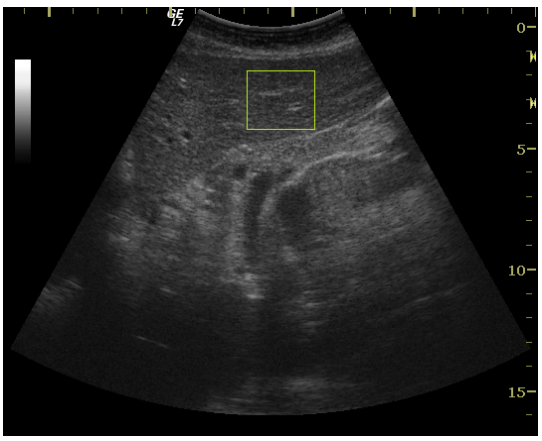


Fig. 1. ROI selected inside the liver affected by cirrhosis.

As cirrhosis is a very serious disease, monitoring its severity presents a great importance. We take into account that type of cirrhosis which has a C virus origin, so, in this case, in order to establish the severity grade of the associated fibrosis, the Metavir grading system is usually considered appropriate. Though, conversion formulas to other scoring systems exist and can be applied, as well. The Metavir classification system assimilates cirrhosis with the last evolution phase of liver fibrosis (Bedossa, 1996). Concerning cirrhosis, some existing medical assessment systems, reveal the next evolution stages for this affection: Stage A – incipient, "compensated"; Stage B – intermediate phase, beginning to decompensate; Stage C – final stage, "decompensated" (Peters, 2013). However, no significant study exists for establishing an objective quantification of the cirrhosis severity grades, using refined information, derived from ultrasound images through advanced computerized methods, so this will be the aim of the current work.

The most relevant existing studies, regarding the detection and characterization of some important diseases and of their severity grades, from medical images, are detailed below. An important approach concerning the recognition of the diffuse liver diseases from ultrasound images through supervised classification methods was described in (Cavouras, 1997). Thus, in order to differentiate among the diffuse liver diseases, the authors employed a hierarchical tree, structured in the following manner: at the first level, the differentiation between the normal and the unhealthy liver was performed, using textural features such as the GLCM mean and energy; at the second level, the difference between steatosis and cirrhosis was highlighted, using features such as the GLCM mean and variance; at the last level, the authors differentiated between multiple severity grades of steatosis and cirrhosis, using GLCM features such as the variance, entropy, sum entropy, difference entropy. For supervised classification, a Multilayer Perceptron (MLP) classifier was employed (Cavouras, 1997). The differentiation between healthy and cirrhotic liver regions, based on MRI images, through unsupervised classifiers, was performed in (Lee, 2008). The implemented methods were a k-means clustering classifier and textural features based on finite differences of the intensity function. The performance, measured through the area under ROC, was 0.704. In (Mala, 2010) the authors aimed to differentiate between cirrhotic and fatty livers, using the information provided by computerized tomography (CT) images, by employing Wavelet-based statistical textural features and several neural-network based classifiers. The maximum resulted accuracy was 96%. The most recent approach involving the differentiation between the cirrhotic and the normal liver, from M-mode ultrasound images, in supervised manner, was presented in (Fujino, 2014). For this purpose, the authors adopted higher-order local autocorrelation (HLAC) features, feature selection methods and a nearest neighbour classifier. The average error rate resulted in these conditions was 12.11%. Concerning the activity of grading the severity of other affections, the most important approaches involved detecting the evolution stages of the glioma tumours, by combining hierarchical classification methods and dimensionality reduction techniques (Yang, 2014), respectively distinguishing HCC in incipient phase through a combination between supervised and unsupervised classification methods (Ciocchetta, 2000).

As it results from the methods described above, no significant research exists concerning the detection of the cirrhosis grades from ultrasound images, through unsupervised classification methods. Thus, in our work, the objective was to discover the cirrhosis evolution stages, by using textural features and clustering methods. In (D. Mitrea, 2013), a preliminary study was performed in order to detect the cirrhosis grades form ultrasound images, in unsupervised manner.

In this paper, the focus was put on the role that the Haralick features derived from the newly defined TMCM matrix of order two and three had in the above-mentioned process. For performing unsupervised classification, clustering methods, such as Expectation Maximization (EM) (Witten, 2011) and X-means (Pelleg, 2000) were chosen, and also an unsupervised meta-classifier was considered, in order to obtain dense clusters (Zaki, 2014). *A textural model* of the cirrhosis evolution stages was defined, consisting of the relevant textural features for class separation and of their specific values for each class: mean, standard deviation, probability distribution. Thus, specific methods for the selection of the relevant textural features, in the context of the unsupervised classification process, were developed and appropriate techniques for determining the specific values of the relevant textural features were employed. The effect of the Principal Component Analysis (PCA) dimensionality reduction technique upon the performances of the unsupervised classification process was also analysed. The correlation of the most important textural features with the physical and visual properties of the cirrhotic parenchyma was discussed, as well. The newly developed methods were assessed through performance parameters specific for the unsupervised classification, such as the log-likelihood, for the EM method (Witten, 2011), respectively the distortion, for X-means clustering (Pelleg, 2000). The results were also validated by performing data visualization through Self Organizing Maps (SOM) (Yin, 2008) and by employing supervised classification methods, well known for their performance, such as the Multilayer Perceptron (MLP), the Support Vector Machines (SVM), decision trees, AdaBoost and multiclass meta-classifiers in combination with the basic learners (Witten, 2011).

This paper was structured in the following manner: after the introduction, which described the state of the art and the contribution of the current work, the proposed solution was presented, by describing in details the developed and applied methods; then, the performed experiments were detailed and the obtained results were discussed; at the end, the conclusions were stated and the possibilities of further research were proposed.

2. THE PROPOSED METHODS

In our work, the aim was to discover the existing cirrhosis grades, corresponding to the classes (clusters) that existed in the data derived from ultrasound images, as well as the characteristics of each grade, in terms of the relevant textural

features, able to provide the best class separation. The specific values of the textural features for each resulted class were determined, for the purpose of computer-assisted diagnosis. Thus, *the textural model of the cirrhosis severity grades was defined*, consisting of the relevant features for class separation, respectively of the specific values associated to the relevant textural features: arithmetic mean, standard deviation, probability distribution. The method was validated by performing cluster visualizations using advanced specific techniques and through supervised classification. The steps of the corresponding methodology were the following:

1. Compute the textural features, by applying texture analysis methods (*the image analysis phase).*
2. Discover the existing classes and their properties, performing the sub-steps described below *(the learning phase*):
   2.1. Apply clustering methods in order to discover the classes (clusters) that exist in the data;
   2.2. Determine the set of the relevant textural features, then re-evaluate the existing clusters by applying the clustering methods again, considering the derived feature set;
   2.3. Compute the specific values of the relevant textural features.
3. Validate the method by performing cluster visualizations using Self Organizing Maps (SOM) (Yin, 2008); apply supervised classification methods, providing at the classifiers' inputs the values of the relevant textural features and assess the classification performance in each case *(the validation phase*).

The methods specific for each step were described in the next sub-sections.

2.1. The computation of the textural features (the image analysis phase)

Already existing methods for texture analysis (Toennies, 2012) were implemented, and new methods were defined, as well. First order statistics of the grey levels, and also the Grey Level Co-occurrence Matrix (GLCM) of order two, together with the associated Haralick features (Davis, 1981) were employed. Edge based statistics, the autocorrelation index, and also first order statistics based on textural microstructures (arithmetic mean and frequency), detected by applying the Laws' convolution filters (Laws, 1980) were adopted as well. Also, the Textural Microstructure Co-occurrence Matrix (TMCM) was defined, as described in (1), the intention being that of providing a more refined texture characterization:

$$\begin{aligned} &C_D(c_1,c_2,..,c_n) = \#\{((x_1,y_1),(x_2,y_2),(x_3,y_3),..,(x_n,y_n)):\\ &A(x_1,y_1)=c_1, A(x_2,y_2)=c_2,..,A(x_n,y_n)=c_n,\\ &|x_2-x_1| = |\vec{dx}_1|, x_3-x_1| = |\vec{dx}_2|,..,|x_n-x_1| = |\vec{dx}_{n-1}|,\\ &|y_2-y_1| = |\vec{dy}_1|,|y_3-y_1| = |\vec{dy}_2|,..,|y_n-y_1| = |\vec{dy}_{n-1}|,\\ &\mathrm{sgn}((x_2-x_1)(y_2-y_1)) = \mathrm{sgn}(\vec{dx}_1\cdot\vec{dy}_1),..\\ &,\mathrm{sgn}((x_n-x_1)(y_n-y_1)) = \mathrm{sgn}(\vec{dx}_{n-1}\cdot\vec{dy}_{n-1})\} \end{aligned} \quad (1)$$

In (1), #*S* is the size of the set *S* (the number of elements) and

$$d = \{(\vec{dx}_1,\vec{dy}_1),(\vec{dx}_2,\vec{dy}_2),...,(\vec{dx}_{n-1},\vec{dy}_{n-1})\} \quad (2)$$

is the set of the displacement vectors.

The TMCM matrix was determined after applying a certain Laws' convolution filter (S_5S_5, R_5R_5, S_5R_5, or R_5S_5) (Laws, 1980), for detecting textural microstructures, such as spots and ripples or their combinations, as the frequencies of these elements resulted as relevant textural features in our previous work (D. Mitrea, 2013; D. Mitrea, 2012). The definition of this matrix was firstly provided in (D. Mitrea, 2014). However, in the current work, a k-means clustering method (Witten, 2011) was also employed, after the convolution process. Each element of the newly defined TMCM matrix represents the number of n-tuples of pixels having the spatial coordinates (x_i, y_i), $i \in \{1,.., n\}$, respectively the values of the cluster labels c_i, $i \in \{1,..,n\}$, assigned by k-means clustering. *A* represents the attribute associated to each pixel - in our case, the appropriate cluster label. The spatial relation between the pixels is defined by the vectors depicted in (2).

The TMCM matrices of order two and three were determined. In the first case, the following directions were considered for the vectors of displacement: 0°; 90°; 180°; 270°. The absolute values of these vectors were 0 or 2, the resulting Haralick features being computed as the arithmetic means of the parameters corresponding to the individual matrices. For TMCM of order three, the following combinations of directions were taken into account, in the case of the co-linearity of the pixels: (0°, 180°); (90°, 270°); (45°, 225°); (135°, 315°). Also, the following direction combinations were considered: (0°, 90°); (90°, 180°); (180°, 270°); (0°, 270°); (45°, 135°); (135°, 225°); (225°, 315°); (45°, 315°), when the three considered pixels were constituted in a right angle triangle, having the current pixel in the right angle vertex position. The absolute values for the displacement vectors were 0 or 2, as in the previous case. At the end, the extended Haralick parameters, defined as in (D. Mitrea, 2012) were averaged.

2.2 Discovering the cirrhosis grades and their specific properties (the learning phase)

2.2.1. Clustering methods

The *Expectation Maximization (EM)* method (Witten, 2011) was adopted, as it is a well-known, powerful and flexible technique, which iteratively estimates the desired parameters, by maximizing the log-likelihood of the model.

The likelihood was computed using the formula (3), the log-likelihood being the natural logarithm of the likelihood (Witten, 2011):

$$Likelihood = \prod_{i=1}^{n}\sum_{j=1}^{m} p_j P_r[x_i \mid j] \quad (3)$$

In (3), *n* is the number of instances, *m* is the number of clusters, p_j are the cluster priors and $P_r[x_i|j]$ represents the conditional probability that the instance $\boldsymbol{x}_i$ appears within cluster *j*. The parameters to be estimated in our context were

the number of clusters and the sample distributions within the clusters.

The *X-means clustering method* was also employed, as it is an improved version of the k-means clustering technique. The classical k-means clustering method (Witten, 2005) has some drawbacks, the most important being the execution speed, as well as the fact that a fixed value for *k* must be provided a-priori. The method of X-means clustering expects a maximum and a minimum value for the *k* parameter and it performs the following steps:

(1.) Run conventional k-means to convergence, for a fixed value of *k*.

(2.) Decide whether new cluster centroids should appear or not, by splitting the old centroids into two.

(3.) If $k>k_{max}$, then stop and report the best model identified during the algorithm, according to the Bayesian Information Criterion – BIC (Pelleg, 2000).

The BIC criterion is used both in order to decide which centroids have to be split, and also to identify the best resulted model.

The overall algorithm performance is estimated by the distortion measure, expressing the average squared distance from the points to their centroids, for the best model (Pelleg, 2000).

Specific methods for obtaining dense clusters were additionally used, for result confirmation. These methods are based on the principle that the clusters are dense point regions separated by sparse point regions and aim to discover classes for which the distances between the points and the corresponding cluster centroid are as small as possible. Also, the density-based clusters don't make any prior assumption concerning the point distributions, trying to provide the best fitting probability density.

An unsupervised meta-classifier, in combination with the above-mentioned elementary methods, was used in the current work, in order to wrap each clustering method and to make it return dense clusters together with the corresponding probability distributions. The performance of this method was measured by the log-likelihood, associated to the best-resulted model (Zaki, 2014).

2.2.2 Dimensionality reduction techniques applied in the context of unsupervised classification

2.2.2.1. Method for relevant feature selection

The relevant features for the unsupervised classification process were selected, in order to provide best class separation, meaning that the overlapping region between two clusters must be as small as possible. For each textural feature *f*, a relevance score was defined according to our formula described in (4):

$$\mathrm{Re}levance\,(f) = \sum_{i,j} (1 - Overlapping\_reg\_size_{i,j}) \tag{4}$$

In (4), *i* and *j* are neighboring classes (clusters). The relevance of each feature is dependent on the cumulated size of the overlapping regions that exist between each pair of neighboring clusters. The overlapping region size was determined as described in (5), assuming a Gaussian distribution for each feature:

$$Overlap\_reg\_size = \begin{cases} S,\ if\ S > 0\ and\ \mu_i \neq \mu_j\ (overlap.clust.) \\ 0,\ if\ S \leq 0\ and\ \mu_i \neq \mu_j\ -\ (non-overlap.clust.) \\ 2 \cdot \dfrac{\sigma\_min_f}{Max_f - Min_f},\ if\ \mu_i = \mu_j \end{cases} \tag{5}$$

where

$$S = \frac{(\mu\_\min_f + 2\sigma\_\min_f) - (\mu\_\max_f - 2\sigma\_\max_f)}{Max_f - Min_f} \tag{6}$$

In (6), $\mu_{\_min\_f}$ represents the minimum value of the arithmetic mean for the feature *f,* considering the two clusters, *i* and *j*; $\sigma_{\_min\_f}$ stands for standard deviation of the textural feature *f* considering the cluster where μ is minimum, $\mu_{\_max\_f}$ is the maximum arithmetic mean of *f, within the clusters i and j,* while $\sigma_{\mu\_max\_f}$ is the standard deviation of *f* inside the cluster where μ has maximum value. Formula (5) estimates the size of the overlapping region for the two clusters, *i* and *j*, with respect to the feature *f*. Thus, if $S \leq 0$, no overlapping region exists for the two considered clusters. If the centers of the clusters *i* and *j* are overlapped considering the feature *f* (the arithmetic means have the same values), then the size of this overlapping region will be twice the minimum standard deviation.

2.2.2.2. Principal Component Analysis (PCA)

The method of Principal Component Analysis (PCA) is a well-known dimensionality reduction technique, which maps the original dataset in a new space, where the main variation modes are emphasized. In mathematical terms, the classical PCA technique aims to determine a linear mapping *M*, in order to maximize the quantity $M^{T}cov(X)M$, where $cov(X)$ is the covariance matrix of the dataset *X*. The matrix *M* is formed by the first *d* eigenvectors that correspond to the highest eigenvalues of the covariance matrix, from which the arithmetic means are subtracted. In our work, the classical PCA algorithm was implemented and its generalized version, kernel PCA, with Gaussian and polynomial kernels, was considered as well (van der Maaten, 2008), their effects upon the performance of the unsupervised classification methods being analyzed.

2.2.3. Detecting the specific values of the relevant textural features

The arithmetic means, corresponding to the cluster centers and the standard deviations for each relevant textural feature were determined in the case of each cluster. Graphical representations of the normalized arithmetic means of the most important textural features, for all the existing clusters, were performed, in order to study the evolution of these values with the severity grade of cirrhosis. The specific intervals of variation for each feature were determined as well, using the probability distribution tables provided by the Bayesian Belief Networks technique (Witten, 2011). This

method detects causal influences among the features, by building a dependency network, represented as a directed, acyclic graph. In this graph, the nodes stand for the considered features, while the edges correspond to the causal influences between them, having associated the values of the conditional probabilities that correspond to certain dependences. Every node *X in* the graph has a set of parents *P*, respectively a set of children *C*. The probabilities of the nodes are determined through a complex inference mechanism. Thus, in a Bayesian Belief Network, each node has a probability distribution table, indicating the specific intervals of values for that node given the values of its parents. Considering the textural features that are relevant for each class, this method provides the most probable variation intervals for each of them, depending on the value of the class.

Establishing correlations between the newly defined TMCM features in the case of cirrhosis, was another objective of this work. The types of correlations considered for assessment were: linear, logarithmic, inverse, quadratic, cubic, power, compound, S, logistic, growth, exponential. The strength of these correlations were measured by the *R-Square* coefficient (Jalobeanu, 2001).

2.3. The validation phase

2.3.1. Validation through supervised classification

Each sample in the experimental set was annotated with the label of the class determined by using the clustering methods, applied after the selection of the most relevant textural features. Then, supervised classifiers were employed, in order to estimate the accuracy of the automatic diagnosis for the cirrhosis grades. Basic classification methods were implemented, which provided natural extensions to the case of multiple classes, as well as meta-classifiers, which reduced the multi-class case to a set of binary classifications, using specific strategies (Aly, 2005).

As basic learners, the methods of Multilayer Perceptron (MLP), Support Vector Machines (SVM), and the C4.5 decision trees algorithm (Witten, 2011), respectively the AdaBoost meta-classifier in combination with C4.5, which provided the best results in our former experiments (D. Mitrea, 2012; D. Mitrea, 2014), were taken into account. The accuracies (recognition rates) obtained when considering the entire group of textural features (old and newly defined), respectively only the formerly existing features, were compared. Other performance parameters, such as the sensitivity (TP rate), the specificity (TN rate) and the area under the Receiver Operating Characteristic (ROC) curve (Witten, 2011) were also estimated. The method of Correlation based Feature Selection (CFS) was also used in the context of supervised classification, in order to refine the set of the relevant textural features specific for each severity grade. This method determines the best subset of the relevant textural features, which have strongest correlations with the class parameter (Weka, 2015).

2.3.2. Self-Organizing Maps (SOM)

The method of Self Organizing Maps (SOM) (Yin, 2008) represents an unsupervised learning instrument, being a subclass of the techniques based on Artificial Neural Networks (ANN) (Witten, 2011). SOM is designed according to the model of an associative memory, performing topographic mappings of the similar items in the same spatial neighborhood. The usual Kohonen SOM is a feed-forward neural network with a single computational layer, looking like a bi-dimensional map, of rectangular or hexagonal form, built by neurons, each of them being connected to all the instances of the input data. Each neuron, *i,* has associated a weight vector, w_i, of the same dimension *(n)* as the input data. At the beginning of the algorithm, all the weights $\{w_1, w_2, ..., w_m\}$, where *m* is the total number of the neurons, are initialized randomly, to small numbers, these weights being updated during the SOM specific algorithm. This algorithm finds the neuron whose weight is closest to the input data, then updates the weight of this neuron and extends the updates in its neighborhood, as well. At the end, the Self Organizing Map illustrates the similar groups (clusters) that exist in the original data. In our work, the SOM method was used in order to put into evidence how well the classes and the subclasses that corresponded to various severity grades of cirrhosis were emphasized before and after the selection of the relevant textural features (Yin, 2008).

3. EXPERIMENTS AND DISCUSSIONS

The experimental dataset contained 75 patients affected by cirrhosis, biopsied for diagnostic certification. B-mode ultrasound images were acquired for each patient, using an ultrasound machine of type Logiq 7, under the same settings – frequency: 5.5 MHz, gain: 78, depth: 16 cm. A ROI of 50x50 pixels was selected on the cirrhotic tissue, for each ultrasound image. The texture analysis methods, described in section 2.1, were applied on each ROI; these methods were implemented using our software application developed in Visual C++. Most of the clustering methods, as well as the supervised classifiers, the method of Bayesian Belief Networks, and the CFS method for feature selection in the supervised classification context, were employed by using the Weka 3.6 library (Weka, 2015). The classical PCA, as well as the kernel PCA techniques, were employed in the Matlab environment, using the dimensionality reduction toolbox (Dimensionality Reduction, 2008). The correlations between the newly defined textural features were established using the IBM SPSS 20 environment (IBM SPSS Statistics, 2015). The method of SOM was implemented in Matlab, using the SOM toolbox (SOM, 2005).

3.1. Detecting the severity grades through unsupervised classification methods

The methods of Expectation Maximization (EM), respectively X-means clustering (XMeans), were applied for this purpose, individually, as well as in combination with the *Make Density based Clusterer* technique (Weka, 2015). For EM, the *numClusters* parameter was initially set to –1, in

order to automatically detect the number of clusters through cross-validation (Weka, 2015). In this situation, the number of clusters was found as being 4. Then, the number of clusters was also successively set to 2, 3 and 5, the log likelihood being estimated in each situation. When setting the *numClusters* parameter to $n>5$, the difference between the proportions of the resulted clusters was very increased, so the result was not considered as being relevant. The values of the log-likelihood, together with the maximum difference between the cluster proportions, for each value of the *numClusters* parameter, are depicted in Table 1.

Table 1. The classification performance obtained for the EM method.

Number of clusters	Log Likelihood	Maximum difference between cluster proportions
2	4.50	32
3	10.64	47
4	14.05	19
5	15.24	23

X-means clustering, with Cebyshev distance, which provided the best results, was also applied, after setting the *kd-trees* option to true, in order to increase the execution speed. The performance parameters, obtained for each value of the resulted number of clusters, are illustrated in Table 2.

Considering the results depicted within Table 1 and Table 2, the number of clusters obtained when setting the –1 value of the *numClusters* parameter for the EM method, respectively the fact that the difference between the cluster proportions should not be very large, the conclusion was that the final number of clusters was 4.

Table 2. The classification performance obtained for the X-means clustering method.

Number of clusters	Distortion	Maximum difference between cluster proportions
2	33.27	15
3	31.63	24
4	30.61	17
5	29.73	18

The *Make Density based Clusterer technique* (Weka, 2015), in combination with the above mentioned unsupervised classifiers provided the same results, validating the existence of four clusters within the data. This confirmed, as well, the result previously obtained in (D. Mitrea, 2013), and also the previously existing medical hypotheses expressed in the introduction, considering the fact that also cirrhotic liver parenchyma surrounding incipient HCC was taken into account in our dataset.

3.2. Dimensionality reduction techniques

3.2.1. Deriving the set of the relevant textural features

Within Table 3, the most relevant textural features were illustrated and the values for the relevance indexes, computed according to the algorithm described in the sub-section 2.2.2.1, under the hypothesis that four clusters existed in the data, were provided as well.

The relevant features were determined for each of the considered unsupervised classification methods and the final relevance score resulted as an average between the individual scores obtained. Only the features that had a relevance score above *0.65* were taken into consideration.

Table 3. The relevance score for the most important textural features.

Feature	Relevance score
Mean_spots	0.9196
Mean_edges	0.9169
Mean_level	0.9155
Mean_waves	0.9147
Mean_wavelet	0.9110
$TMCM_3$_Correlation_S_5R_5	0.9087
$TMCM_3$_Correlation_R_5S_5	0.9060
Edge_frequency	0.9030
$TMCM_3$_Entropy_S_5R_5	0.8989
Wave_frequency	0.8949
$TMCM_3$_Entropy_R_5S_5	0.8935
Spot_frequency	0.8932
TMCM_Entropy_S_5R_5	0.8904
Wavelet_frequency	0.8871
Level_frequency	0.8814
GLCM_Entropy	0.8718
TMCM_Entropy_R_5S_5	0.8641
GLCM_Correlation	0.8297
$TMCM_3$_Homogeneity_S_5R_5	0.8254
$TMCM_3$_Correlation_ripples	0.803
$TMCM_3$_Homogeneity_R_5S_5	0.787
$TMCM_3$_Correlation_spots	0.7379
TMCM_Entropy_spots	0.726
Edge_orientation_variability	0.722
TMCM_Homogeneity_S_5R_5	0.6895
$TMCM_3$_Variance_S_5R_5	0.6876
TMCM_Variance_spots	0.6854
$TMCM_3$_Variance_spots	0.6854
$TMCM_3$_Variance_R_5S_5	0.6532

According to Table 3, it results that the most relevant parameters were the arithmetic mean and the frequency of the textural microstructures detected by using the Laws' convolution filters, together with the Haralick features derived from the second and third order TMCM matrix. These features highlighted the importance of the first, second and third order statistics based on textural microstructures in characterizing the evolution of cirrhosis. The GLCM parameters entropy and correlation were important as well, illustrating the chaotic structure of the liver tissue in evolved phases of cirrhosis, respectively differences in granularity

between various disease evolution phases. Features like the edge orientation variability also resulted as being relevant, denoting the fact that the complexity of the cirrhotic tissue structure increased during the restructuring process.

After the feature selection process, the performance of the clustering methods was assessed again, under the hypothesis that there were 4 clusters in the data.

Thus, in the case of the EM method, the final value for log-likelihood was 58.57, estimated to a 78.90% probability for a certain data instance to belong to the corresponding cluster, while for X-means clustering, the final value for distortion was 27.15, better (smaller) than before feature selection.

The same values were obtained while employing the *Make Density based Clusterer* technique, in combination with the EM and X-means classifiers. In the case of the combination with X-means clustering, the log-likelihood provided by this method was 13.26 before feature selection, respectively 55.64 after feature selection. Comparing these values with those depicted in Table 1 and Table 2 it results that, in both cases, the performance was better after feature selection.

3.2.2. The effect of Principal Component Analysis (PCA)

The effect of the PCA technique upon the accuracy of the unsupervised classification process was studied as well. For the classical PCA method, the best results were obtained when taken into consideration 54 components, while for generalized PCA, the best results were obtained in the case of the Gaussian kernel, for 50 components. The initial dataset contained 77 attributes.

The values of the classification performance parameters obtained before and after the application of the PCA method were depicted in Table 4. Both the EM and X-means clustering methods were considered, and also the *Make Density based Clusterer* technique combined with the two basic classifiers was evaluated.

Table 4. The effect of PCA upon the unsupervised classification process.

	EM – (log-likelih.)	MakeDen-sity + EM (log – likelihood)	XMeans Distortion	MakeDen-sity + Xmeans Distortion
Origi-nal set	14.05	14.05	30.61	30.61
PCA	-88.44	-88.44	27.75	27.75
Kernel PCA	*34.57*	*34.57*	*27.73*	*27.73*

According to Table 4, it resulted that the classical PCA method led to a decrease of the classification performance, in the case of the EM method, while in the case of the X-means clustering technique, it led to an improvement of this performance. The kernel PCA technique with Gaussian kernel always led to improvements of the classification performance, for all the considered unsupervised classifiers. However, the accuracy of the unsupervised classification process was better after performing feature selection, than in the case of applying the PCA techniques.

3.3. The specific values for the relevant textural features

The values of the normalized arithmetic means for some of the most relevant textural features corresponding to each cluster center were illustrated in the figure below (Fig. 2). As it results from this figure, the c*luster* $\boldsymbol{c_0}$ corresponded, most likely, to the incipient stage of the disease, as the GLCM homogeneity had maximum value, while the entropy, contrast and variance derived from the second and third order TMCM matrices had minimum values. The *cluster* $\boldsymbol{c_3}$ corresponded, most probably, to the most advanced disease stage (when the tumors begin to appear), as the GLCM homogeneity had a minimum value, the entropy computed from the TMCM matrix based on spot microstructures had increased values, while the contrast and variance based on the second and third order TMCM had maximum values, denoting an increased structural complexity, due to the advanced restructuring process. The clusters $\boldsymbol{c_1}$ and $\boldsymbol{c_2}$ corresponded to intermediate evolution stages. In the case of cluster $\boldsymbol{c_2}$, the entropy derived from the TMCM matrix based on spot microstructures had a high value, this fact leading us to the conclusion that class $\boldsymbol{c_2}$ corresponded to a more advanced disease evolution phase, preceding the final stage. The probability density table, obtained by applying the method of Bayesian Belief Networks with K_2 search and BMA estimation (Weka, 2015), for the variance parameter derived from the third order TMCM matrix based on R_5S_5 microstructures is illustrated in Table 5. For the cluster $\boldsymbol{c_0}$ (the incipient stage) the above mentioned parameter was situated, with the highest probability (91.3%), in the interval of the lowest values. For the most advanced stage (cluster c_3), the third order TMCM variance belonged, most likely (*60%*), to the interval that corresponded to the highest values, denoting an increased tissue structure complexity, while for the intermediate stages (clusters c_1 and c_2), the considered third order TMCM variance was situated in the interval of middle values (92% for cluster $\boldsymbol{c_1}$, respectively 97.4% for cluster c_2). The same method was applied as well for the other features, in order to determine their variation intervals for each class, the final purpose being to provide a reliable instrument for computer assisted diagnosis.

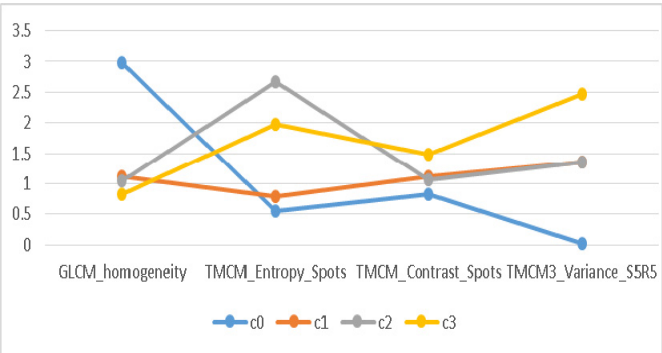


Fig. 2. The specific values of the most relevant textural features, for each cluster.

Table 5. The probability distribution table for the $TMCM_3\_R_5S_5\_variance$.

Class	(-∞, 742.36]	(742.36, 1013.32]	(1013.32, ∞)
c1	0.04	*0.92*	0.04
c2	0.013	*0.974*	0.013
c3	0.067	0.333	*0.6*
c0	*0.913*	0.044	0.043

The existing correlations between the newly defined TMCM features in the case of cirrhosis were also assessed, using the functions of the IBM SPSS Statistics environment. The strongest correlations (R-Square > 0.990) were noticed among the variances of the second order TMCM matrix, respectively of the third order TMCM matrix, based on the same textural microstructures. Other cases of relevant correlations resulted between the parameters derived from the third order TMCM matrix based on S_5R_5 microstructures, respectively from the third order TMCM matrix based on R_5S_5 microstructures, especially between the entropies and homogeneities.

Thus, in the case of the entropies, all the correlation types had increased values, the R-Square coefficient being around 0.990. In the case of the homogeneities, the linear and logarithmic correlations had associated the most increased values of the R-Square parameter (0.779 for linear correlation, respectively 0.772 for logarithmic correlation). These results were graphically illustrated in the next figure (Fig. 3) and indicated that, although the R_5S_5 and S_5R_5 convolution kernels were mathematically different, the entropy and homogeneity features varied in the same way in the case of cirrhosis.

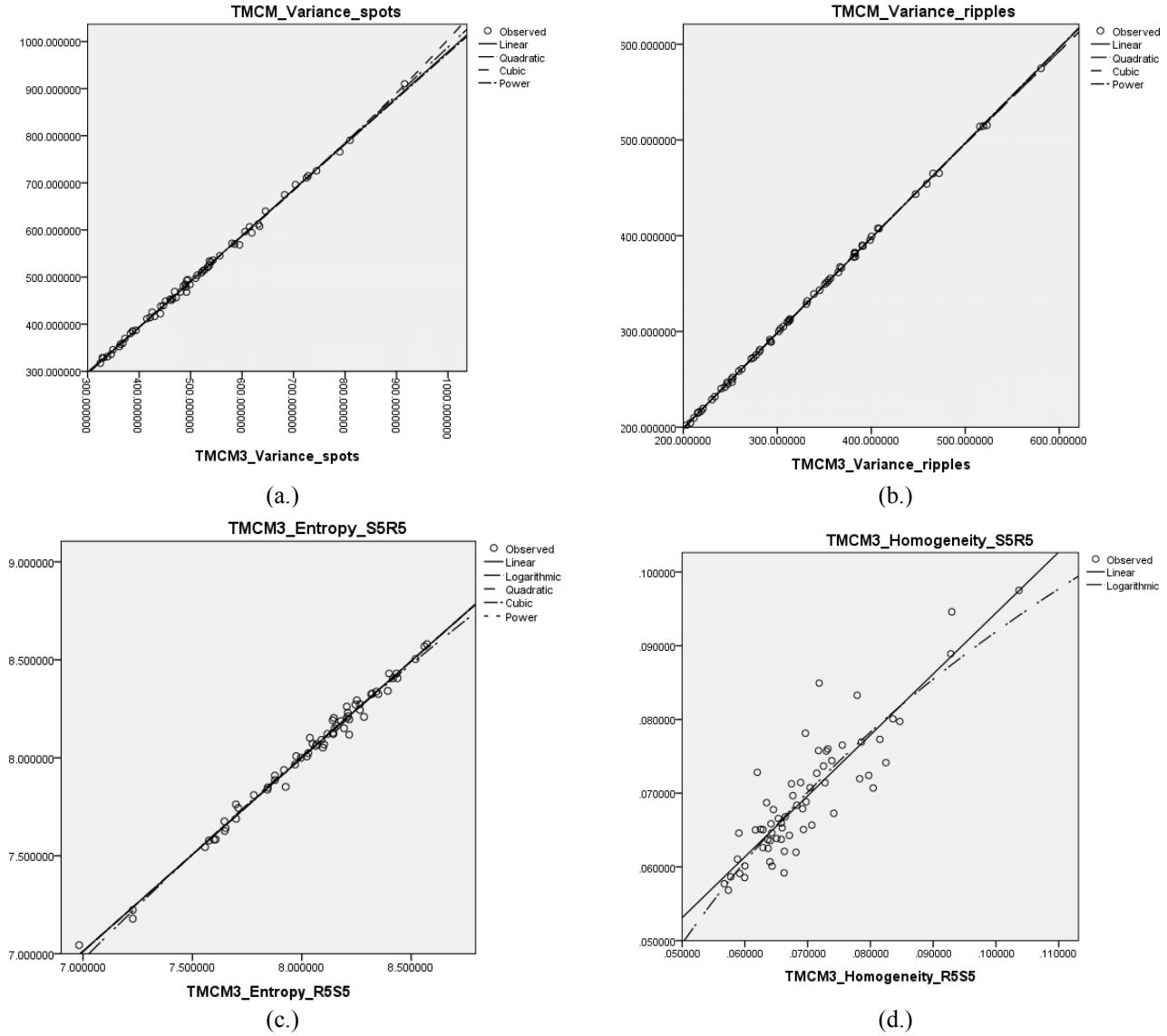


Fig. 3. The correlations between the parameters: (a.) *TMCM_Variance_Spots* and *$TMCM_3$_Variance_Spots* (b.) *TMCM_Variance_Ripples* and *$TMCM_3$_Variance_Ripples* (c.) *$TMCM_3$_Entropy_S_5R_5*, and *$TMCM_3$_Entropy_R_5S_5*; (d.) *$TMCM_3$_Homogeneity_S_5R_5*, respectively *$TMCM_3$_Homogeneity_R_5S_5*.

3.4. The validation phase

3.4.1. Validation through supervised classification

In order to assess the value of the model discovered through unsupervised classification, respectively the power of the newly defined textural features in the context of cirrhosis severity grade detection, the following classifiers from the Weka 3.6 library were employed: the SMO classifier, standing for the SVM method (Weka, 2015), having a 3rd degree polynomial kernel, this configuration corresponding to the best obtained results; the Multilayer Perceptron (MLP) classification technique, having a learning rate of 0.2, a momentum (α) of 0.8, and the number of nodes from the unique hidden layer equal with *a*, where *a=(nr_input_features + nr_classes)/2*; the J48 method, the Weka equivalent of the C4.5 algorithm for decision trees. Also, the *multiclass* meta-classifier was employed, which reduced the classification process to a set of binary classifications, in combination with the above-mentioned elementary classifiers and the Exhaustive correction code strategy. A cross-validation process with 5 folds was performed in order to assess the classification performance.

Concerning the classification performance parameters in the case when taking into consideration the set containing both the formerly existing and the newly defined textural features, the maximum recognition rate, of 96.29%, was obtained in the case of the Multilayer Perceptron (MLP) classifier. The comparison between the classification accuracies obtained when considering only the old textural feature set, respectively the set formed by the old and new textural features is illustrated in Fig. 4. It results an increase in accuracy, for all the classifiers, due to the newly defined textural features.

The accuracy of distinguishing the most advanced evolution phase (cluster c_3) from the other severity grades of cirrhosis was assessed as well. The AdaBoostM1 meta-classifier in combination with J48 was taken into consideration, as well (Weka, 2015). The classification performance is illustrated in Table 6. The maximum accuracy (recognition rate), of 93.75%, respectively the maximum AuC, of 96.6%, resulted in the case of the MLP classification technique. Increased specificity (TN rate) values can be observed for all the classifiers, expressing a reduced risk of an erroneous diagnostic in the case of the healthy patients.

The set of the relevant textural features that best separate the most advanced evolution phase of cirrhosis, from the other evolution phases, was also derived, using CFS method, specific for supervised classification (Weka, 2015). The resulted relevant feature set is provided in (7).

{GLCM_Energy, TMCM_Contrast_S5R5, TMCM3_Contrast_S5R5, TMCM_Homogeneity_S5R5, TMCM3_Entropy_S5S5, TMCM_Entropy_R5R5, TMCM3_Entropy_S5R5, TMCM3_Entropy_R5S5, Level_Frequency} (7)

Table 6. The classification performance resulted when distinguishing class c_3 from the other classes.

	Recog. Rate	TP Rate	TN Rate	AuC
MLP	93.75%	90.9%	94.3%	96.6%
SVM	92.18%	81.8%	94.3%	88.1%
J48	89.06%	63.6%	94.3%	79.5%
AdaBoost+J48	92.18%	72.7%	96.2%	81.7%

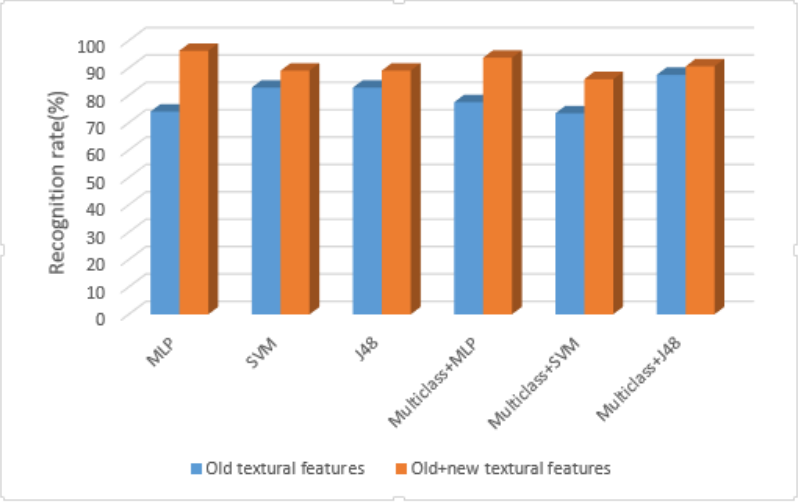


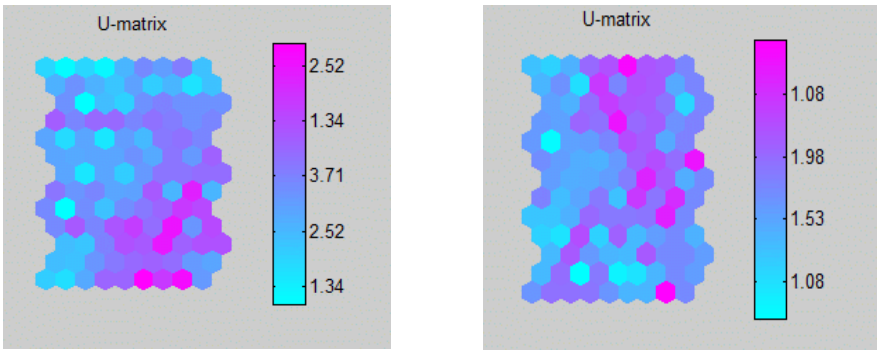
Fig. 4. The increase in accuracy due to the newly defined textural features.

In (7), the presence of the GLCM energy, of the contrast derived from the second and third order TMCM matrix based on combined microstructures (S_5R_5), of the entropies derived from the second and third order TMCM matrices based on spot, ripple and combined microstructures, of the homogeneity derived from the 3rd order TMCM matrix based on combined microstructures (*S5R5*), as well as of the level frequency, was noticed. All these features emphasized the chaotic, heterogeneous and complex structure of the liver parenchyma affected by the last cirrhosis stage and also the importance of the TMCM based features in characterizing the evolution of the restructuring process.

3.4.2. Modeling the clusters using SOM

In order to visualize the data structure, with respect to the existing classes, the method of Self Organizing Maps (SOM) was applied, before and after the selection of the relevant textural features. For SOM generation, the default option was used, corresponding to the batch-training algorithm.

This implied the existence of a first training phase, in which both the learning rate and the neighborhood radius were large, and also of a second phase, of refined training, when both the learning rate and the neighborhood radius took small values (SOM, 2005).



(a.) (b.)

Fig. 5. The U-Matrices corresponding to the Self Organizing Maps (SOM) obtained before (a.) and after (b.) feature selection.

As visualization form, the Unified Distance Matrix (U-Matrix) was chosen, in order to emphasize the distances between neighboring map units, and to provide the cluster structure of the map. In this context, high values of the U-matrix indicated a cluster border, while uniform areas of low values indicated the clusters themselves (SOM, 2005). The SOM structures, in the form of the U-matrices for the original data (a.), as well as for the data obtained after the selection of the relevant textural features (b.), are illustrated in Fig. 5. Thus, in Fig. 5 (a.), the existing clusters can be hardly distinguished, while in Fig. 5 (b.), these classes are more obvious, as the cluster edges, marked by high intensity elements (most of them being traced with dark pink) are better emphasized and the color contrast between the various groups is higher.

4. CONCLUSIONS

The textural features, especially those derived from the TMCM matrices, proved to be efficient concerning the detection and characterization of the cirrhosis severity grades, improving the previously obtained results in this domain. The unsupervised classifiers revealed the existence of four evolution stages for cirrhosis. The relevant textural features and their specific values for each severity grade confirmed the existing medical knowledge in the domain and brought new information, concerning the differences in granularity between the evolution phases of the considered disease. Also, the newly defined textural features led to an increase in accuracy in comparison with the old textural feature set, the maximum obtained recognition rates being above 90%. In the context of our future research, the role that the co-occurrence matrices based on more complex textural microstructures have in the detection and characterization of the cirrhosis severity grades will be also analyzed. Other dimensionality reduction methods, such as Multidimensional Scaling (MDS), Laplacian Eigenmaps (Van der Maaten, 2008), Independent Component Analysis (ICA) (Yang, 2014), respectively the Curvilinear Component Analysis (CCA) (Demartines, 1997) will be employed as well and their performance will be compared with that corresponding to the already employed equivalent technique (PCA). More powerful unsupervised classification methods based on optimization techniques (Das, 2008) will be considered, as well. The experiments will be performed on larger datasets, for the validation of the results. A score that quantifies the cirrhosis severity grades will be also derived, based on a weighted mean that considers the normalized values of the relevant textural features and their relevance index, as described by (8), where f_i are the normalized values of the relevant textural features, r_i is the relevance of the feature f_i and n is the number of the relevant textural features:

$$R = \sum_{i=1}^{n} r_i f_i \tag{8}$$

Using the formula (8) and considering the specific values of the relevant textural features for each cirrhosis severity grade, as indicated by technique of Bayesian Belief Networks, the value R can be associated to a score that indicates the cirrhosis severity grade, each severity grade corresponding to a specific interval of values of R. Thus, this score will be employed in order to monitor the disease evolution and when its value indicates certain risks, the medical treatment, as well as the corresponding examination frequency can be adapted accordingly. For validation, this score will be compared with the already existing scores for cirrhosis severity quantification, such as the Child-Turcotte-Pugh score, or the Model for End Stage Liver Disease (MELD) score (Peters, 2013). Establishing correlations between the newly defined score and the already existing scores, as well as their combined use, will be also taken into account, the final purpose being the increase in accuracy of the cirrhosis severity grading and the improvement of the survival chances for the affected patients.

ACKNOWLEDGEMENT

This paper was supported by the Post-Doctoral Programme POSDRU/159/1.5/S/137516, project co-funded from European Social Fund through the Human Resources Sectorial Operational Programme 2007-2013.

REFERENCES

Aly, M. (2005). Survey on multiclass classification methods. Online: http://www.vision.caltech.edu/ malaa/research/aly05multiclass.pdf

Bedossa, P. Poynard T. (1996). An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology*, 24, 289-293.

Carr, B.I. (2010). *Hepatocellular carcinoma: diagnosis and treatment (2nd edition)*. Humana Press, a part of Springer Science + Business Media. New York.

Cavouras, D., Kandarakis, I., and Other (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. *Medical Informatics Europe*, 522-528. IOS Press.

Ciocchetta, F., Dell' Anna, R. and Other (2000). Combining Supervised and Unsupervised Methods to Support Early Diagnosis of Hepatocellular Carcinoma. *Artificial Intelligence in Medicine,* 2780, 239-243.

Das, S. and Others (2008). Particle Swarm Optimization and Differential Evolution Algorithms: Technical Analysis, Applications and Hybridization Perspectives. *Studies in Computational Intelligence,* 116, 1–38 .

Davis, L.S. (1981). Image Texture Analysis Techniques – A Survey. *Digital Image Processing*, Simon and R. M. Haralick (eds.), 189-201.

Demartines, P., Herault, J. (1997). Curvilinear Component Analysis: A Self-Organizing Neural Network for Nonlinear Mapping of Data Sets*, IEEE Transacyions on Neural Networks*, 8(1), 148-154

Fujino, K., Mitani, Y. (2014). Liver Cirrhosis Classification on M-Mode Ultrasound Images by Higher-Order Local Auto-Correlation Features. *Journal of Medical and Bioengineering,* 3(1), 29-32.

IBM SPSS Statistics (2015). Online:http://www-01.ibm.com/software/analytics/spss/products/statistics/

Jalobeanu, C., Rasa, I. (2001). *Incertitudine si decizie. Statistica si probabilitati aplicate in mangement*. Editura UT Press, Cluj-Napoca.

Laws, K. (1980). Rapid texture identificaton. *SPIE: Image Processing for Missile Guidance,* 238, 376-380.

Lee, G., Kanematsu, M., Kato, H. (2008). Unsupervised classification of cirrhotic livers using MRI data. *Proceedings of SPIE*, 6915, 69514-1 – 69514-8.

Mala, K., Sadasivam, V. (2010). Classification of fatty and cirrhosis liver using wavelet-based statistical texture features and neural network classifier. *International Journal on Software Informatics*, 4(2), 151-163.

Matlab Toolbox for Dimensionality Reduction, ver 7.0., 2008. Online: http://lvdmaaten.github.io/drtoolbox/

Mitrea, D., Nedevschi, S., Socaciu, M., Badea, R. (2012), The role of the superior order GLCM in the characterization and recognition of the liver tumors from ultrasound images. *Radioengineering Journal*, 21(1), 79-85.

Mitrea, D., Platon-Lupsor, M, Nedevschi, S., Badea, R. (2013). Discovering the cirrhosis grades from ultrasound images by using textural features and clustering methods. *Proc. of the 36th International Conference on Telecommunications and Signal Processing (TSP),* 2013, 633-637.

Mitrea, D., Nedevschi, S., Badea, R. (2014). The Role of the Textural Microstructure Coocurrence Matrices in the Classification of the Abdominal Tumors. *Proc. of the 10th International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, September 4-6 2014,* 187-190, Cluj-Napoca.

Pelleg, D., Moore, A. (2000). X-means: Extending K-means with Efficient Estimation of the Number of Clusters, *Proc. of International Conference on Machine Learning (ICML)*, 727-734. Stanford, CA, USA.

Peters, M., Carbone, J. (2013). Evaluation and prognosis of patients with cirrhosis. Online: https://depts.washington.edu/hepstudy/presentations/uploads/19/m2_l5_evaluation_and_prognosis_of_patients_with_cirrhosis.pdf

Shiha, G., Zalata, K. (2011). Ishak versus METAVIR: Terminology, Convertibility and Correlation with Laboratory Changes in Chronic Hepatitis C. Technical report. Online: http://cdn.intechopen.com/pdfs-wm/18773.pdf

SOM 2.0 toolbox for Matlab (2005). Online: http://www.cis.hut.fi/somtoolbox/.

Toennies, K., D. (2012). *Guide to medical image analysis. Methods and algorithms*, Springer Verlag, London.

Van der Maaten, L.J.P, Postma, E.O. (2008). Dimensionality reduction: A comparative review. Online:http://www.iai.uni-bonn.de/~jz/dimensionality_reduction_a_comparative_review.pdf

Weka 3, Data Mining Software in Java (2015). Online: http://www.cs.waikato.ac.nz/ml/weka/.

Witten, I., Frank, E. (2011). *Data Mining – Practical Machine Learning Tools and Techniques (third edition)*, Morgan Kaufmann, Massachusetts.

Yang, G., Raschke, F., Barrick, T., Howe, F. (2014). Manifold Learning in MR spectroscopy using nonlinear dimensionality reduction and unsupervised clustering, *Magnetic Resonance in Medicine*. Online: http://onlinelibrary.wiley.com/doi/10.1002/mrm.25447/abstract

Yin, H., (2008). The Self-Organizing Maps: Background, Theories, Extensions and Applications. *Computational Intelligence*, 115, 715–762.

Zaki, M, Meira, W (2014). *Data Mining and Analysis. Fundamental Concepts and Algorithms*. Cambridge University Press. Cambridge.