Revista Romana de Interactiune Om-Calculator 9 (1) 2016, 53-70

Software system for the automatic and computer assisted diagnosis of some severe abdominal affections, based on ultrasound images

Delia Mitrea[1], Sergiu Nedevschi[1], Radu Badea[2]

[1] Technical University Cluj-Napoca, Faculty of Automation and Computer Science, Department of Computer Science
Baritiu Str., No. 26-28, Cluj-Napoca
E-mail: delia.mitrea@cs.utcluj.ro

[2]Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Department of Medical Imaging
V. Babes Str., No.8, Cluj-Napoca
E-mail: rbadea@umfutcluj.ro

Abstract. In this article we describe specific methods for the characterization and computer assisted diagnosis of some severe abdominal diseases, based on ultrasound images and the corresponding software system. The objective of our research is to build an instrument for the non-invasive diagnosis, based on virtual biopsy, of some severe abdominal diseases such as the cirrhosis, the hepatocellular carcinoma (HCC) and the colorectal tumors. The gold standard for the diagnosis of these affections is the classical biopsy, but this is invasive, dangerous. Ultrasonography is a medical-imaging based investigation method that is noninvasive, inexpensive, safe and thus repeatable. Texture is an important property of the human body tissues, able to reveal subtle characteristics upon the pathology. We defined the imagistic textural model of these affections, using the information obtained from ultrasound images. Based on the imagistic textural model, we assessed the possibility of performing automatic recognition of the considered affections both in supervised and unsupervised manner, the purpose of the latter approach being that of discovering the severity grades of the diseases. The resulted accuracy was above 90% in most of the cases.

Keywords: Automatic and computer assisted diagnosis, non-invasive characterization, ultrasound images, texture, virtual biopsy, classification performance.

1. Introduction

The malignant tumors represent an important issue nowadays, as they can irremediably affect the human body. The hepatocellular carcinoma (HCC) is the most frequent malignant liver tumor, appearing in 75% of the liver

cancer cases. It evolves from cirrhosis, after a liver parenchyma restructuring phase at the end of which dysplastic nodules (future malignant tumors) result (Bruix, 2003). The colorectal tumors are malignant tumors as well, that frequently affect the population of the developed countries (Ruess, 2000). The early detection of all these affections could lead to an increase of the patient survival rate and also to metastasis prevention. The biopsy is among the few reliable methods for the diagnosis of these diseases, representing nowadays the « gold standard » in this domain, being, however, invasive, dangerous for the patient, as it could lead to infections and to the spread of the tumor inside the human body. Ultrasonography is a non-invasive method for patient examination, non-invasive, safe and thus repeatable, appropriate for a periodical monitoring of the patient state. Other examination methods, such as the Computer Tomography (CT), the Magnetic Resonance Imaging (MRI), the endoscopy or the Contrast Enhanced Ultrasonography (CEUS) could be expensive, invasive or dangerous. Considering the objective of non-invasive diagnosis, aiming to exceed the capacities of the medical specialist eye, we developed computerized methods for the characterization and automatic recognition of the considered affections, based only on the information derived from ultrasound images. Our methods will contribute to the achievement of a virtual biopsy system, in the context of the prevention and early diagnosis of the above mentioned severe affections aiming the following functions: the *characterization of the analyzed tissue through textural features, correlated with the physical and chemical properties of this tissue; the detection of the pathology class where the tissue belongs.* Texture being an important property of the human body tissues, able to reveal subtle information concerning the pathology, we adopted existing texture analysis methods in our approach and we developed, as well, original, advanced texture analysis methods, based on superior order, generalized co-occurrence matrices. We defined the imagistic textural model of the considered affections, consisting of the complete set of the relevant textural features able to characterize these affections and to distinguish them from visually similar tissues, respectively of the specific values associated to the relevant textural features (mean, standard deviation, probability distribution, confidence intervals). The imagistic textural model was validated through supervised classification methods and the accuracy of the automatic diagnosis was assessed in this manner. Also, unsupervised classification methods were implemented in

order to discover the severity grades of some affections. This research was performed in the context of the collaboration with the collective of specialist physicians from the Third Medical Clinic of Cluj-Napoca.

The texture-based methods were widely employed for computer-aided diagnosis based on medical images (Sujana el al, 1996), (Chikui et al, 2005), (Yoshida et al, 2003), (Madhabushi et al, 2005). However, the imagistic textural model was not approached before in this domain and there doesn't exist any significant study concerning the unsupervised detection of the evolution stages of the considered affections, based on ultrasound images. The affections considered for analysis were the following: cirrhosis, HCC, and the colorectal tumors. We also included in the experimental dataset the cirrhotic parenchyma around HCC, benign liver tumors (hemangioma) and Inflammatory Bowel Diseases (IBD), as they have a similar visual aspect with the colorectal tumors. The software system for the characterization and diagnosis of the above mentioned affections performs the functions:

- The computation of the textural features through classical and advanced texture analysis methods;
- The computation of the relevant textural feature set, by applying dimensionality reduction methods and specific techniques for feature selection (Van der Maaten, 2008);
- The visualization of the effects that various texture analysis methods produce upon the image, respectively the export of values of the relevant textural features in various formats: "arff", the specific Weka format (Weka, 2016), respectively the Excel format;
- The computation of the specific values for the relevant textural features and their storage in the database in association to the corresponding class: arithmetic mean, standard deviation, probability distribution, confidence intervals;
- Training of performant supervised classifiers using the values of the relevant textural features;
- The implementation of the trained classifiers in the context of the automatic diagnosis of the considered affections.
- The function concerning the unsupervised detection of the severity grades is still under construction.

The content of the rest of this article is structured in the following

manner: first, the state of the art is presented and the contribution of the current work is highlighted; then the adopted methods and the software system are described in details; the experiments and the results are presented and analyzed; at the end, the conclusions are drawn and further development possibilities are proposed.

2. The state of the art

The texture-based methods were widely employed nowadays, for the characterization and recognition of various diseases, based on ultrasound images. Considering the purpose of malignant tumor recognition, multiple methods of increased complexity were implemented, in conjunction with performant classifiers. In (Sujana et al, 1996), the authors determined first order statistics of the grey levels, such as the grey level mean and variance, the Haralick parameters derived from the Grey Level Cooccurrence Matrix (GLCM) and also the Run-Length matrix parameters, together with a neural network classifier, as well as with a classifier based on linear discriminants, for the differentiation between the hepatic malignant tumors, the benign tumors and the normal liver. In the case of the linear discriminant based classifier, the recognition rate was 79.6%, while in the case of the neural network classifier, a recognition rate of 100% resulted (Sujana et al, 1996). The fractal based methods were implemented in (Chikui et al, 2003), in order to detect the salivary gland tumors from ultrasound images. The Wavelet transform was implemented as well, in order to analyze the values of the textural features at multiple resolutions, the objective being that of differentiating the malignant and benign liver tumors. This method provided satisfying results, the area under the ROC curve being 90%. (Yoshida et al, 2003). Aiming the automatic detection of the prostatic adenocarcinoma, based on magnetic resonance images (MRI), a method based on the Gabor transform, combined with the Haralick features derived from GLCM, and with Bayesian classifiers, was implemented (Madabhushi, 2005). The sensitivity was 42.35% in this case, better than that of the human experts (36.41%), while the specificity was 97.25%. An important approach concerning the recognition of the diffuse liver diseases from ultrasound images through supervised classification methods was described in (Cavouras et al, 1997). In order to differentiate among the diffuse liver

diseases, the authors employed a hierarchical tree: 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 et al, 1997). Concerning the unsupervised detection of the tumor evolution phases, several approaches exist, the most relevant being (Atupelage et al, 2013), (Ciocchetta et al, 2000), but no significant research exists regarding the automatic grading of the HCC severity, based on ultrasound images, in unsupervised manner. In (Atupelage et al, 2013), the authors determined the evolution stages of HCC, in supervised manner, from histological images, using newly defined textural features, derived through multifractal analysis. The final resulted accuracy was 95%. Another approach combined the supervised and unsupervised classification techniques, in order to detect the HCC tumor in incipient phase, using histological features (Ciocchetta et al, 2000). Regarding the unsupervised classification of other malignant diseases, the authors assessed the role of the dimensionality reduction methods, in the context of the differentiation among the glioma brain tumor evolution phases, based on spectroscopic image (Resmi et al, 2010).

Taking into consideration all these methods, it results the importance of the texture based techniques concerning the detection of the various diseases from medical images. Thus, the computerized, texture-based methods can substantially contribute to the early, non-invasive detection of some severe diseases, which could threaten the human life. However, there doesn't exist any significant approach based on ultrasound images that defines an imagistic textural model for the considered affections and for their evolution phases, in order to highlight the set of the relevant textural features and the specific values associated to them for each class; also the accuracy of the already existing methods can be further improved. In our approach, we aimed to define the imagistic textural model of the considered diseases and to perform the correlation between the textural features and the structural properties of the tissue affected by disease, the final purpose

being the virtual biopsy. The software system that will be described in the following paragraphs implements the method that generates and assesses this imagistic textural model, in both cases of supervised and unsupervised classification.

3. The proposed solution

3.1. The imagistic textural model

The imagistic textural model of HCC consists of: *the complete set of the relevant textural features* for the characterization of a certain disease, considered in their optimal representation, after applying the appropriate dimensionality reduction methods (feature extraction and feature selection techniques); *the specific values of the relevant textural features:* arithmetic mean, standard deviation, probability distribution (Mitrea D. et al, 2012). In order to generate the imagistic textural model, the following phases are necessary: (1) an image analysis phase, when the potentially relevant textural features are computed; (2) a learning phase, involving the application of the dimensionality reduction methods, the computation of the specific values of the relevant textural features and the automatic discovery of the disease evolution phases; (3) a validation phase, consisting of applying supervised classifiers and data representation methods in order to assess the classification performance due to the imagistic textural model.

3.2. Methods used during the image analysis phase

With the purpose of characterizing with maximum accuracy the tissue affected by disease, as it appears in ultrasound images, already existing, classical texture analysis methods were applied and advanced texture analysis methods were elaborated as well. The following features were computed through classical texture analysis methods : first order statistics of the grey levels, edge based statistics and other gradient based features (directional gradient variance), statistics of the textural microstructures, obtained after applying the Laws' convolution filters, the Gray Level Cooccurrence Matrix (GLCM), the Shannon entropy computed after applying the Wavelet transform, at two resolution levels. In order to refine

the discriminative power of the textural features, we defined generalized cooccurrence matrices of second and superior order, as described by (1):

$$\begin{aligned} &C_D(\nu_1,\nu_2,\nu_3,..,\nu_n) = \#\{((x_1,y_1),(x_2,y_2),(x_3,y_3),..,(x_n,y_n)):\\ &A(x_1,y_1)=\nu_1, A(x_2,y_2)=\nu_2,..,A(x_n,y_n)=\nu_n,\\ &|x_2-x_1| = |\vec{dx}_1|, x_3-x_1| = |\vec{dx}_2|,..,|x_n-x_1| = |\vec{dx}_{n-1}|,\\ &|y_2-y_1| = |\vec{dy}_1|,|y_3-y_1| = |\vec{dy}_2|,..,|y_n-y_1| = |\vec{dy}_{n-1}|,\\ &\mathrm{sgn}((x_2-x_1)(y_2-y_1)) = \mathrm{sgn}(\vec{dx}_1\cdot\vec{dy}_1),..\\ &,\mathrm{sgn}((x_n-x_1)(y_n-y_1)) = \mathrm{sgn}(\vec{dx}_{n-1}\cdot\vec{dy}_{n-1})\} \end{aligned} \quad (1)$$

In (1), #S represents the size of the set S, each element of this matrix being equal with the number of n-tuples of pixels, having the spatial coordinates (xi, yi) and the value υi for the attribute A. The spatial relation between the pixels (xi, yi) was defined by the set of the displacement vectors, depicted in (2):

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

The attribute A is the grey level of the pixel in the case of the superior order GLCM matrix (Mitrea D. et al, 2012), the edge orientation in the case of the Edge Orientation Co-occurrence Matrix (EOCM) (Mitrea D. et al, 2012), respectively the label associated to a complex textural microstructure, in the case of the Complex Textural Microstructure Co-occurrence Matrix (CTMCM), respectively in that of the Complex Extended Textural Microstructure Cooccurrence Matrix (CETMCM). The CTMCM matrix was defined after applying the Laws' convolution filters, followed by an improved k-means clustering algorithm (Mitrea D. et al, 2015). The clusters provided by this algorithm stood for the complex textural microstructures. The CETMCM matrix was derived after applying both the Laws' convolution filters and the edge detection techniques, followed by the improved k-means clustering algorithm (Mitrea D. et al, 2014). The GLCM matrix of order 2, 3, 5, 7, as well as the EOCM, CTMCM and CETMCM matrices of order 2 and 3, were computed as in (Mitrea D. et al, 2012).

3.3. The learning phase

During the learning phase, the data dimensionality reduction was performed

first, in most of the cases, by applying both feature extraction and feature selection methods. By using the feature selection methods we aimed to separate the relevant textural features from the non-relevant ones. By using the feature extraction methods, we aimed to transpose the initial dataset into a new space, where to highlight some properties of the data, such as the main variation modes (in the case of PCA) and their capacity to separate among classes, in the LDA and MDA cases (Van der Maaten et al, 2008). All the considered methods, as well as their combinations, were compared from the performance point of view. The best resulted techniques were the feature selection methods, especially the Correlation based Feature Selection (CFS) and the Information Gain Attribute Evaluation (Hall, 2003). The probability distributions of the relevant features were determined using the method of Bayesian Belief Networks (Witten et al, 2005). Unsupervised classification methods were also applied in this phase, when aiming to discover the cirrhosis and HCC evolution phases. We considered the following methods: k-means and X-means clustering, Expectation Maximization (EM), Particle Swarm Optimization (PSO) combined with k-means clustering (Mitrea D. et al, 2016). In this case, the feature selection was performed after the cluster discovery process (Mitrea D. et al, 2016) and the final clusters were established afterwards.

3.4. The validation phase

In order to evaluate the model resulted during the learning phase, the relevant textural feature values were provided as inputs of supervised classifiers. Several classification techniques, including both basic learners and classifier combination schemes (boosting, bagging and stacking) were compared in our experiments, but the following methods provided the best results: the Multilayer Perceptron (MLP), the Support Vector Machines (SVM), the C4.5 and Random Forest (RF) algorithms based on decision trees, and also AdaBoost combined with C4.5 (Witten et al, 2005). The strategy of cross-validation with 5 folds was applied and the following parameters were considered: the recognition rate (classification accuracy), the sensitivity (TP rate), specificity (TN rate), the area under the ROC curve (AUC).

4. The software system that implements the imagistic textural model

4.1. Purpose and functionalities

The software system that implements the imagistic textural model performs both the characterization of the ultrasound images belonging to patients, which are diagnosed or suspect of the considered affections and also the automatic recognition of these diseases. The types of users considered during the design and implementation are described further. ***The administrator*** generates the values of the textural features resulted after the application of the texture analysis methods upon the images in the training set, each of these having a region of interest inside the tissue of interest (tumor, liver parenchyma or bowel wall) which are automatically stored into the database; applies the dimensionality reduction methods upon the generated set of values, obtaining relevant textural features, in the most appropriate data representation space; determines the specific values of the relevant textural features (mean, standard deviation, confidence intervals, probability distribution); trains the classifier using the set of the generated values corresponding to the relevant textural features. ***The medical specialist*** selects a rectangular region of interest (ROI) inside the suspect area; visualizes the values of the textural features that correspond to that ROI and the effect of the texture analysis methods upon the image; lists the specific values of the relevant textural features and the probability densities; detects the pathology class where the tissue of the ROI belongs.

4.2. The structure of the software system

The system architecture consists of three main component parts: the database; the interface for the insertion of the patient data and of the ultrasound images; the software application for the imagistic textural model generation. The structure of the software system is illustrated in Figure 1.

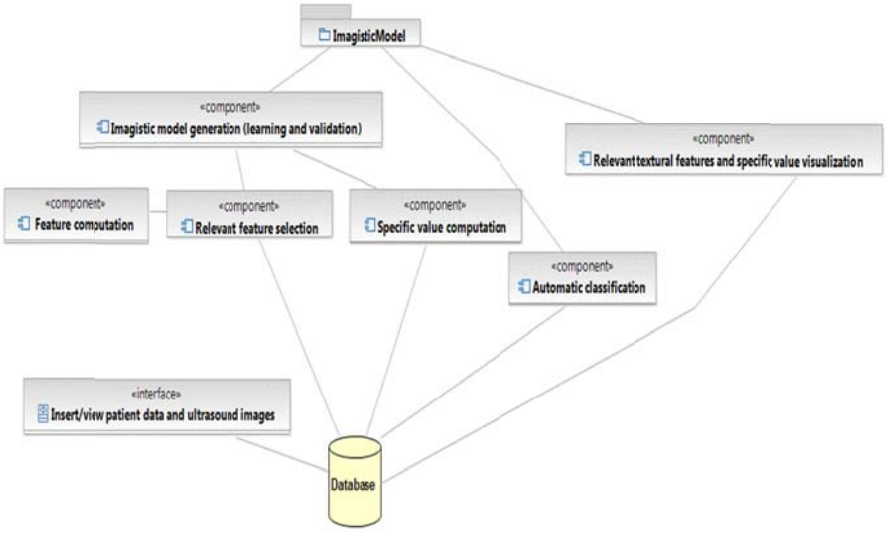


Figure 1. The structure of the software system

Both the database and the interface for the direct communication with the database were achieved in Microsoft Access. The interface for the communication with the database provides the possibility to insert and edit the patient data, as well as the data resulted after the examination, facilitates the choice of the ultrasound images from the disk for their storage into the database, offering also the possibility to select rectangular regions of interest on the ultrasound images, to specify the tissue type (class), respectively the characteristics of the region of interest.

The software application that builds the imagistic textural model consists of three main modules: the module corresponding to the generation (learning and validation) of the imagistic textural model; the module that contributes to the implementation of the model in the domain of computer aided design (provides the possibility to display the set of the relevant textural features and their specific values and also to visualize the effect of the texture analysis methods upon the images); the module corresponding to the automatic classification (automatic diagnosis), performed through any of the most performant classifiers detected through our experiments. The module for the generation of the imagistic textural model uses the components for relevant feature selection and specific value computation

and the relevant feature selection component communicates with the component that performs the computation of the textural features. The last version of this application was built in the Visual Studio 2010 environment, using the Visual C++ programming language. The functions corresponding to the learning and evaluation phases of the imagistic textural model were implemented using the Weka libray (Weka 3, 2016). The module that corresponds to the automatic diagnosis provides the possibility to select the region of interest inside the tissue affected by pathology and to perform the automatic classification of that tissue region using a trained classifier. The considered classifier was trained with the values of the relevant textural features, according to the imagistic textural model of the analyzed affection. The sub-menu corresponding to this module is illustrated in Figure 2.

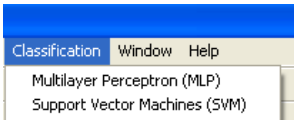


Figure 2. The sub-menu for automatic diagnosis (classification)

The classification result is provided then in a dialog box, where the patient diagnostic is stated, as indicated in Figure 3. The newly obtained result can also be stored in the training set for the further improvement of the classification performance.

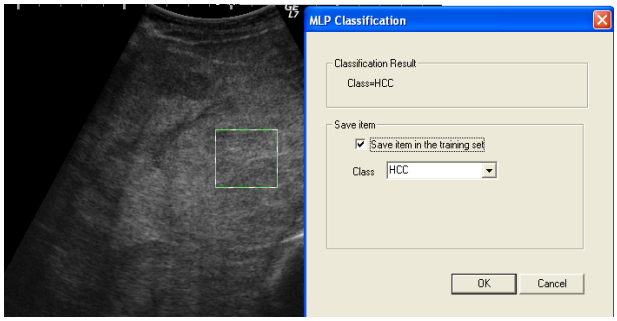


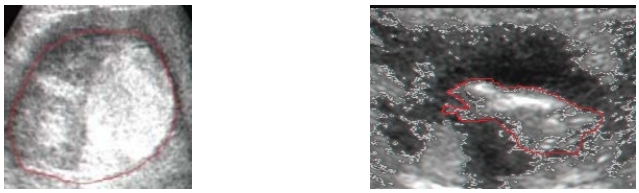
Figure 3. The result of the automatic diagnosis viewed in a dialog box

The software system was installed at the Third Medical Clinic in Cluj-Napoca, being used by the specialists in medical ultrasonography in the context of the computer aided diagnosis. The functions concerning the

automatic discovery of the evolution phases of the affections (cirrhosis and HCC) is still under construction. However, lots of experiments were performed in order to discover these evolution stages based on ultrasound images. The final purpose is to define a score, based on the relevant textural features and on their specific values, in order to non-invasively assess, with maximum accuracy, the severity grade of each considered disease. This score will be comparable with already existing scores, such as METAVIR, Child-Turcotte-Pugh, respectively the Model for End Stage Liver Disease (MELD).

5. Experimental results

The dataset taken into consideration for the experiments consisted of 300 cases (patients) of HCC, 100 cases of hemangioma (benign tumor), 100 cases of cirrhosis, 65 cases of colorectal tumors and 65 cases of Inflammatory Bowel Diseases (IBD). The images were acquired using a Logiq 7 ultrasound machine, under the same settings: frequency of 5.5 MHz, gain of 78, depth of 16 cm. For each patient, three images were considered, corresponding to various orientations of the transducer. In Figure 4, the maps of complex extended textural microstructures, resulted after the application of the Laws' convolution filters, of the edge detection filters and of the improved k-means clustering algorithm (Mitrea D. et al, 2014) are illustrated in the cases of HCC and colorectal tumors. As it results from this figure, this method is able to put into evidence important structures within the tumor tissue and thus it results that the Complex Extended Textural Microstructure Cooccurrence Matrix is appropriate for malignant tumor characterization.



(a) (b)

Figure 4. The complex extended textural microstructure maps in the cases of: (a) HCC (marked contour); (b) colorectal tumor (marked contour)

5.1. Experiments performed in the context of the supervised classification of the malignant tumors

For the selection of the relevant textural features, the techniques *CfsSubsetEval*, corresponding to the CFS method combined with genetic search, respectively the *InfoGainAttributeEval* method combined with the *Ranker* search method were adopted. In order to compute the probability distributions of the relevant textural features, the method of Bayesian Belief Networks, with K2 search and Bayesian Model Average (BMA) estimation was employed (Weka, 2016). Concerning the classification methods, the John's Platt Sequential Minimal Optimization Algorithm (SMO) of Weka 3.6, was adopted for implementing the SVM method, with a polynomial kernel of third degree and normalized input data. For the classifier of Multilayer Perceptron (MLP), we considered the specific method of Weka 3.6 (MultilayerPerceptron), containing, within the hidden layer, a number of nodes equal with the arithmetic mean between the number of features and the number of classes ; the learning rate was tuned to 0.2 and the momentum (α) was 0.8. The Random Forest (RF) classificatier, with 10 trees, was employed as well. The AdaBoostM1 meta-classifier of Weka was also implemented in conjunction with the J48 method, the equivalent of the C4.5 technique (Weka 3, 2016). Concerning the relevant textural features resulted in the case of tumor tissue characterization, the performed experiments indicated that the following features were the most important concerning the separation of the malignant tissue from the benign one: the homogeneity and energy based on the second and superior order GLCM, computed both in single and multiresolution manner; these features denoted the difference in homogeneity and echogenicity between the malignant and benign tissues, being well known that the malignant tumors have, in advanced evolution phases, a heterogeneous, hyperechogenic character, due to the coexistence of various region types (necrosis, fibrosis, active growth regions, fat cells). Another important feature was the GLCM entropy, denoting the chaotic character of the tumor tissue. The autocorrelation index was also important, putting into evidence differences in granularity between the malignant and benign tissues. Features like the Hurst fractal index, the edge frequency, the edge orientation variability, the density and frequency of the Laws' textural microstructures, as well as the parameters derived

from the Edge Orientation Cooccurrence Matrix (EOCM), from the Complex Textural Microstructure Cooccurrence Matrix and from the Complex Extended Textural Microstructure Cooccurrence Matrix (CETMCM) were relevant as well, illustrating the increased complexity of the tumor tissue structure. Figures 5 and 6 demonstrate the increase in accuracy due to the latest defined features based on the CTMCM, respectively on the CETMCM matrices, in the case of HCC and colorectal tumor recognition. The set formed by the old textural features and the CETMCM textural features was compared with the set formed by the old textural features and the CTMCM features, respectively with the set formed only by the old textural features. The sets including the CETMCM features and the CTMCM features led to an accuracy increase in most of the situations. In the case of the differentiation between HCC and the cirrhotic parenchyma on which HCC had evolved (Figure 5), the maximum recognition rate, of 84.33%, was obtained for the AdaBoost meta-classifier combined with the J48 method, which corresponded to the C4.5 method based on decision trees, when taking into account the CETMCM features. In the case of differentiation between HCC and hemangioma, the maximum recognition rate was 98.8% and resulted in the case of the SVM classifier (Mitrea D. et al, 2014).

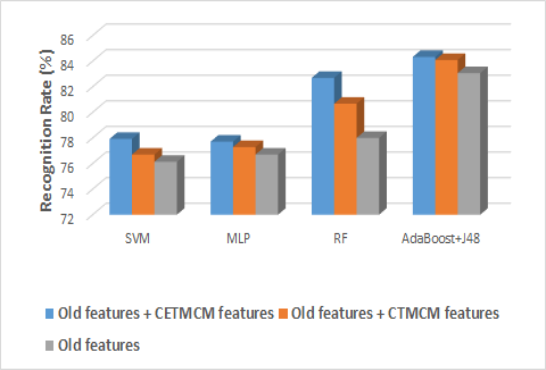


Figure 5. The role of the CTMCM and CETMCM features in the improvement of the classification accuracy, in the case of HCC/cirrhotic parenchyma differentiation

In the case of the differentiation between the colorectal tumors and IBD Figure 6), the maximum recognition rate, of 98.33%, resulted in the case of the Support Vector Machines (SVM) method (Mitrea D. et al, 2014).

Concerning the poorly classified cases, in both of the above described situations, these corresponded to tumor instances which were visually similar with other tissue classes (cirrhotic parenchyma on which HCC had evolved, hemangima, IBD). For further accuracy improvement, we aim to develop more advanced multiresolution techniques, based on the Wavelet and Gabor transforms.

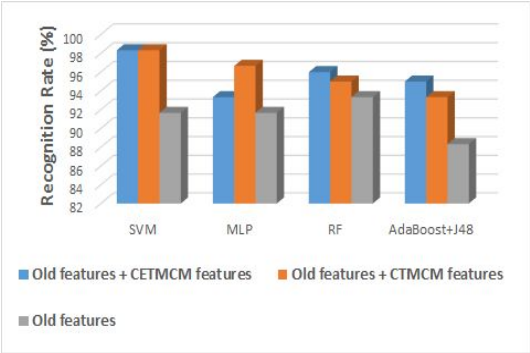


Figure 6. The role of the CETMCM features in the improvement of the classification accuracy, in the case of colorectal tumors/IBD differentiation

5.2. Experiments performed in the context of disease evolution phase discovery through unsupervised classification

As resulted from our experiments, four severity grades could be distinguished in the case of cirrhosis, based on ultrasound image information (Mitrea D. et al, 2013), while in the case of HCC, five evolution phases could be detected with maximum accuracy (Mitrea D. et al, 2016). Figure 7 depicts the five HCC evolution phases and the values of the arithmetic means of the most relevant textural features for each cluster). Cluster c_2 corresponded to the incipient phase, as the GLCM homogeneity had maximum value and the maximum value for the grey levels had small value, denoting hypoechogenicity. In this case, the third order CETCM energy had small values and the CETMCM cluster prominence had increased values, denoting the small density of the complex textural microstructures in these cases. Cluster c_5 corresponded to an advanced evolution phase, as the GLCM homogeneity had minimum value and the

maximum of the grey levels had increased value, due to the fat cells; the third order CETMCM energy had an increased value in this case, while the CETMCM cluster prominence had a small value, expressing an increased density of the complex extended textural microstructures. The other clusters (c_1, c_3 and c_4) corresponded to intermediate phases (Mitrea D. et al, 2016).

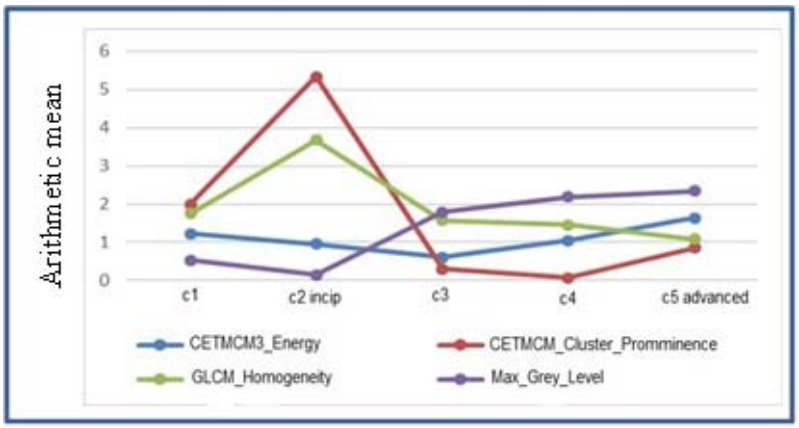


Figure 7. The values of the arithmetic means in the case of HCC evolution phase discovery

During the validation phase, when the obtained results were assessed through supervised classifiers, the maximum recognition rate in the case of cirrhosis severity grade detection was 97.87%, obtained for the J48 classifier, while in the case of HCC evolution phase detection, the maximum recognition rate, of 93.35%, was achieved in the case of the MLP classifier.

6. Conclusions and future work

The imagistic textural model proved to be efficient in the context of the automatic and computer assisted diagnosis of the considered affections. The newly defined textural features led to an increase in accuracy in comparison with the classical textural features, the final resulted accuracy being above 90%. After the validation of the experimental results on larger datasets, the functionalities concerning the automatic discovery and unsupervised detection of the disease evolution phases and of the corresponding relevant textural features will be integrated within the system as well, by storing the

corresponding data in the database and training the supervised classifiers in appropriate manner. Our software system was trained and experimented considering the cirrhosis, HCC and colo-rectal tumors, but it can be used also for characterizing and diagnosing other similar affections.

References

Atupelage, C., Nagahashi, H., 2013. Computational grading of hepatocellular carcinoma using multifractal feature description, Computers in Medical Images and Graphics, 37, 61-71

Bruix, J., 2003. Hepatitis B Virus and Hepatocellular Carcinoma, *Hepatology*, 39(1), S59-S63.

Cavouras, D., Kandarakis, I., 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.

Chikui, T., Tokumori, K., 2005. Sonographic texture characterization of salivary gland tumors by fractal analysis, *Ultrasound in Med. & Biol*., 31(10), 1297–1304.

Ciocchetta, F., et al, 2000. Combining Supervised and Unsupervised Methods to Support Early Diagnosis of Hepatocellular Carcinoma. *Artificial Intelligence in Medicine*, 2780, 239-243.

Hall, M., 2003. Benchmarking attribute selection techniques for discrete class data mining, *IEEE Trans On Knowledge and Data Eng.,* 15(3), 1-16.

Madabhushi, A., Felman, M.D., 2005. Automated Detection of Prostatic Adenocarcinoma From High-Resolution Ex Vivo MRI, *IEEE Trans. on Medical Imaging*, 1611-1626.

Mitrea, D.Mitrea, P., Nedevschi, S., et al, 2012. Abdominal tumor characterization and recognition using superior order cooccurrence matrices, based on ultrasound image. *Computational and Mathematical Methods in Medicine*. Online: http://www.hindawi.com/.

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, Roma, 633-637.

Mitrea, D., Nedevschi S., Abrudean, M., Badea, R., 2014. Abdominal tumor recognition from ultrasound images using Complex Extended Textural Microstructure Co-occurrence Matrices. *Automation, Computers and Applied Mathematics*, 23(1), 9-17.

Mitrea, D., Nedevschi, S., Abrudean, M., Badea, R., 2015. Colorectal cancer recognition from ultrasound images, using complex textural microstructure co-occurrence matrices, based on Laws' features, *Procs of the 38th IEEE International Conference on*

Telecommunications and Signal Processing (TSP), 2015, Praga, 458-462.

Mitrea, D., Nedevschi, S., Badea, R., 2016. The Role of the Complex Extended Textural Microstructure Co-occurrence Matrix in the Unsupervised Detection of the HCC Evolution Phases, based on Ultrasound Images. *Procs. of the 5th Conference on Pattern Recognition Applications and Methods (ICPRAM)*, 2016, Roma, 698-705.

Resmi, A., 2010. Texture description of low grade and high grade glioma using statistical features in brain MRIs, *Recent Trends in Eng. and Technology*, 4(3), 27-33.

Ruess, L., 2000. Inflammatory bowel diseases in children and young adults: correlation of sonographic and clinical parameters during treatment, *American Journal on Roentgenology*, 175(1), 79–84.

Sujana, H., Swarnamani, S., 1996. Application of Artificial Neural Networks for the classification of liver lesions by texture parameters, *Ultrasound in Med. & Biol.*, 22(9), 1177- 1181.

Van der Maaten, L., Postma, E., 2008. Dimensionality reduction: A comparative review, January 2008, 1-22.

Weka 3, Data Mining Software in Java (2016), http://www.cs.waikato.ac.nz/ml/weka/

Witten, I., Frank, E., 2005. Data Mining. Practical Machine Learning Tools and Techniques (3rd edition), Morgan Kaufmann.

Yoshida, H., Casalino, D., 2003. Wavelet-packet-based texture analysis for differentiation between benign and malignant liver tumors in ultrasound images, *Physics in Med. and Biol.*, 48, 3735-3753.

.