

The Role of the Superior Order GLCM in the Characterization and Recognition of the Liver Tumors from Ultrasound Images

Delia MITREA¹, Sergiu NEDEVSCHI¹, Mihai SOCACIU², Radu BADEA²

¹ Dept. of Computer Science, Technical Univ. of Cluj-Napoca, G. Baritiu Str. No.26-28, 400659 Cluj-Napoca, Romania

² Dept. of Ultrasonography, I. Hatieganu Univ. of Medicine and Pharmacy, V. Babes Str. 8, 400012 Cluj-Napoca, Romania

Delia.Mitrea@cs.utcluj.ro

Abstract. *The hepatocellular carcinoma (HCC) is the most frequent malignant liver tumor. It often has a similar visual aspect with the cirrhotic parenchyma on which it evolves and with the benign liver tumors. The golden standard for HCC diagnosis is the needle biopsy, but this is an invasive, dangerous method. We aim to develop computerized, non-invasive techniques for the automatic diagnosis of HCC, based on information obtained from ultrasound images. The texture is an important property of the internal organ tissues, able to provide subtle information about the pathology. We previously defined the textural model of HCC, consisting in the exhaustive set of the relevant textural features, appropriate for HCC characterization and in the specific values of these features. In this work, we analyze the role that the superior order Grey Level Cooccurrence Matrices (GLCM) and the associated parameters have in the improvement of HCC characterization and automatic diagnosis. We also determine the best spatial relations between the pixels that lead to the highest performances, for the third, fifth and seventh order GLCM. The following classes will be considered: HCC, cirrhotic liver parenchyma on which it evolves and benign liver tumors.*

Keywords

Superior order GLCM, ultrasound images, hepatocellular carcinoma (HCC), benign liver tumors, non-invasive diagnosis.

1. Introduction

The hepatocellular carcinoma is the most frequent malignant liver tumor (75% of liver cancer cases). The human observations are not enough for a reliable diagnosis, and the biopsy is an invasive, dangerous method. Thus, a more subtle analysis is due, and we perform this by using computerized methods applied on ultrasound images. We previously defined the imagistic textural model of HCC [1], consisting in the most relevant textural features, able to separate this tumor from the visually similar tissues, and in

the specific values associated to the relevant features. In this work, we analyze new methods for textural feature computation, based on the superior order GLCM [2], the purpose being that of improving the textural model of HCC, under both aspects of relevant textural features and automatic diagnosis accuracy. We analyzed the role that the superior order GLCM has concerning both the subtle characterization of HCC and of the benign liver tumors, as well as their automatic diagnosis. Feature selection methods such as the Correlation based Feature Selection (CFS) [3] and the evaluation of the attributes through the gain ratio with respect to the class [4] were applied for the detection of the relevant textural features. Powerful classifiers, that gave the best results in our former experiments [1], such as the Decision Trees, the Multilayer Perceptron [5] and the support Vector Machines (SVM) [5], as well as classifier combination schemes [5], were adopted for the recognition accuracy assessment. In this paper I extend the concepts and the experiments presented in [6], where the second, third and fifth order GLCM were compared, using ultrasound images belonging to the classes of HCC and cirrhotic parenchyma on which HCC had evolved. Here, we consider also the seventh order GLCM and we assess the superior order GLCM on a more extended dataset, formed by the classes of HCC, cirrhotic parenchyma on which HCC had evolved and benign liver tumors.

2. The State of the Art

The most frequently used methods in the field of texture-based characterization of the malignant tumors are the Grey Levels Cooccurrence Matrix (GLCM) and the associated Haralick parameters, the Run-Length Matrix parameters [7], fractal-based methods [8], the wavelet [9] and Gabor transforms [10], combined with the k-nn classifiers, Bayesian classifiers [10], Artificial Neural Networks, Fisher Linear Discriminants [7], Support Vector Machines [7]. In [7] the authors computed the first order statistics, the Grey Level Co-occurrence Matrix and the Run-Length Matrix parameters, which were used in combination with Artificial Neural Networks, as well as with Linear Discriminants, for the classification of the liver lesions. The

fractal-based methods were used in [8], for detecting the salivary gland tumors in ultrasound images. The wavelet transform was also implemented for differentiating the malignant and benign liver tumors in ultrasound images [9]. Concerning the implementation of the superior order GLCM, the theoretical aspects were best approached in [2]. The third order GLCM was applied in [11] for the analysis of the trabecular bones within proximal femur radiographs. Here, the third order GLCM matrix was computed, considering the three corresponding pixels collinear, or forming a right angle triangle, being situated at distances of 1, 2, 3 or 4 from each-other. The final textural features were obtained by averaging the values of the third order GLCM features that resulted in each case. Only moderate correlations were found between the third order GLCM features and the bio-mechanical properties of the bones ($R^2=0.38$, $p<0.05$), while the correlation of the third order GLCM features with the bone mineral density parameter was not significant ($p>0.05$). The third order GLCM was also applied in [12] for crop classification within radar images and provided accuracy results of 86%. As we can notice, there are not any significant achievements concerning the implementation of the superior order GLCM for ultrasound image analysis or for the characterization of the malignant tumors. The fifth order and seventh order GLCM are also less implemented. In our research, we applied the third order, fifth order and seventh order GLCM for finding subtle properties of the HCC tissue, based on ultrasound images, aiming to improve the textural model of HCC and the automatic recognition.

3. HCC and the Benign Liver Tumors. Main Characteristics and Visual Aspect in Ultrasound Images

The most relevant oncogenic agent for HCC development is the chronic viral infection with the B virus or C virus of hepatitis, the next evolution phase, preceding HCC, being cirrhosis [13]. HCC evolves from cirrhosis, after a restructuring phase at the end of which dysplastic nodules (future malignant tumors) result. In incipient phase, HCC appears like a small region having a different texture than the other parts of the tissue and a diameter of 1.5 cm to 2 cm. In the case of an evolved HCC (Fig.1), the essential textural attribute is the heterogeneity, due to the co-existence of the necrosis, fibrosis, and of the regions with active growths. HCC is also characterized through a complex vessel structure [13]. Concerning the benign liver tumors, the most frequent types are hemangioma, focal nodular hyperplasia and hepatocellular adenoma. Hemangiomas consist in a mass of abnormal blood vessels.

Up to 5% of the adults in the United States have small hemangiomas in their liver. Hepatocellular adenomas (HA) are benign liver tumors that occur mostly in women. They consist of sheets of hepatocytes without bile ducts or portal areas. The focal nodular hyperplasia (FNH) is another

frequent benign liver tumor. The histological studies detected the existence of some fibrous bands, of multiple biliary ducts, and of some fibrous scars, of stellar shape. In ultrasound images, the FNH tumor appears as an isoechogetic or hypoechogetic structure. The benign liver tumors are visually similar with HCC in most cases [14].

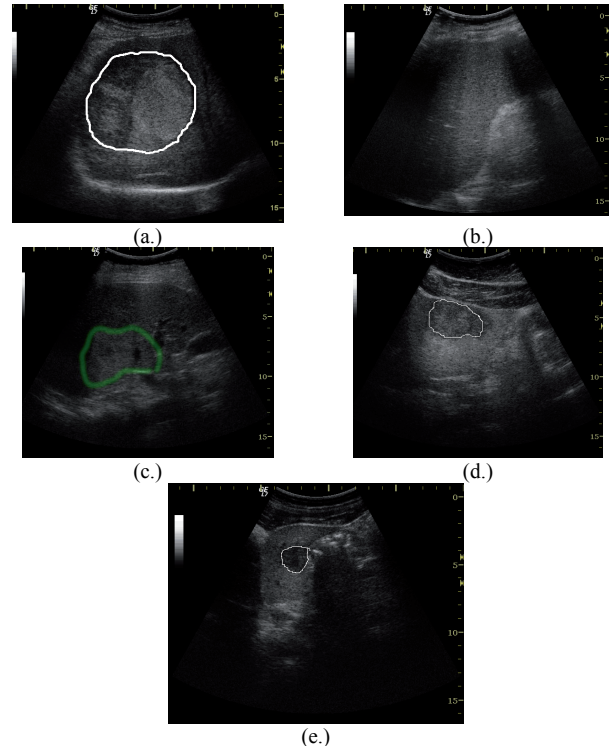


Fig. 1. (a.) HCC in evolved phase; (b) Cirrhotic parenchyma without HCC; (c.) Hemangioma; (d.) Hepatocellular adenoma (HA); (e.) Focal Nodular Hyperplasia (FNH).

4. The Proposed Solution

4.1 The Textural Model of HCC

The textural model of HCC consists in the exhaustive and non-redundant set of relevant textural features, able to distinguish this tumor from the cirrhotic liver parenchyma and from the benign tumors. The specific, statistical values of the textural features (mean, standard deviation and probability distribution) are part of the model. In order to build a reliable textural model, first, the image selection for the training set building is necessary. For each considered type of tissue a corresponding class is built. Then, an image analysis phase is due: the textural feature computation using specific methods for texture analysis is involved in this process. The learning phase is important in order to perform dimensionality reduction through feature extraction and relevant feature selection, to eliminate the redundant features, to determine the specific, statistical values and the corresponding probability distributions. At the end, we perform the validation phase, involving the evaluation of the model by providing the selected features at classifi-

ers inputs and estimating the accuracy of each classifier [1].

4.2 Improvement of the Image Analysis Phase Using Superior Order GLCM

During the image analysis phase, specific methods for texture analysis were applied [1], resulting the following features: the mean value of the grey levels [15], the Grey Levels Cooccurrence Matrix (GLCM) and the associated Haralick parameters [16], edge based statistics, the frequency and density of the textural microstructures, detected by applying the Laws convolution filters [15], the Shannon entropy [17] computed after applying the wavelet transform, at two levels of resolution [18]. In this work, we implemented the superior order GLCM, for obtaining more refined textural features. The Grey Levels Cooccurrence Matrix (GLCM) is defined, in the classical approach [16], as containing, in its elements, the number of pairs of pixels, having two specific values of the grey levels, g_1 and g_2 , being situated at a distance defined by a displacement vector. We extended this definition, in order to describe the n -th order GLCM, in the following manner:

$$C_D(g_1, g_2, g_3, \dots, g_n) = \#\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)\}: \\ f(x_1, y_1) = g_1, f(x_2, y_2) = g_2, \dots, f(x_n, y_n) = g_n, |x_2 - x_1| = |\vec{d}x_1|, \\ |x_3 - x_1| = |\vec{d}x_2|, \dots, |x_n - x_1| = |\vec{d}x_{n-1}|, |y_2 - y_1| = |\vec{d}y_1|, |y_3 - y_1| \\ = |\vec{d}y_2|, \dots, |y_n - y_1| = |\vec{d}y_{n-1}|, \\ \text{sgn}(x_2 - x_1)(y_2 - y_1) = \text{sgn}(\vec{d}x_1 \cdot \vec{d}y_1), \dots \\ \text{sgn}(x_n - x_1)(y_n - y_1) = \text{sgn}(\vec{d}x_{n-1} \cdot \vec{d}y_{n-1})\} \quad (1)$$

In (1), #S is the number of the elements in set S, while

$$\vec{d} = ((\vec{d}x_1, \vec{d}y_1), (\vec{d}x_2, \vec{d}y_2), \dots, (\vec{d}x_{n-1}, \vec{d}y_{n-1})) \quad (2)$$

is the set of the displacement vectors. In practice, the GLCM probability (3) will be used:

$$p(g_1, g_2, \dots, g_n) = \frac{C_D(g_1, g_2, \dots, g_n)}{\sum_{g_1=0}^{N_g-1} \sum_{g_2=0}^{N_g-1} \dots \sum_{g_n=0}^{N_g-1} C_D(g_1, g_2, \dots, g_n)} \quad (3)$$

In (3), N_g is the total number of the grey levels in the image. Based on the n -th order GLCM, we computed the following parameters: energy, entropy, local homogeneity, correlation, contrast, variance [2]. We also computed the maximum probability for an n -tuple of grey levels to appear in the image, as indicated in (4):

$$p_{\max} = \text{Max} (p(g_1, g_2, \dots, g_n), 0 < g_1 < N_g, \dots, 0 < g_n < N_g) \quad (4)$$

• The third order GLCM

In the case of the third order GLCM, we considered specific orientations of the displacement vectors, related to the direction of the ultrasound signal propagation. The corresponding three pixels were either collinear, having the

current pixel, of coordinates (x_1, y_1) , situated in the central position, or they formed a right angle triangle, with the current pixel placed in the position of the right angle. In the case of the collinearity of the pixels, the direction pairs were $(0^\circ, 180^\circ)$, $(90^\circ, 270^\circ)$, $(45^\circ, 225^\circ)$, $(135^\circ, 315^\circ)$, while in the second case, the following direction pairs were considered: $(0^\circ, 90^\circ)$, $(90^\circ, 180^\circ)$, $(180^\circ, 270^\circ)$, $(0^\circ, 270^\circ)$, $(45^\circ, 135^\circ)$, $(135^\circ, 225^\circ)$, $(225^\circ, 315^\circ)$, $(45^\circ, 315^\circ)$. The values of $|\vec{d}x_i|$ and $|\vec{d}y_i|$ were 0 or 2, with $i \in \{1, 2\}$.

• The fifth order GLCM

For the fifth order GLCM, the following groups of directions were taken into account: $(0^\circ, 180^\circ, 90^\circ, 270^\circ)$, respectively $(45^\circ, 225^\circ, 135^\circ, 315^\circ)$. The current pixel was situated in the central position. The values of $|\vec{d}x_i|$ and $|\vec{d}y_i|$ were 0 or 2, with $i \in \{1, 2, 3, 4\}$.

• The seventh order GLCM

In the case of the seventh order GLCM, combinations of the horizontal and vertical directions with each of the diagonal directions were considered. The following groups of directions were obtained: $(0^\circ, 180^\circ, 90^\circ, 270^\circ, 45^\circ, 225^\circ)$, $(0^\circ, 180^\circ, 90^\circ, 270^\circ, 135^\circ, 315^\circ)$, $(0^\circ, 180^\circ, 90^\circ, 270^\circ, 45^\circ, 135^\circ)$, $(0^\circ, 180^\circ, 90^\circ, 270^\circ, 225^\circ, 315^\circ)$, $(0^\circ, 180^\circ, 90^\circ, 270^\circ, 45^\circ, 315^\circ)$, $(0^\circ, 180^\circ, 90^\circ, 270^\circ, 225^\circ, 135^\circ)$. Also, combinations of the diagonal directions with each of the horizontal and vertical directions were taken into account, resulting the following groups: $(45^\circ, 225^\circ, 135^\circ, 315^\circ, 0^\circ, 180^\circ)$, $(45^\circ, 225^\circ, 135^\circ, 315^\circ, 90^\circ, 270^\circ)$, $(45^\circ, 225^\circ, 135^\circ, 315^\circ, 0^\circ, 90^\circ)$, $(45^\circ, 225^\circ, 135^\circ, 315^\circ, 0^\circ, 270^\circ)$, $(45^\circ, 225^\circ, 135^\circ, 315^\circ, 90^\circ, 180^\circ)$, $(45^\circ, 225^\circ, 135^\circ, 315^\circ, 180^\circ, 270^\circ)$. The values of $|\vec{d}x_i|$ and $|\vec{d}y_i|$ were 0 or 2, with $i \in \{1, 2, 3, 4, 5, 6\}$.

4.3 The Learning Phase

During the learning phase, we performed the selection of the relevant textural features by using the method of Correlation based Feature Selection (CFS), in combination with Genetic Search [3], for retaining those textural features that were mostly correlated with the class parameter and less correlated with the other textural features. For each group of features, a merit was computed. Another feature selection method, that computed, for each textural feature, the gain ratio with respect to the class, and ranked the features accordingly [3], was also implemented.

4.4 The Validation Phase

The validation phase consisted in providing the textural features at the inputs of powerful classifiers and in analyzing the effect of the new textural features on the classification accuracy improvement. We adopted the classifiers of Multilayer Perceptron (MLP) [5], of Support Vector Machines (SVM) [5], and of Decision Trees [5], as they led to the best results in our former experiments [1].

The AdaBoost, bagging and stacking combination schemes [5] were also implemented. The following parameters were used in order to assess the performance: the recognition rate, the sensitivity (TP rate), the specificity (TN rate), the area under the Receiver Operating Characteristic (ROC) curve (AUC) [5], and the time for model building [4].

5. Experiments and Discussions

The experimental dataset contained the classes of HCC, cirrhotic parenchyma on which HCC had evolved, and benign liver tumors. This dataset was formed from 300 cases of HCC and 150 cases of benign tumors of the following type: hemangioma, hepatocellular adenoma (HA), focal nodular hyperplasia (FNH). For each patient, three images were considered. The ultrasound images were acquired using various orientations of the transducer, under the same settings of the Logiq 7 ultrasound machine: the frequency was 5.5 MHz, the gain was 78 dB and the depth was 16 cm. Each class was formed by regions of interest selected on the desired type of tissue. The classes were combined in equal proportions. An instance of the dataset consisted of the values of the textural features considered for the experiment. The training set was formed by 80% of the data instances, while the test set consisted of 20% of the data instances. These subsets were provided by the cross-validation method. During *the image analysis phase*, the textural features were computed after applying an averaging filter for noise reduction, on rectangular regions of interest, of 50x50 pixels in size. These features were independent of orientation, of illumination, and were scaled with the size of the region of interest. During the *learning phase*, the method of Correlation based Feature Selection (CFS) in combination with genetic search, and the Gain Ratio Attribute Evaluation with the Ranker method belonging to the Weka library [4], were applied for the selection of the relevant features. The validation phase consisted in the evaluation of the new textural features by using the following classifiers of the Weka library [4]: the Support Vector Machines, with a polynomial kernel of third degree, the Multilayer Peceptron (MLP), with the learning rate of 0.2, the momentum of 0.8, the number of nodes in the single hidden layer being the arithmetic mean between the number of classes and the number of input features, these parameters being chosen with the purpose of achieving both speed and accuracy, and in order to avoid the over-training. The J48 method [4] that implements the C4.5 algorithm, and also the Random Forest method [4] based on Decision Trees were experimented, as well. The AdaBoost combination scheme [4], with 10 iterations, was also implemented. The bagging method with 100% bag size percent and the StackingC combination scheme using the linear regression as a meta-classifier [4] were also tested for classification accuracy improvement. The stratified cross-validation with 5 folds was used for the assessment of the classification performance.

• Discriminating HCC from the cirrhotic liver parenchyma on which it evolves

In the case of the third order GLCM, the selection of the relevant textural features was performed first, using the methods of Correlation-based Feature Selection (CFS), respectively of Gain Ratio Attribute Evaluation, after applying the third order GLCM for each considered pair of directions separately. The set of relevant textural features showed the prevalence of the $(0^\circ, 90^\circ)$, $(0^\circ, 180^\circ)$ and $(0^\circ, 270^\circ)$ combinations. The best classification results were obtained when considering only the $(0^\circ, 270^\circ)$ combination of directions. When combining the third order GLCM parameters obtained for the $(0^\circ, 270^\circ)$ directions with the other old textural features, the accuracy was better, in the case of the SVM classifier, than in the case of combining the second order GLCM features with the other old textural features. When combining the considered third order GLCM parameters with all the other textural features (including the second order GLCM features) we obtained the best classification performance.

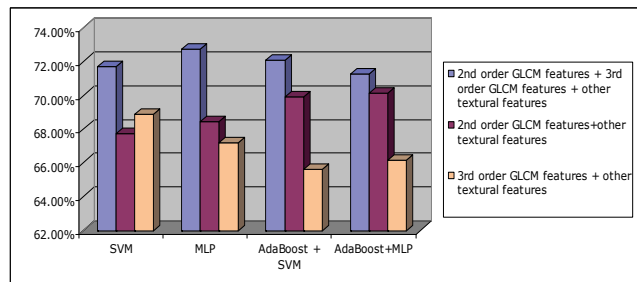


Fig.2. The recognition rate obtained when considering the previous textural features, and various combinations of the old textural features with the third order GLCM features.

Fig. 2 illustrates the results concerning the improvement in the recognition rate. The maximum recognition rate, of 72.73%, was obtained for the Multilayer Peceptron.

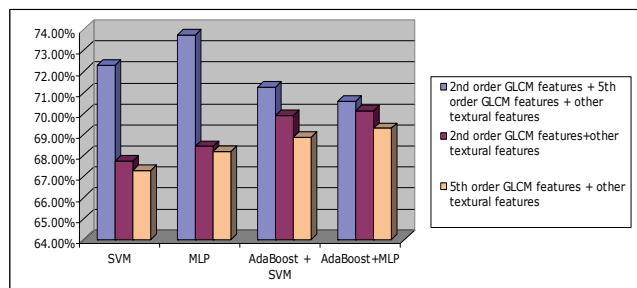


Fig.3. The recognition rate obtained when considering the previous textural features, and combinations of the old textural features with the fifth order GLCM features.

The classification accuracy results obtained when using only the fifth order GLCM features were also compared by considering various groups of the displacement vector directions. The best performance was achieved for

the (0°, 180°, 90°, 270°) directions. The classification accuracy that resulted after combining the fifth order GLCM parameters with the other textural features is illustrated in Fig. 3.

The maximum recognition rate, of 73.75%, was obtained in the case of the Multilayer Perceptron (MLP) classifier, and corresponded to the combination between the second order GLCM features, the fifth order GLCM features, computed for the (0°, 180°, 90°, 270°) group of directions, and the other textural features. In the case of the seventh order GLCM, the best combination of directions, that led to the highest recognition rate, was that performed between the vertical and two diagonal directions of the displacement vectors: (45°, 225°, 135°, 315°, 90°, 270°).

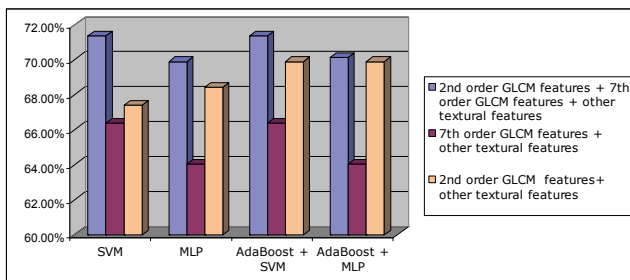


Fig. 4. The recognition rate obtained when considering the old textural features, and combinations of the previous features with the seventh order GLCM features.

After the combination of the seventh order GLCM features with the old textural features, the highest recognition rate, of 71.38%, was obtained for the SVM classifier, respectively for the AdaBoost combination scheme that used the SVM method as basic classifier. The comparison of the recognition rates that corresponded to various groups of textural features is illustrated in Fig. 4. We can notice that the group including the second order GLCM features, the seventh order GLCM features and the old textural features overpassed the accuracy that corresponded to the other groups of features. Finally, the second order, third order, fifth order, seventh order GLCM features and the other textural features were combined in a single group. In the set of relevant textural features, derived after applying the feature selection methods (Correlation based Feature Selection and Gain Ratio Attribute Evaluation) we remarked the presence of the third order GLCM features (correlation, homogeneity and contrast), of the fifth order GLCM features (entropy, correlation and variance), respectively of the seventh order GLCM features (entropy and variance). The GLCM correlation denoted differences in granularity between the malignant tumor of HCC and the cirrhotic liver parenchyma, the homogeneity and entropy emphasized the heterogeneous, chaotic structure of the tumor tissue, while the variance and contrast denoted the more increased variety of the grey levels in the case of HCC. Textural features such as the edge orientation variability, the mean of the waves microstructures obtained by applying the Laws convolution filters, the entropy computed after applying the wavelet transform, also resulted as being relevant, emphasizing the complex structure of HCC

and its chaotic character.

| | Recogn. Rate | TP Rate | TN Rate | AuC | Time |
|-------------------------------------|--------------|---------|---------|-------|--------|
| Random Forest | 77.26% | 82.8% | 71.7% | 85% | 1.44s |
| SVM | 77.48% | 75.8% | 79.2% | 77.5% | 17.23s |
| Bagging + Random Forest | 81.78% | 83% | 80.5% | 90.8% | 9.78s |
| Bagging + SVM | 77.3% | 75.7% | 79.4% | 77.6% | 17.23s |
| AdaBoost + Random Forest | 78.50% | 81% | 76% | 83.8% | 5.56s |
| AdaBoost + SVM | 77.5% | 75.9% | 79.3% | 77.5% | 17.23s |
| Stacking: MLP + SVM + Random Forest | 79.185% | 79.6% | 78.7% | 87.4% | 47.33s |

Tab.1. The classification performance obtained for the final set of the relevant textural features.

The highest values of the recognition accuracy parameters, obtained for the final group of relevant textural features, which resulted after performing the union between the two groups of features that were obtained by applying the feature selection methods separately, are depicted in Tab. 1. The best recognition rate, of 81.78%, was obtained in the case of the bagging combination scheme, which used the Random Forest method as basic classifier.

• Discriminating HCC from the benign liver tumors

In the case of the third order GLCM, the most efficient pair of directions was (0°, 90°). The increase of the recognition rate, obtained after combining the old textural features with the third order GLCM features, corresponding to the case of differentiation between HCC and the benign liver tumors, is illustrated in Fig. 5. Thus, the group formed by the third order GLCM features and by the old textural features, as well as the group formed by the second order textural features, by the third order textural features and by the old textural features, led to an increase in the recognition rate, compared with the case when only the second order GLCM features and the old textural features were used.

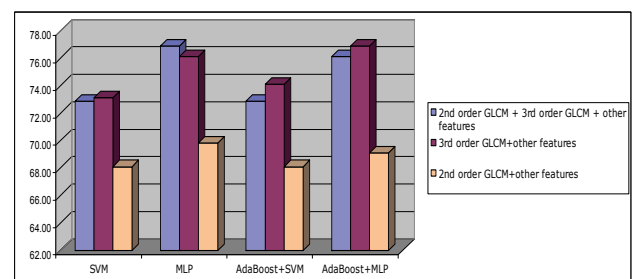


Fig. 5. The recognition rate obtained when considering the previous textural features, and combinations of the old textural features with the third order GLCM features.

The best recognition rate, of 76.88%, resulted in the case of the combination between the third order GLCM features and the other textural features, when applying the AdaBoost combination scheme with the MLP as basic classifier. In the case of the fifth order GLCM matrix, applied for the differentiation between HCC and the benign liver tumors, the combination of the features corresponding to both direction groups: (0°, 180°, 90°, 270°), respectively

(45°, 225°, 135°, 315°), provided the best accuracy results. Fig. 6 illustrates the comparison of the recognition rates that correspond to various groups of features, involving or not the fifth order GLCM features. The groups formed by the fifth order GLCM features and the old textural features, respectively by the second order GLCM features, by the fifth order GLCM features and by the old textural features, overpassed in recognition accuracy the group formed by the second order GLCM features and by the other textural features. The best results were obtained using both the third order GLCM features, the fifth order GLCM features and the old textural features, the highest recognition rate, of 75.90%, corresponding to the MLP classifier, and also to the AdaBoost scheme that used the MLP method as basic learner.

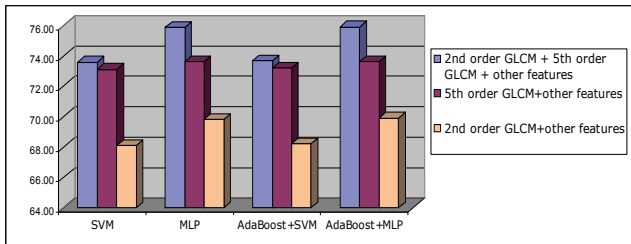


Fig. 6. The recognition rate obtained when considering the previous textural features and combinations of the old textural features with the fifth order GLCM features.

When experimenting the seventh order GLCM matrix for various groups of the displacement vector directions, the combination (45°, 225°, 135°, 315°, 0°, 270°) provided the best results. Fig. 7 shows the comparison between the recognition rates corresponding to various feature groups. Both groups, the first being formed by the seventh order GLCM features and the other old textural features, and the second being constituted by the second order GLCM features, the seventh order GLCM features and the other old textural features, led to an accuracy improvement, compared with the old features.

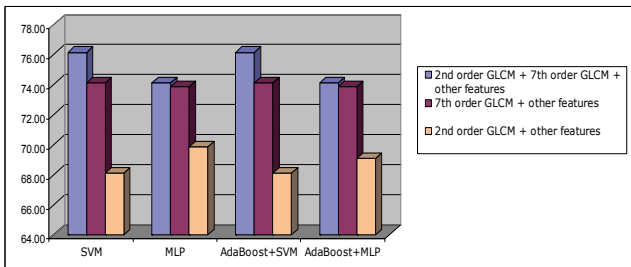


Fig. 7. The recognition rate obtained considering the previous textural features, and combinations of the old textural features with the seventh order GLCM features.

The highest recognition rate, of 76.130%, was obtained in the case of the SVM classifier, when using both the second and seventh order textural features, in combination with the other old textural features. Then, the second order, third order, fifth order, seventh order GLCM features and the other old textural features were combined, the feature selection methods were applied and the classifica-

tion performance was assessed by providing the union of the relevant feature sets at the classifier inputs. The results are illustrated in Tab. 2. The highest accuracy values were achieved for the stacking scheme, which used the MLP, SVM and Random Forest methods as basic learners, the recognition rate being 87.07%, and the AuC parameter being above 93%.

| | Recogn. Rate | TP Rate | TN Rate | AuC | Time |
|-------------------------------------|--------------|---------|---------|-------|--------|
| Random Forest | 81.97% | 89.1% | 74.8% | 90.5% | 0.77s |
| SVM | 83.67% | 80.3% | 87.1% | 83.7% | 0.22s |
| MLP | 82.99% | 80.3% | 85.7% | 91.1% | 46.8s |
| Bagging + Random Forest | 85.37% | 83.7% | 87.1% | 92.3% | 3.72s |
| Bagging + SVM | 82.99% | 80.3% | 85.7% | 88.9% | 2.5s |
| AdaBoost + Random Forest | 84.35% | 87.1% | 81.6% | 90.9% | 0.33s |
| AdaBoost + SVM | 80.61% | 78.9% | 82.3% | 88.6% | 2.81s |
| Stacking: MLP + SVM + Random Forest | 87.07% | 87.1% | 87.1% | 93.2% | 98.84s |

Tab. 2. The classification performance obtained for the final set of the relevant textural features.

The final set of the relevant textural features contained, in the case of differentiating between HCC and the benign liver tumors, the second order GLCM features, the third order GLCM features, the fifth order GLCM features, as well as the seventh order GLCM features. The GLCM homogeneity, energy and entropy emphasized the inhomogeneous, irregular, chaotic structure of the malignant tumor tissue, while the GLCM correlation pointed out a difference in granularity between the malignant and the benign tumors. The entropy computed at the second level, after applying the wavelet transform, was also important. The mean and the frequency of the textural microstructures, obtained after applying the Laws filters, resulted to be relevant, as well.

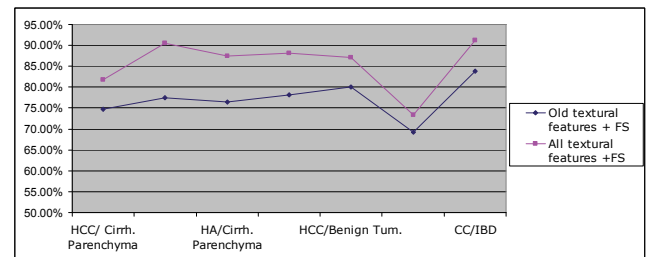


Fig. 8. The comparison of the recognition rates obtained by using the former set and the new set of features.

A global result, illustrating the role of the new textural features in the improvement of the classification accuracy, is depicted in Fig. 8. Here, the comparison between the recognition rates obtained for the old textural features, after feature selection, respectively the recognition rates achieved for the entire set of the new and old relevant textural features, was performed using the following pairs of classes: HCC/cirrhotic parenchyma on which HCC had evolved; Hepatocellular adenoma (HA)/ cirrhotic parenchyma; Hemangioma/cirrhotic parenchyma; HNF/cirrhotic parenchyma; HCC/benign liver tumors. The results obtained in the context of Inflammatory Bowel Diseases

differentiation (considering the Crohn's disease and the ulcero-hemorrhagic recto-colitis), respectively in the case of discriminating between IBD and colo-rectal cancer (CC) [19], were also represented. Concerning the values of the maximum probability parameter, computed for the second, third, fifth and seventh order GLCM, for each considered pair of classes, we obtained the following results: in the case of the second order GLCM, the average value of the maximum probability was 0.05, for the third order GLCM this value was 0.01, for the fifth order GLCM the mean value of the maximum probability was 0.0025, while in the case of the seventh order GLCM, this value was 0.0005. As we can notice, the value of the maximum probability decreased, while the order n of the matrix increased. Also, the maximum probability had, in the case of HCC, lower values than in the cases of the cirrhotic parenchyma and of the benign liver tumors, due to the irregular, chaotic character of the malignant tumors. For the HCC tumor, groups of pixels having hypoechoogenic values appeared frequently. In the case of the fifth order GLCM, this group was (57, 57, 57, 57, 57) and corresponded to areas of pure tumor tissue (Fig. 9).

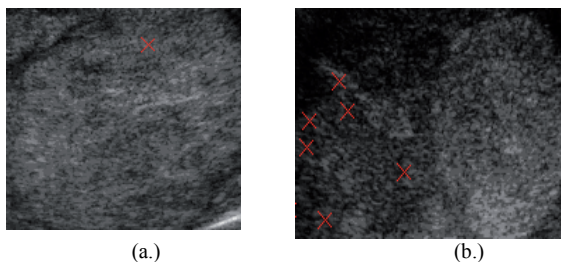


Fig. 9. Groups of five hypoechoogenic values putting into evidence tumor regions with active growth: (a.) inside isoechoic HCC; (b.) inside hyperechoic HCC.

6. Conclusions

The textural features proved to be appropriate for the characterization and recognition of HCC. The superior order GLCM features led to a recognition rate situated above 81%. The groups of specific grey level values met the most frequently inside the tumors corresponded to the structural components of this kind of tissue. The best directions of the displacement vectors, which led to the highest recognition accuracy, corresponded to the axes of the structural components, being also parallel or perpendicular on the directions of the ultrasound signal propagation. In our future research, we aim to confirm the efficiency of the superior order GLCM matrix using real-world texture classes, as well. The extension of the multi-resolution methods and their combination with the superior order GLCM are expected to further increase the recognition performances.

References

[1] MITREA, D., NEDEVSCHI, S., LUPSOR, M., SOCACIU, M., BADEA, R. Advanced classification methods for improving the

- automatic diagnosis of the hepatocellular carcinoma, based on ultrasound images. In *Proceedings of the 2010 IEEE International Conference on Automation, Quality and Testing, Robotics*. Cluj-Napoca (Romania), May 28-30, 2010, p. 265-271.
- [2] TALLA TANKAM, N., AKONO, A., TONYE, E. A generalized new algorithm for the evaluation of textural parameters of order n . In *Proc. of the Fringe Workshop*. Frascati (Italy), 2005, p. 71-77.
- [3] HALL, M. Benchmarking attribute selection techniques for discrete class data mining. *IEEE Transactions on Knowledge and Data Engineering*, 2003, vol. 15, no.3, p. 1-16.
- [4] Weka 3, Data Mining Software in Java. [Online] 2011. Available at: <http://www.cs.waikato.ac.nz/ml/weka/>
- [5] DUDA, R. *Pattern Classification*. Wiley Interscience, 2003.
- [6] MITREA, D., NEDEVSCHI, S., BADEA, R. The role of the superior order GLCM in improving the characterization and recognition of the hepatocellular carcinoma, based on ultrasound images. In *Proceedings of the 3rd International Conference on Telecommunications and Signal Processing (TSP)*. Budapest (Hungary), 2011, p. 602-607.
- [7] SUJANA, H., SWARNAMANI, S. Application of Artificial Neural Networks for the classification of liver lesions by texture parameters. *Ultrasound in Medicine & Biology*, 1996, vol. 22, no. 9, p. 1177-1181.
- [8] CHIKUI, T., TOKUMORI, K., YOSHIURA, K. Sonographic texture characterization of salivary gland tumors by fractal analysis. *Ultrasound in Medicine & Biology*, 2005, vol. 31, no.10, p. 1297-1304.
- [9] YOSHIDA, H., CASALINO, D. Wavelet packet based texture analysis for differentiation between benign and malignant liver tumours in ultrasound images. *Physics in Medicine & Biology*, 2003, vol. 48, p. 3735-3753.
- [10] MADABHUSHI, A., FELMAN, M. Automated detection of prostatic adenocarcinoma from high-resolution ex vivo MRI. *IEEE Trans. on Medical Imaging*, 2005, p.1611-1626.
- [11] HUBER, M. B., CARBALLIDO-GAMIO, J. Development and testing of texture discriminators for the analysis of trabecular bone in proximal femur radiographs. *Medical Physics*, 2009, vol. 36, no. 11, p. 5089 – 5098.
- [12] ANYS, H., HE, D. Evaluation of textural and multipolarization radar features for crop classification. *IEEE Trans. on Geosciences and Remote Sensing*, 1995, vol. 33, no. 5, p. 1170-1181.
- [13] BRUIX, J. Hepatitis B virus and hepatocellular carcinoma. *Journal of Hepatology*, 2003, vol. 39, no.1, p. S59- S63.
- [14] American liver foundation. [Online]. Available at: <http://www.liverfoundation.org/education/info/benign tumors/>
- [15] JAIN, A. K. *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.
- [16] CLAUSI, D. An analysis of cooccurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing*, 2002, vol. 28, no. 1, p. 45-62.
- [17] STOLLNITZ, E., DEROSE, T. Wavelets for computer graphics. *IEEE Computer Graphics & Applications*, 2005, vol. 15, no. 3, p. 76-84.
- [18] PARKER, J. R. *Algorithms of Image Processing and Computer Vision*. Wiley Computer Publishing, 1996.
- [19] MITREA, D., NEDEVSCHI, S., BADEA, R., SOCACIU, M., GOLEA, A. Computerized characterization and diagnosis of the inflammatory bowel diseases and of the colon cancer based on ultrasound images, using superior order cooccurrence matrices. *Automation, Computers, Applied Mathematics (ACAM) Scientific Journal*, December, 2011, vol. 20, no. 2.