

Texture analysis of masses in digitized mammograms using Gleason and Menhinick Diversity Indexes

Simara Vieira da Rocha*, Geraldo Braz Junior, Aristófanés Corrêa Silva, Anselmo Cardoso de Paiva

Abstract Introduction: Breast cancer is the second most common type of cancer in the world, being more common among women and representing 22% of all new cancer cases every year. The sooner it is diagnosed, the better the chances of a successful treatment are. Mammography is one way to detect non-palpable tumors that cause breast cancer. However, it is known that the sensitivity of this exam can vary considerably due to factors such as the specialist's experience, the patient's age and the quality of the images obtained in the exam. The use of computational techniques involving artificial intelligence and image processing has contributed more and more to support the specialists in obtaining a more precise diagnosis. **Methods:** This paper proposes a methodology that exclusively uses texture analysis to describe features of masses in digitized mammograms. To increase the efficiency of texture feature extraction, the diversity index's capability to detect patterns of species co-occurrence is used. For this purpose, the Gleason and Menhinick indexes are used. Finally, the extracted texture is classified using the Support Vector Machine, looking to differentiate the malignant masses from the benign. **Results:** The best result was obtained using the Gleason index, with 86.66% accuracy, 90% sensitivity, 83.33% specificity and an area under the ROC Curve (Az) of 0.86. **Conclusion:** Both indexes showed statistically similar performance; however, the Gleason index was slightly superior.

Keywords Breast cancer, Medical images, Gleason Diversity Index, Menhinick Diversity Index, Computer-aided diagnosis.

Introduction

Breast cancer is the second most common type of cancer in the world, being more common among women and representing 22% of all new cancer cases every year. The occurrence of this type of cancer has grown 3.1% per year (American..., 2011). Mammography is one way of detecting non-palpable tumors that cause breast cancer. On average, it detects 80% to 90% of breast cancers in asymptomatic women (American..., 2011). Since this exam began being used as a routine check-up, a reduction in the mortality rate related to this pathology has been observed. This can be largely explained by a mammogram's ability to detect the cancer in its initial stage, when treatment can be more efficient and a cure is more likely. However, it is known that the sensitivity of this exam can vary considerably due to factors such as the specialist's experience, the patient's age and the quality of the images obtained in the exam.

The use of computational techniques of artificial intelligence and image processing has contributed more and more to support specialists in obtaining more precise diagnoses. In this way, since the last decade, the interest in systems for detection (Computer-Aided Detection – CAD) and diagnosis (Computer-Aided Diagnosis – CADx) has been growing, in

order to support the radiologists in interpreting the mammograms (Kinoshita et al., 2004).

One of the visibly detectable anomalies in mammographic images is the cell masses, or clusters of cells that create a denser band than the surrounding tissue. Masses can be caused by both benign and malignant conditions. Therefore, characterizing mass features such as size, shape and margin disposition is fundamental in establishing the probability of malignancy (Kopans, 2000).

Because texture is an attribute that is hard for the human analyst to interpret, it is common to use features of the outlines of the masses to diagnose these regions. However, such features are not always clear in these exams. For example, there can be lesions without a well-defined outline that can overlap regions of masses and calcifications, preventing their proper visualization. This difficulty contributes to an increase in the number of biopsies with negative results. Therefore, the development of techniques of texture feature extraction can help specialists produce more precise diagnoses. Many studies have been developed to perform the texture analysis, aiming to differentiate suspicious regions in mammography exams in order to establish their benign or malignant

*e-mail: simara@deinf.ufma.br

behavior. In this paper, several of these systems will be described along with their respective techniques.

One strategy largely used to extract texture features is the use of gray levels co-occurrence matrixes (GLCM), whose features can be described using the measures of Haralick et al. (1973). Various works adopt this technique, including Lim and Er (2004) and Rangayyan et al. (2010). In Mavroforakis et al. (2006), gray level run length matrixes (GLRLM) were used along with the GLCM matrixes as a texture descriptor. In the methods proposed in this work, we use the GLCM matrix, the GLRLM matrixes and the gray level gap length matrixes (GLGLM), though we use it only as a way of representing the region of interest (ROI), as will be explained in the methodology.

The feature extraction stage can be performed both by texture analysis and by mass geometry. However, given the difficulty of differentiation between benign and malignant patterns, many studies usually combine texture features and geometry to perform this task, such as Sahiner et al. (2001), Shi et al. (2007), Suganthi and Madheswaran (2010), Varela et al. (2006), Mu et al. (2008) and Liu et al. (2010, 2011).

In Table 1, a summary of related works is presented, containing the technique used to extract texture, the classifier used, the adopted image base, the number of test samples (with M = malignant and B = benign) and the results in percentages.

Here, we see a necessity for developing a technique with exclusive usage of texture analysis for feature extraction to analyze mammographic images with ill-defined mass outlines to determine malignancy.

The goal of this paper is to propose a methodology to differentiate between malignancy and benignancy using the texture of masses in digitized mammograms using Gleason and Menhinick diversity indexes over regions of interest, represented in the form of gray level co-occurrence matrixes (GLCM), gray level run length matrixes (GLRLM) and gray level gap length matrixes (GLGLM). Furthermore, the Support Vector Machine is used to discriminate whether the features produced by the masses should be in the malignant or benign classes.

Methods

The steps of the proposed methodology for differentiation between the benign and malignant classes in digitized mammography masses, using texture characterization through the diversity indexes of Gleason and Menhinick, are presented in Figure 1. They are Image Acquisition, Pre-processing, Image Representation, Feature Extraction and Pattern Recognition.

Image acquisition

The first stage of the methodology was dedicated to obtaining the mammographic images that were used in the tests. For this purpose, the Digital Database for Screening Mammography (DDSM) public database of digitized mammograms, available on the Internet, (Heath et al., 2001) was used. Because the focus of this research was to characterize the mass textures through diversity indexes and determine their malignant nature, we did not use the complete mammogram image. For the sample selection, we adopted the same approach used by Braz et al. (2009). With this

Table 1. Summary of related work.

Work	Technique Texture	Classifier	Base	ROIs (M/B)	Accuracy (%)	Az
Lim and Er (2004)	GLCM	Neural Network	DDSM	343(163/180)	70	-
Rangayyan et al. (2010)	GLCM	FLD	Own	111(46/65)	-	0.75
Mavroforakis et al. (2006)	GLCM, GLRLM	SVM, ANN	DDSM	130(84/46)	83.9	-
Liu et al. (2010)	GLCM, Geometry	SVM	DDSM	309(167/142)	65	0.7
Liu et al. (2010)	GLCM, Geometry	SVM, LDA	DDSM	309(167/142)	76	
Mu et al. (2008)	GLCM, Geometry	SVM	Own	111(46/65)	-	0.93
Varela et al. (2006)	GLCM, Geometry, Histogram	Neural Network Backpropagation	DDSM	1076(590/486)	-	0.81
Sahiner et al. (2001)	Histogram, Geometry	Leave-one-out	Own	249(127/129)	-	0.87
Shi et al. (2007)	Histogram, Geometry	LDA	Own	909(451/458)	-	0.83
Suganthi and Madheswaran (2010)	GLRLM, Geometry, Histogram	Neural Network Backpropagation	DDSM	350(175/175)	99.5	0.95

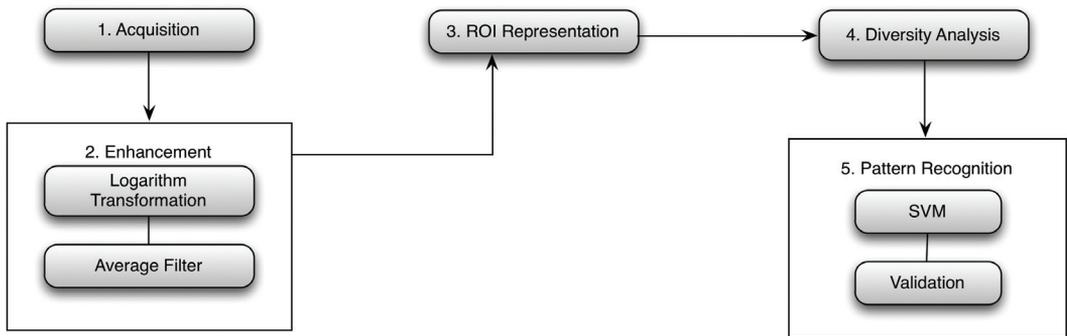


Figure 1. Steps of the proposed methodology.

approach, from the markings performed by specialists in mammography, we extracted the bounding box (minimum area-enclosing rectangle) containing only regions that have masses, yielding a total of 3559 ROIs.

For the tests performed in this project, a subset of 300 ROIs was used, all chosen at random, totaling 160 malignant masses and 140 benign masses. Although this number of samples does not represent the total set of samples in the DDSM database, it does not compromise the generalization capacity of the classifier because the selection is performed at random and there is not a significant imbalance between the two groups. In addition, it allows comparison with other works that use the same database and sample size.

Pre-processing

The goal of this stage is to improve the contrast of the object of interest in relation to the background that may exist in the ROIs and consequently provide a better description of their texture. In this way, we use the logarithmic transformation and average filter (Gonzalez and Woods, 2002).

The logarithmic transformation is defined by the equation:

$$g_i(x, y) = G \lg_{10}(g(x, y) + 1) \quad (1)$$

where $g_i(x, y)$ is the new value of gray level at the point (x, y) , $g(x, y)$ is the original gray level value, and G is the factor defined by the maximum and minimum limits of the image. This factor aims to guarantee that the new values are between zero and the maximum gray level allowed for the image representation. In this work, G will be equal to 255 because the ROIs have 8 bits per pixel.

The logarithmic transformation is used to give more relevance to dark gray levels, which are rarer in mass images (ROIs) of mammograms (lower frequency of occurrence). This way, when we apply

the logarithmic transformation in an ROI, the darker gray levels are grouped to increase their quantitative importance. This process is illustrated in Figure 2.

In applying the logarithmic transformation, the noise in the ROI is amplified (Figure 2b); therefore, we use an average filter with a window size of 5×5 to soften the noise, smoothing the ROI.

Image representation

This stage was introduced in the proposed methodology due to the necessity of adopting the concept of ecological diversity as a technique to perform the feature extraction. Thus, based on the image enhancement performed in the previous stage, the image was represented using first-order statistics in order to calculate the diversity indexes based on the ROI's histogram (Gonzalez and Woods, 2002); second-order statistics through the GLCM matrix (Haralick et al., 1973); and superior-order statistics using the GLRLM matrixes (Galloway, 1975) and GLGLM (Xinli et al., 1994).

Feature extraction

This stage aims to produce descriptive measures for the images, which will form the feature vectors that will be used in the classifying stage. In this work, texture analysis was performed using the statistical approach in Gonzalez and Woods (2002), through adapting the concept of the Ecological Diversity index.

Ecological diversity indexes

In ecology, the term 'diversity' is used to refer to the variety of species present in a community, habitat or region. A community is defined as a set of species that occur in a specific time and place (Magurran, 2004). In this way, the usage of indexes, even though it does not represent the total composition of a community, makes it possible to measure the richness, equality and diversity of species in the different studied

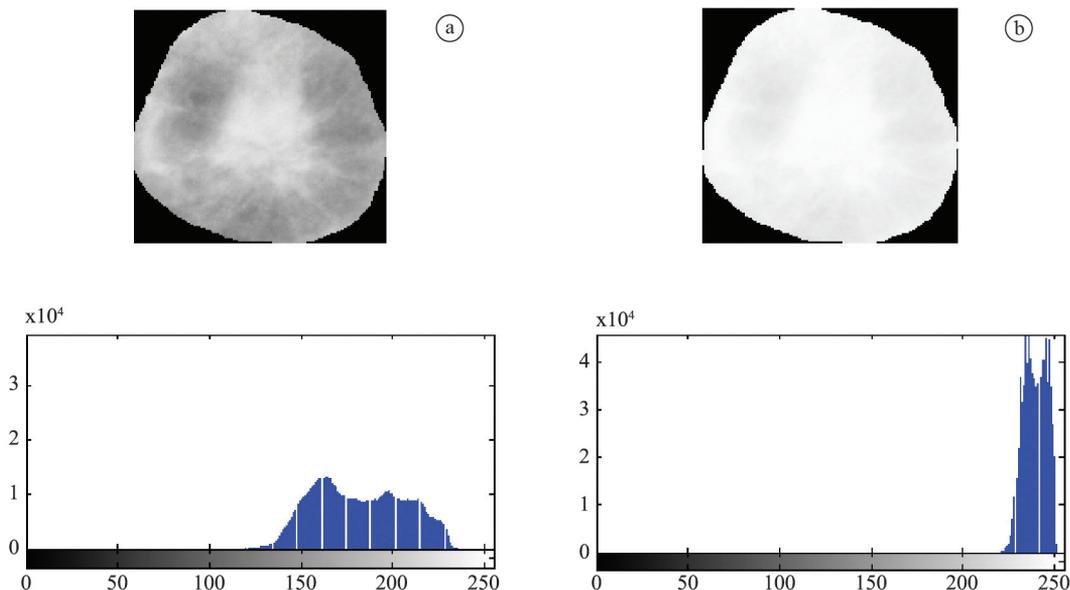


Figure 2. Logarithmic transformation. (a) Original image with its histogram; (b) Enhanced image with its histogram.

environments. This tool is useful to monitor and predict environmental changes.

The concept of diversity involves two parameters: richness, which represents the number of species; and relative abundance, which represents the number of individuals of a particular species that occur in a place or sample. This way, communities with the same richness can differ in diversity depending on the distribution of individuals among species (Mcintosh, 1967). In this paper, we used Gleason and Menhinick's Indexes.

The Gleason Diversity Index considers only the number of species (s) and the logarithm (base 10 or natural) of the total number of individuals (Brower et al., 1997). This index is defined by the equation:

$$D_g = \frac{s}{\text{Log}N} \quad (2)$$

where s is the number of sampled species and N is the total number of individuals in all species.

Another diversity index used in this paper is Menhinick (1964), which considers only the number of species (s) and the square root of the total number of individuals and is calculated by the equation:

$$D_b = \frac{s}{\sqrt{N}} \quad (3)$$

where s is the number of sampled species, and N is the total number of individuals in all species.

Adaptation of the concept of ecological diversity

In order to adapt the concept of ecological diversity for our purposes, two abstractions are used. The first considers that a community will be formed by the pixels of the ROI (each pixel is an individual, and its value defines the species). The second considers that the community is defined by the internal elements of the co-occurrence matrixes of calculated species, and the species will be formed by the values of those elements. This way, we can investigate whether, within the ROIs, there is a dominance of some gray levels over others. Independent of the index, the procedure to extract the features remains the same.

First, the highlighted samples were quantized from 256 to 128, 64, 32, 16 and 8 gray levels, considering each ROI as a community of 256, 128, 64, 32, 16 and 8 species. Along with the quantization, we looked to reproduce the ROIs' representations in different scales of gray levels in such a way to make the description of the texture in those scales possible. After that, these ROIs had their diversity indexes calculated.

Through the histogram of the ROI, we registered the frequency of each gray level (species). This made it possible to extract the richness of species (d) through the quantity of non-null entries (bins) of the histogram, as well as the relative abundance of each species through the value of each bin. The produced feature vector presents 6 variations because we calculated the diversity value for each quantization.

The idea of using the GLCM matrix (Haralick et al., 1973) as a way of representing the ROI was to verify the diversity of the dominance of some pairs of gray levels over others. Two approaches are used to represent the community of gray level pairs. In the first, the individuals correspond to the occurrence of a pair of pixels (i, j) with the same value (gray level) separated by a distance, d , and positioned in a direction θ . The population of each species is represented in the main diagonal of the GLCM matrix (Figure 3b). In the second, we consider the occurrence of members of the community as pairs of pixels (i, j) with different values. This way, the elements outside the main diagonal of the GLCM matrix represent the population of individuals (Figure 3c).

For the θ direction, we adopted the values 0° , 45° , 90° and 135° . For the distance d , the values were 1, 2, 3, 4 and 5. The feature vector generated presented 120 texture attributes (5 distances \times 4 directions \times 6 quantizations), because a GLCM is needed for each θ and d , and 6 quantizations were considered.

The goal of using the GLRLM matrix (Galloway, 1975) is to analyze whether there is a predominance of relatively long run lengths in relation to the short runs or vice-versa. Therefore, the community was formed by the occurrence of consecutive and collinear sequences of n pixels of the same value and a direction θ . This scheme is presented in Figure 4.

The usage of diversity indexes with the GLGLM matrix (Xinli et al., 1994) seeks to investigate whether a mass presents, in a general way, a more homogeneous texture than other mass. It is possible that it contains a higher concentration of homogeneous neighbors, suggesting low diversity. Otherwise, if it possesses a lower concentration of homogeneous neighbors, it is likely to have a high diversity. This way, a community is composed of pixels with intensity i when this

pixel is found only in the beginning and the end of a sequence of consecutive and collinear pixels in a direction θ (Figure 5).

For both the GLRLM matrix and the GLGLM matrix, the values adopted for θ were equal to 0° , 45° , 90° and 135° . In both situations, because a matrix is necessary for each direction, and six quantizations were considered, the resulting feature vector presented 24 variables.

Pattern recognition

To analyze whether the produced features differentiate between a benign and a malignant pattern, a pattern recognition stage was included in the following methodology, which will be detailed in sequence.

Support Vector Machine

To validate the proposed methodology and classify the masses as benign or malignant, we used the Support Vector Machine (SVM) (Vapnik, 1998). This technique has performed well when applied to image processing of mammograms, especially to distinguish patterns of the mammogram in mass or normal tissue, as reported in Braz et al. (2009), Carvalho et al. (2012) and Martins et al. (2010). Previously, in Rocha et al. (2012), the SVM was used successfully for diagnosing breast regions as benign and malignant.

Broadly, given the set of training samples (x_i, y_i) , the input vector is $x_i \in \mathfrak{R}^n$, the correct classification of the samples is y_i , and the index of each sample point is $i = 1, \dots, n$. The aim of the classification is to estimate the function $f: \mathfrak{R}^n \rightarrow \{\pm 1\}$, which correctly separates the test samples into distinct classes. Each sample x is mapped to a feature space of the highest dimension through the transformation function $z = \Phi(x)$. The hypothesis is that in this new space, the samples can be discriminated by linear iterations. This way, we

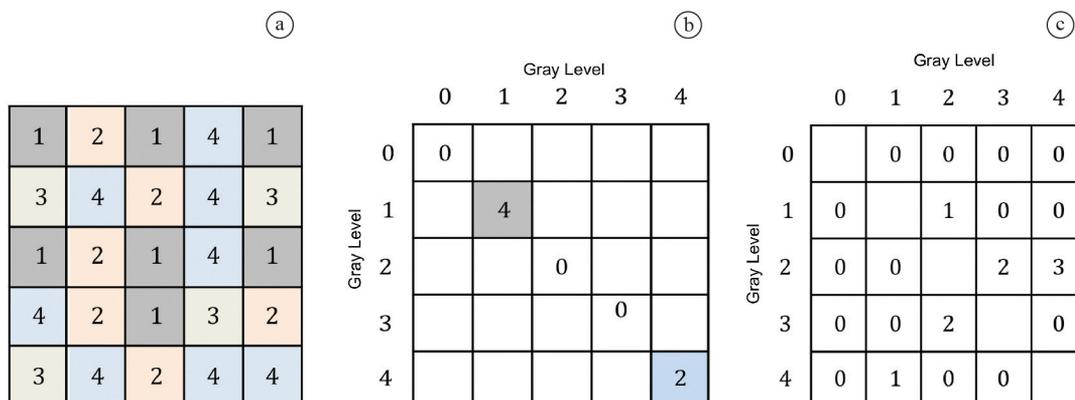


Figure 3. GLCM matrix calculation for $\theta = 0^\circ$ and $d = 2$. (a) ROI 5×5 ; (b) Occurrences of pairs of pixels with the same gray level values; (c) Occurrences of pairs of pixels with different gray level values.

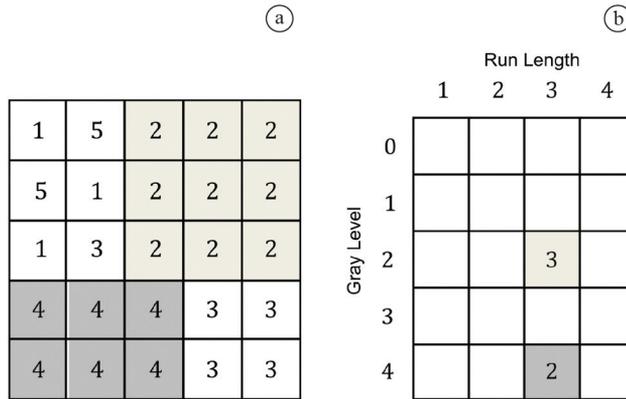


Figure 4. GLRLM matrix calculation for $\theta = 0^\circ$. (a) ROI 5×5 ; (b) Occurrences of gray levels with run length of $k = 3$.

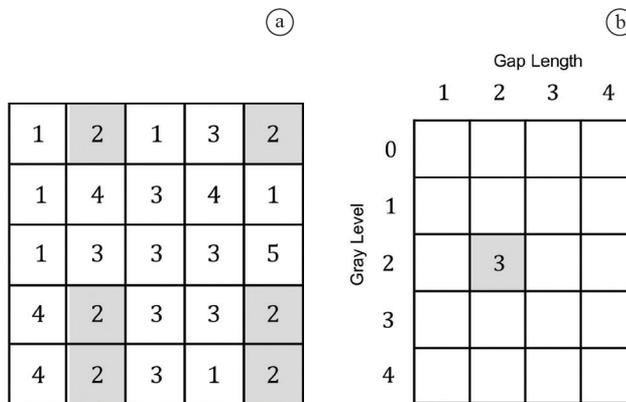


Figure 5. GLGLM matrix calculation for $\theta = 0^\circ$. (a) ROI 5×5 ; (b) Occurrences of gray levels with gap length of $k = 2$.

can say that the decision function can be improved with support from the kernel function $K(x, y)$, which is represented by $\Phi(x)$ (Haykin and Engel, 2001). Therefore, the SVM issue is treated as a linear function of optimization represented by:

$$f(x) = \sum_{i=1}^l \alpha_i y_i K(x, x_i) + b \quad (4)$$

where $K(x, x_i) = \Phi(x)$, and that the coefficient α_i and the variable b are obtained through the optimization of a dual quadratic system based on the equation:

$$\min_{(w, b, \xi)} 1/2 w^T \cdot w + C \sum_{i=1}^l \xi_i \quad (5)$$

$$\text{subject to: } y_{(i)} [w^T \cdot \Phi(x) + b] \geq 1 - \xi_{(i)} \quad \xi_i \geq 0 \quad (6)$$

where $C > 0$ is a parameter estimated by the user, which corresponds to a classifying error penalty, and ξ_i are the slack variables that penalize training errors.

SVM can perform very complex boundary separations. To do so, we use a function of space transformation (kernel function), which transforms the data in the feature space in a space of a higher dimension tending to infinity, where it is possible

to trace a separation hyperplane. In this paper, we investigated the linear function ($x_i^T x_j$) and the Radial Basis Function (RBF) ($\exp(-\gamma \|x_i - x_j\|^2)$). The goal was to analyze which of them better discriminates the studied pattern.

In this paper, to perform the experiments, various criteria were adopted for the division of the training and test bases in the training stage of the SVM. The criteria were 50/50, 60/40, 70/30, and 80/20. Independent of the adopted proportion, for each configuration, the test was repeated 5 times at random. Because it is a random selection, each experiment had the cost parameters (C) and complexity level of the mapping function γ , which is used when the chosen kernel is RBF of the SVM estimated.

The goal was to analyze whether the accuracies, in all repetitions, behave in a similar way, evidencing how the approach represents the texture pattern in the samples of benign and malignant masses. We used the implementation of the SVM available in the LIBSVM library (Chang and Lin, 2011) to conduct this stage.

Result validation

To measure the performance of the proposed methodology, we calculated certain statistics of the test results. The statistics were accuracy (A), sensitivity (S) and specificity (Sp). Accuracy is defined as $(TP + TN)/(TP + FP + TN + FN)$. Sensitivity is given by $TP/(TP + FN)$. Specificity is defined as: $TN/(TN + FP)$, where TP (true positive), TN (true negative), FP (false positive) and FN (false negative) (Bland, 2000).

Finally, we also performed the performance evaluation of the classifiers by analyzing the Receiver Operating Characteristic (ROC) curve, which relates the sensitivity and specificity of the classifier (Mazurowski et al., 2008).

The index used by this paper in the analysis of the ROC curve was the Az , which represents the area under the ROC curve. The closer the Az index is to 1, the better the discrimination performed by the classifier between benign and malignant classes in the test samples. When the index Az equals 0.5, it means that the classifier could not differentiate between the benign and malignant classes (Brown and Davis, 2006).

Results

The results produced by applying the proposed methodology when performing the tests can be found in sequence. As described in the methods section, for each proportion of training and testing, 5 repetitions were performed. The parameters used for each test were estimated using genetic methods of selection, implemented in the reference library used. The parameters are unique for each test and represent the mapping function of the feature vectors for support vectors. Thus, due to the random selection of the train and test bases, the parameters cannot be reused. We also present the average accuracy, sensitivity, and specificity of each proportion with their standard deviation and present the best result obtained by the experiment.

Table 2 presents the results produced by the experiment using the Gleason index. The approach that presented the best result was the GLCM from the representation of the ROI performed by the main diagonal of the matrix, combined with the RBF kernel in 80/20 proportion, presenting an average accuracy of 84.33%. At this proportion, the best result was 86.66% accuracy, 90% sensitivity, 83.33% specificity and an area Az of 0.86.

The results obtained using the Menhinick Index are listed in Table 3. In this configuration, the technique with the best results was GLCM using the whole matrix to represent the ROI and RBF kernel but at the proportion 50/50, with an 83.33% average of

accuracy. The best result generated at this proportion had 85.33% accuracy, 81.71% sensitivity, 89.70% specificity and an area Az of 0.85.

Discussion

As we found from the obtained results, the two indexes present very similar performances; however, in general, the Gleason index was slightly superior, if we consider the greater average of accuracy.

By analyzing all results produced in both experiments, we noticed that the best performance of each index was obtained through the combination of RBF kernel with GLCM matrixes. In the Gleason index, the standard deviations of the attained average accuracies were 2.02% and 0.44% for diagonal GLCM and whole matrix, respectively. In the Menhinick Index case, the standard deviations of the averages were 0.99% and 0.91%, respectively, for the diagonal GLCM and whole matrix. We can observe that, in both situations, there were no discrepancies in the averages. This shows that the results behave in a similar way, evidencing that the tested approaches represent well the texture pattern of the samples of benign and malignant masses.

To have a more detailed analysis of the results produced by this work, we chose, at random, 5 malignant ROIs and 5 benign ROIs from the set of samples used in the tests of the DDSM database.

As we can observe from Figures 6a and b, the masses visually have a spiculated outline, indicating a high probability of malignancy. Figures 7a and b have a regular outline, suggesting a high probability of benignancy. However, only by visual analysis of Figures 6c, d, e, 7c, d and e, it is not possible for the specialist to perform a precise diagnosis because those samples do not have well-defined outline features and have similar textures.

Given the difficulty of differentiation between malignant and benign patterns of the masses, the ideal for a more precise classification is to combine geometry and texture features to perform this task. However, the proposed method shows good results if compared to related works (Table 1), even using only texture analysis to describe the masses.

To provide evidence of the relevance of the proposed method, from the results presented in Table 2 and Table 3, graphs were generated for some of the produced texture features, in the best and worst case of each diversity index using the RBF kernel, for all samples in Figure 6 and Figure 7.

The worst result presented by the Gleason index was for the feature extraction performed from the diversity calculated by the histogram. This can be verified by the graph in Figure 8a, where the 6 produced texture features (one for each quantization), for both

Table 2. Results Gleason Index.

Technique	Kernel	Proportion	Average accuracy	Average sensitivity	Average specificity
Histogram	Linear	50/50	57.06 ± 5.64	54.08 ± 17.23	60.05 ± 12.08
		60/40	59.66 ± 2.56	60.17 ± 9.48	59.07 ± 7.95
		70/30	57.99 ± 4.24	65.42 ± 11.95	50.47 ± 13.97
		80/20	57.00 ± 5.41	55.53 ± 11.04	58.57 ± 9.63
	RBF	50/50	56.40 ± 3.83	59.26 ± 5.44	53.49 ± 5.88
		60/40	58.66 ± 2.72	58.02 ± 7.81	59.33 ± 7.32
		70/30	56.44 ± 4.63	72.92 ± 6.68	39.93 ± 6.22
		80/20	56.00 ± 4.90	61.28 ± 12.94	50.80 ± 12.40
GLCM Main Diagonal	Linear	50/50	72.13 ± 10.32	88.40 ± 8.38	55.90 ± 28.45
		60/40	76.50 ± 6.57	84.37 ± 7.06	68.55 ± 17.05
		70/30	63.33 ± 10.54	98.21 ± 2.65	28.58 ± 25.37
		80/20	82.92 ± 5.45	86.86 ± 5.99	78.97 ± 15.10
	RBF	50/50	79.33 ± 1.40	83.13 ± 3.99	75.35 ± 4.60
		60/40	80.16 ± 3.09	84.16 ± 7.32	76.07 ± 6.20
		70/30	82.89 ± 1.94	85.43 ± 7.62	80.29 ± 3.99
		80/20	84.33 ± 3.59	82.75 ± 7.32	85.86 ± 3.74
GLCM Whole Matrix	Linear	50/50	55.73 ± 1.91	96.16 ± 4.71	15.34 ± 10.24
		60/40	67.00 ± 8.14	85.98 ± 9.00	47.93 ± 20.77
		70/30	58.88 ± 9.35	89.18 ± 3.94	28.61 ± 19.59
		80/20	60.33 ± 7.33	89.20 ± 7.51	31.46 ± 18.27
	RBF	50/50	77.86 ± 2.68	83.86 ± 7.85	71.82 ± 5.29
		60/40	79.00 ± 4.70	81.64 ± 7.47	76.34 ± 7.53
		70/30	78.44 ± 2.29	81.70 ± 8.57	75.28 ± 7.86
		80/20	78.00 ± 2.67	82.14 ± 6.77	73.92 ± 6.76
GLGLM	Linear	50/50	56.66 ± 3.91	85.75 ± 17.49	27.49 ± 28.12
		60/40	55.00 ± 4.34	74.98 ± 15.07	34.96 ± 23.27
		70/30	58.88 ± 5.35	87.50 ± 10.54	30.19 ± 22.74
		80/20	57.00 ± 5.21	86.25 ± 18.60	27.86 ± 23.97
	RBF	50/50	59.60 ± 3.54	69.25 ± 4.21	50.07 ± 8.06
		60/40	60.83 ± 6.43	71.25 ± 11.67	50.43 ± 10.27
		70/30	60.66 ± 4.53	77.92 ± 4.29	43.35 ± 11.21
		80/20	63.66 ± 4.52	73.53 ± 8.05	53.82 ± 2.66
GLRLM	Linear	50/50	53.06 ± 2.41	75.43 ± 20.53	30.68 ± 23.75
		60/40	49.50 ± 3.68	89.23 ± 16.34	9.80 ± 19.59
		70/30	49.33 ± 6.07	79.90 ± 19.78	18.85 ± 24.07
		80/20	50.33 ± 5.81	81.82 ± 22.52	18.90 ± 24.77
	RBF	50/50	53.33 ± 4.28	64.56 ± 10.28	42.03 ± 6.27
		60/40	52.33 ± 4.13	72.32 ± 12.41	32.33 ± 15.05
		70/30	52.66 ± 6.23	71.24 ± 12.00	34.09 ± 16.67
		80/20	55.66 ± 8.27	68.52 ± 14.75	42.74 ± 9.79

the malignant samples and the benign, present many values in the same range, making the differentiation of the sample classes difficult by the classifier.

Figure 8b shows the graph generated by the best result produced by the Gleason index, obtained through the GLCM matrix parting from the diagonal of the matrix. In the construction of the graph, 10 features were used (1 quantization × 5 distances × 2 directions). The quantization used 8 gray levels, distance $d = 1, 2, 3, 4$ and 5, and $\theta = 0^\circ$ and 45° . We can observe that, for

each line in the graph, this form of representation of the ROI produces the majority of the features of the malignant and benign classes with different value ranges, making it possible for the classifier to have an average accuracy of 84.33%.

The Menhinick index, in the RBF kernel, as well as the Gleason index also presented the worst results, through feature extraction performed via diversity calculation, using the histogram (Figure 9a). However, the best results were obtained through the diversity calculation with the GLCM approach, using

Table 3. Results Menhinick Index.

Technique	Kernel	Proportion	Average accuracy	Average sensitivity	Average specificity
Histogram	Linear	50/50	54.53 ± 4.53	60.97 ± 21.86	47.44 ± 22.45
		60/40	53.00 ± 2.15	77.85 ± 29.39	27.38 ± 34.19
		70/30	54.44 ± 7.13	65.67 ± 21.02	43.17 ± 22.41
		80/20	55.42 ± 5.12	75.96 ± 21.36	34.93 ± 29.00
	RBF	50/50	56.53 ± 4.35	66.96 ± 11.96	45.82 ± 8.51
		60/40	60.00 ± 2.04	71.15 ± 7.21	48.46 ± 8.25
		70/30	59.50 ± 6.46	65.29 ± 16.41	53.62 ± 5.45
		80/20	57.66 ± 1.34	64.66 ± 7.94	49.96 ± 10.32
GLCM Main Diagonal	Linear	50/50	71.20 ± 11.03	90.27 ± 8.46	52.23 ± 27.27
		60/40	72.99 ± 6.29	94.35 ± 4.01	51.50 ± 14.96
		70/30	68.22 ± 13.06	90.90 ± 9.31	45.54 ± 30.82
		80/20	69.00 ± 12.45	98.05 ± 1.62	40.00 ± 22.53
	RBF	50/50	78.00 ± 2.39	81.13 ± 5.01	74.37 ± 4.02
		60/40	80.66 ± 4.20	83.71 ± 5.12	77.36 ± 5.90
		70/30	79.77 ± 3.33	77.97 ± 3.73	81.62 ± 6.31
		80/20	78.86 ± 2.63	84.35 ± 6.58	72.88 ± 6.33
GLCM Whole Matrix	Linear	50/50	70.00 ± 10.46	89.26 ± 7.37	50.37 ± 15.91
		60/40	76.99 ± 10.39	85.53 ± 6.86	67.95 ± 21.07
		70/30	77.33 ± 4.13	94.85 ± 1.81	59.78 ± 6.42
		80/20	77.33 ± 5.64	96.24 ± 2.66	58.50 ± 10.42
	RBF	50/50	83.33 ± 1.52	87.40 ± 3.96	79.12 ± 6.25
		60/40	81.00 ± 3.81	82.00 ± 6.13	79.90 ± 7.22
		70/30	82.00 ± 3.54	87.36 ± 5.37	76.78 ± 7.22
		80/20	83.00 ± 1.63	82.19 ± 4.86	83.66 ± 4.61
GLGLM	Linear	50/50	60.93 ± 6.75	80.51 ± 14.88	40.97 ± 29.50
		60/40	60.33 ± 5.26	86.94 ± 8.97	33.78 ± 23.35
		70/30	54.88 ± 6.83	94.92 ± 6.19	14.85 ± 8.45
		80/20	60.66 ± 6.55	80.89 ± 15.28	39.91 ± 29.03
	RBF	50/50	67.33 ± 1.46	68.18 ± 8.34	66.50 ± 9.83
		60/40	66.16 ± 3.86	63.15 ± 7.86	69.25 ± 2.60
		70/30	69.77 ± 1.78	69.34 ± 5.78	70.25 ± 6.95
		80/20	66.33 ± 2.45	65.71 ± 2.21	66.82 ± 4.78
GLRLM	Linear	50/50	51.06 ± 2.25	54.98 ± 16.05	47.36 ± 16.18
		60/40	53.00 ± 5.34	58.77 ± 14.66	47.29 ± 12.43
		70/30	55.77 ± 2.93	64.71 ± 15.80	46.64 ± 14.50
		80/20	58.00 ± 4.14	87.97 ± 14.44	27.93 ± 21.43
	RBF	50/50	55.46 ± 2.04	57.95 ± 6.51	53.01 ± 5.81
		60/40	53.16 ± 4.79	58.46 ± 13.51	47.76 ± 13.72
		70/30	59.77 ± 4.35	74.66 ± 8.67	44.61 ± 11.05
		80/20	58.33 ± 6.24	71.70 ± 8.44	44.85 ± 7.81

the whole matrix to represent the ROI. In the graph construction of Figures 9a and b, we respected the same requisites used in elaborating the graphs referring to the Gleason index.

The overlap among the values of the malignant and benign samples shown in the graphs of Figure 8b and Figure 9b happen because the two techniques produce some texture features with the same range of values. These data are possibly explained by the fact that the two researched indexes only take into consideration two parameters in the diversity calculation: the number

of species and the total number of individuals. Factors such as the size of samples and weight given to rare species are not considered in the calculation of those indexes. However, as previously demonstrated, the textures of malignant and benign masses are similar; therefore, such factors can be determinants for a better discrimination between the mass classes. This way, it is necessary to research other diversity indexes in order to prove this supposition.

Comparing results produced by this work with related ones (presented in the introduction) was not

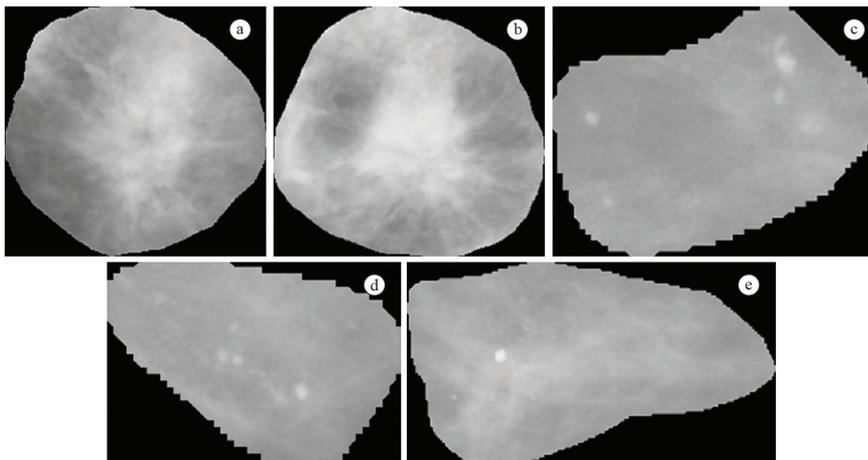


Figure 6. Malignant ROI. (a) to (e) correspond to the malignant samples labeled M1 to M5.

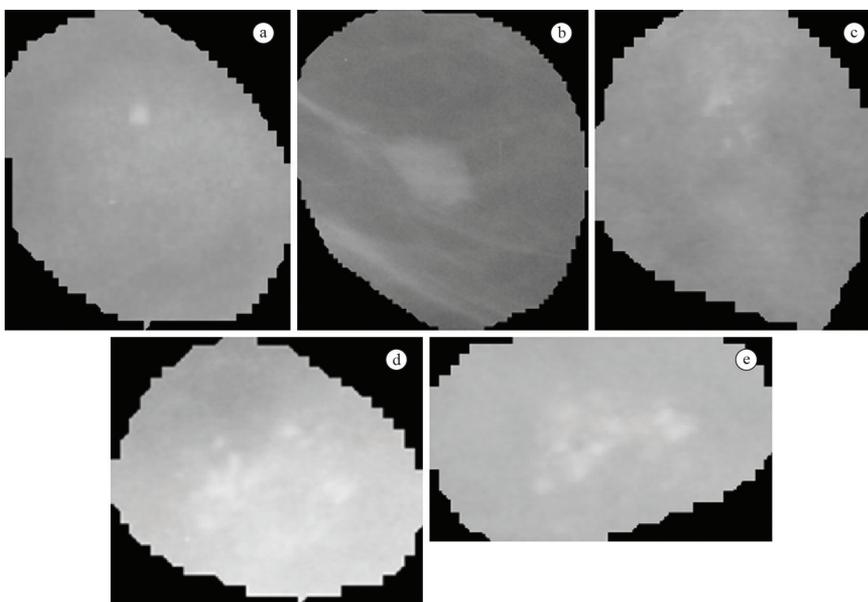


Figure 7. Benignant ROIs. (a) to (e) correspond to the benignant samples labeled B1 to B5.

a simple task because, as previously discussed, the works present different methodologies, image bases and number of samples used in the experiments. However, by analyzing Table 1, it is possible to establish some conclusions. The first is that, from approaches that extract features only via texture analysis, the performance of the methodology proposed in this work is, in general, superior to the results presented by other works. Even when comparing the results of this work with those that combine texture and geometry features, we can observe that the results generated here are superior to the majority of the related works, outlining that the proposed methodology is quite promising.

Another point that deserves attention is that, in this work, a detailed analysis of the results shows

the average accuracy, sensitivity, and specificity obtained. In other words, not only isolated results are considered. This way, it is possible to verify the consistency of the produced results for discriminating between malignant and benign patterns.

Overall, although the two indexes presented statistically similar performances, the Gleason index was slightly superior. The Gleason diversity index, combined with the approach of the GLCM matrix, resulted in 86.66% accuracy, 90% sensitivity, 83.33% specificity and an area A_z of 0.86. However, it is still necessary to not only perform more tests but also investigate the performance in other bases of regions of interest, as well as other diversity indexes, in order to perform a more detailed analysis of the proposed methodology.

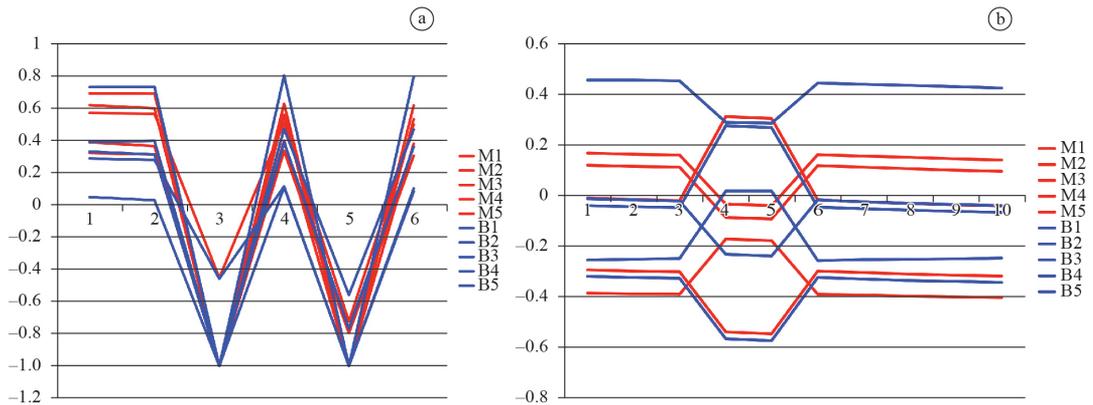


Figure 8. Graphics for Gleason's Index. (a) Histogram; (b) GLCM main diagonal.

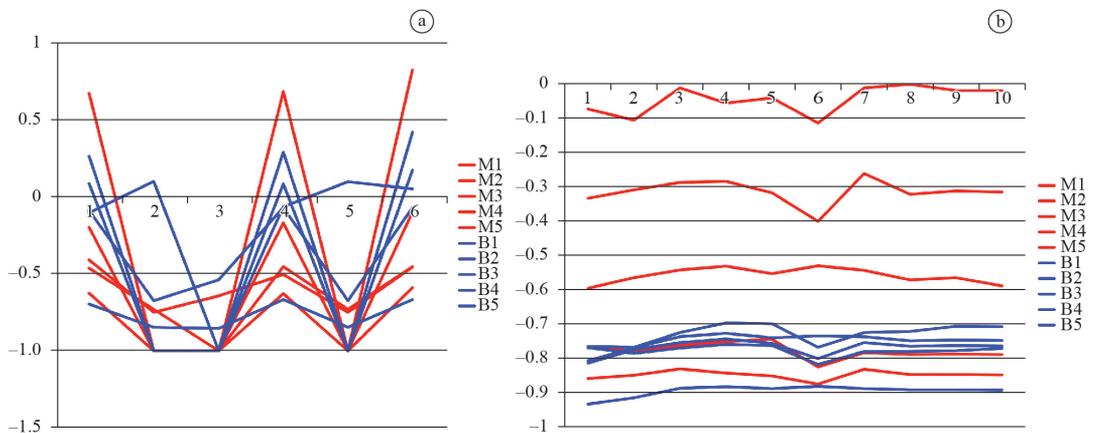


Figure 9. Graphics for Menhinick's Index. (a) Histogram; (b) GLCM whole matrix.

Acknowledgements

The authors acknowledge CAPES, CNPq and FAPEMA for the financial support.

References

American Cancer Society - ACS. Learn about breast cancer. [cited 2011 nov 17]. Available from: <http://www.cancer.org>.

Bland M. An Introduction to medical statistics. 3rd ed. Oxford University Press; 2000.

Braz JG, Silva AC, Paiva AC, Oliveira ACM. Classification of breast tissues using Moran's index and Geary's coefficient as texture signatures and SVM. *Computers in Biology and Medicine*. 2009; 39:1063-72. PMID:19800057. <http://dx.doi.org/10.1016/j.combiomed.2009.08.009>

Brower JE, Zar JH, Ende CV. Field and laboratory methods for general ecology. 4th ed. Dubuque: McGraw-hill Science; 1997.

Carvalho PMS, Paiva AC, Silva AC. Classification of breast tissues in mammographic images in mass and non-mass using McIntosh's Diversity Index and SVM. In: *Machine Learning and Data Mining in Pattern Recognition*. In: MLDM 2012: Proceedings of the 8th international

conference on Machine Learning and Data Mining in Pattern Recognition – LNCS; 2012; Berlin. Berlin: MLDM; 2012. p. 482-94.

Chang C, Lin C. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*. 2011; 2(3):27-27pages. [cited 2012 Oct 6]. Available from: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

Brown CD, Davis HT. Receiver operating characteristics curves and related decision measures: a tutorial. *Chemometrics and Intelligent Laboratory Systems*. 2006; 80:24-38. <http://dx.doi.org/10.1016/j.chemolab.2005.05.004>

Galloway MM. Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*. 1975; 4:172-9. [http://dx.doi.org/10.1016/S0146-664X\(75\)80008-6](http://dx.doi.org/10.1016/S0146-664X(75)80008-6)

Gonzalez RC, Woods RE. Digital image processing. 2nd ed. New York: Prentice Hall; 2002.

Haralick RM, Shanmugan K, Dinstein I. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*. 1973; 3(6):610-21. <http://dx.doi.org/10.1109/TSMC.1973.4309314>

Haykin S, Engel P. Redes neurais: princípios e prática. Rio de Janeiro: Bookman; 2001.

- Heath M, Bowyer KW, Kopans D, Moore R, Kegelmeyer WP. DDSM: The Digital Database for Screening Mammography. In: International Workshop on Digital Mammography: Proceedings of the 5th International Workshop on Digital Mammography; 2001. Medical Physics Publishing, 2001. p. 212-8.
- Kinoshita SK, Pereira RR, Honda MO, Rodrigue J, Marques PMA. An automatic method for detection of the nipple and pectoral muscle in digitized mammograms. In: Congresso Latino-Americano de Engenharia Biomédica - CLAEB 2004: Anais do Congresso Latino-Americano de Engenharia Biomédica; 2004 Set 22-25; João Pessoa, Brasil. João Pessoa: International Federation for Medical and Biological Engineering; 2004. p. 1303-6.
- Kopans DB. *Imagens da mama*. 2nd ed. Porto Alegre: Medsi; 2000.
- Lim WK, Er MJ. Classification of mammographic masses using generalized dynamic fuzzy neural networks. *Medical Physics*. 2004; 31(5):1288-95. <http://dx.doi.org/10.1118/1.1708643>
- Liu X, Liu J, Zhilin F. Mass classification in mammography with morphological features and multiple kernel learning. In: *iCBBE 2011: Proceedings of the 5th International Conference on Bioinformatics and Biomedical Engineering*; 2011 May 10-12; Wuhan. Wuhan: IEEE; 2011. p 1-4. <http://dx.doi.org/10.1109/icbbe.2011.5780356>
- Liu X, Liu J, Zhou D, Tang J. A benign and malignant mass classification algorithm based on an improved level set segmentation and texture feature analysis. In: *iCBBE 2010: Proceedings of the 4th International Conference on Bioinformatics and Biomedical Engineering*; 2010 June 18-20; Chengdu. Chengdu: IEEE; 2010. p. 1-4. <http://dx.doi.org/10.1109/ICBBE.2010.5518284>
- Magurran AE. *Measuring biological diversity*. New York: Wiley; 2004.
- Martins LO, Silva AC, Paiva AC, Braz JG. Comparison of Support Vector Machines and Bayesian Neural Networks performance for breast tissues using geostatistical functions in mammographic images. *International Journal on Computational Intelligence and Applications*. 2010; 9:271-88. <http://dx.doi.org/10.1142/S1469026810002914>
- Mavroforakis ME, Georgiou HV, Dimitropoulos N, Cavouras D, Theodoridis S. Mammographic masses characterization based on localized texture and dataset fractal analysis using linear, neural and support vector machine classifiers. *Artificial Intelligence in Medicine*. 2006; 37(2):145-62. PMID:16716579. <http://dx.doi.org/10.1016/j.artmed.2006.03.002>
- Mazurowski MA, Habas PA, Zurada JM, Lo JY, Baker JA, Tourassi GD. Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. *Neural Networks*. 2008; 21:427-36. PMID:18272329 PMID:PMC2346433. <http://dx.doi.org/10.1016/j.neunet.2007.12.031>
- Mcintosh RP. An index of diversity and the relation of certain concepts to diversity. *Ecological Society of America*. 1967; 48:392-404.
- Menhinick EF. A Comparison of some species-individuals diversity indices applied to samples of field insects. *Ecology*. 1964; 45(4):859-61. <http://dx.doi.org/10.2307/1934933>
- Mu T, Nandi AK, Rangayyan RM. Classification of breast masses using selected shape, edge-sharpness, and texture features with linear and kernel-based classifiers. *Journal of Digital Imaging*. 2008; 21(2):153-69. PMID:18306000 PMID:PMC3043867. <http://dx.doi.org/10.1007/s10278-007-9102-z>
- Rangayyan RM, Nguyen TM, Ayres FJ, Nandi AK. Effect of pixel resolution on texture features of breast masses in mammograms. *Journal of Digital Imaging*. 2010; 23(5):547-53. PMID:19756865 PMID:PMC3046677. <http://dx.doi.org/10.1007/s10278-009-9238-0>
- Rocha SV, Braz JG, Paiva AC, Silva AC. Uso da função K de Ripley e Máquina de Vetores de Suporte para diagnóstico de regiões da mama. In: *XXIII CBEB 2012: Anais do XXIII Congresso Brasileiro de Engenharia Biomédica*; 2012 out 1-5; Porto de Galinhas, Brasil. Porto de Galinhas; 2012. p. 1153-7.
- Sahiner B, Chan HP, Petrick N, Helvie MA, Hadjiiski LM. Improvement of mammographic mass characterization using spiculation measures and morphological features. *Medical Physics*. 2001; 28(7):1455-65. PMID:11488579. <http://dx.doi.org/10.1118/1.1381548>
- Shi J, Sahiner B, Chan H, Ge J, Hadjiiski L, Helvie MA, Nees A, Wu Y, Wei J, Zhou C, Zhang Y, Cui J. Improvement of mammographic mass characterization using spiculation measures and morphological features. *Medical physics*. 2007; 35(1):280-90. <http://dx.doi.org/10.1118/1.2820630>
- Suganthi M, Madheswaran M. An improved medical decision support system to identify the breast cancer using mammogram. *Springer Science Business Media*. 2010; 36:79-91.
- Vapnik VN. *Statistical learning theory*. New York: Wiley-Interscience; 1998.
- Varela C, Timp S, Karssemeijer N. Use of border information in the classification of mammographic masses. *Phys. Med. Biol*. 2006; 51(2):425-41. PMID:16394348. <http://dx.doi.org/10.1088/0031-9155/51/2/016>
- Xinli W, Albrechtsen F, Foyb B. Texture features from gray level gap length matrix. In: *Workshop on Machine Vision Applications: Proceedings of the Workshop on Machine Vision Applications*; 1994 Dec 13-15; Kawasaki. Kawasaki; 1994. p. 375-8.

Authors

Simara Vieira da Rocha*, **Geraldo Braz Junior**, **Aristófanés Corrêa Silva**, **Anselmo Cardoso de Paiva**
Applied Computing Group – NCA, Federal University of Maranhão – UFMA, Av. dos Portugueses, s/n,
Campus do Bacanga, São Luís, MA, Brasil.