A False Positive Reduction in Mass Detection Approach using Spatial Diversity Analysis

Geraldo Braz Junior, Simara Vieira da Rocha, Aristófanes Correa Silva, Anselmo Cardoso de Paiva
Applied Computer Group
Federal University of Maranhão
São Luiz - Brazil
geraldo@nca.ufma.br, simara@nca.ufma.br, ari@nca.ufma.br, paiva@nca.ufma.br

Abstract—**Efforts in image processing and pattern recognition have been made in order to help improving the detection accuracy by physicians. In this paper, we present a analysis that study the use of Diversity Indexes in a Spatial approach as a texture measure in order to distinguish suspicious regions previously detected by segmentation scheme. The description of the pattern is based on the fact that important features could be distributed on the region under certain distance, angle and tonalities. And these tonalities represents species that have a particular associations that may be important distinctions between the pattern of mass and non-mass regions helping do false positive reduction and assisting a physician on a task of verify suspicious regions on a mammogram. The computed measures are classified through a Support Vector Machine and reaches a reduction of 75% of false positives on mass detection methodology.**

Keywords-**Mass False Positive Reduction; Pattern Recognition; Diversity Index.**

I. Introduction

According to the estimative of the American Cancer Society (ACS), the chance of a woman having the breast cancer in any period of her live is a just little below of 1 in 8, and the chances of dying from the disease is of 1 in 35 [1]. The screening mammography is the best way for the precocious detection of any kind of lesion in breast, also cancers. An earlier diagnosis improve the chances of cure. It is also know that in early stages, the treatment is more effective. The digitized mammogram is a image that presents overlapping breast tissues obtained by the X-ray exposure. The overlap and physiological characteristics of the patient can generate a mammogram of low contrast which could lead the physician a wrong diagnostic caused by the repetitive task of analyzing the image [2].

Efforts has been made in the use of image processing and pattern recognition techniques to improve the breast cancer diagnosis results. The goal is to increase detection and diagnostic accuracy providing a second opinion also causing a reduction in the rate of false positive biopsies for cases since the sensitivity of mammography exam is around 85% [3]. Mass detection methodologies based on image processing generally has a step of false positive reduction. This is aimed at decreasing the amount of the segmented regions that were marked as suspected mass, when they are actually normal tissues.

We have verified that the false positive reduction of the tissues extracted from the mammogram as mass and non-mass is a crucial stage in the methodologies for detection of breast cancer and that there is a potential to be explored in measurements that describe texture as the diversity indexes applied in a spatial approach. This study began in previous work of the group [4][5][6], which demonstrate the effectiveness of these methods.

Also, the classification of breast masses in mammography images has motivated many research. Some surveys [7][8] analysis different approaches published in literature. Most approaches are based on the extraction of features such as texture (statistics) [9][10][11], geometry (roundness, sphericity, spicularity) [12][13][14][15], morphological or nominal (BI-RADS) [16][17][18]. These features are submitted to a classification using a pattern-recognition machine as Support Vector Machine or Neural Networks.

This work proposes the use of diversity indexes used in Ecology (Total Diversity, Brillouin, Berger-Parker, J, ED, Hill, Buzas-Gibson) in a spatial approach to describe the patterns found in previously segmented images of mammogram. The main goal is to propose an efficient methodology for reducing false positives and assist detection methodologies that are used on Computer Aided Systems to assist the physician on a task of detection breast cancer.

The remainder of this work is organized as follows. Section II brings a brief introduction to Diversity Indexes. Section III presents a detailed description of the proposed methodology and evaluation used in this work. Section IV presents and discusses about the founded results. Finally, Section V presents the final considerations.

II. Diversity Analysis

This work proposes a study of a texture extraction for false positive reduction inspired in biological process using diversity analysis. Following we present the Diversity Indexes used as feature descriptor as well as the mapping of this problem to a traditional image processing task.

This analysis intends to identify the distribution of a group of species and their interrelations. It is used in Ecology to measure the biodiversity in an ecosystem. The diversity refers to the variety of species in a given community or habitat. Biodiversity is the relationship between the number of species (richness), the pattern of distribution of individuals in their species (evenness) and the dominance of one or more species among others (dominance). All these characteristics can be measured and investigated using indexes usually classified as coverage of local analysis (alpha) or between various habitats (beta) [19].

More generally, the diversity indexes can be used to measure the diversity of a population where each member belongs to one single group or species. This paper intends to study the use of Diversity indexes for false positive reduction in pattern recognition process to distinguish mass and non-mass ROI's extracted from a digitalized mammography image. With this goal we adopt that the pixels are the individuals and their tonalities represents the set of specie.

Considering that each image ROI used has a distribution of gray tones varying from 0 to 255 (8 bits per pixel). Thus, any pixel x of image A have a specie s_i. The set given by $x_0, x_1, ..., x_N$ represents the overall population P where N is the total number of individuals and also of pixels. The set S is the total amount of species of the set $s_0, s_1, ..., s_i$ where i represents a specific specie. The number of individuals of each species is represented by n_i and p_i is the total proportion of the sample belonging to species i, that is $p_i = \frac{n_i}{N}$. This paper investigates the application of the following diversity indexes for texture characterization of mammography images.

The Total Diversity Index [20] estimates the total richness of a population based on species variation. This measure is given by:

$$Td = \sum_{i=1}^{S} w_i(p_i(1 - p_i)) \qquad (1)$$

where w_i is the weight or importance given individually to a specie characterized by $\frac{1}{n_i}$.

The Brillouin Index [21] measures the richness of a known population and it is recommended when that population is not random. Also, this index tends to inform similar results comparable to Shannon-Wienner's Index [22] used in a not completely known population. It is defined by:

$$Hb = (\frac{1}{N})(logN! - \sum_{i=1}^{S} logn_i!) \qquad (2)$$

The Berger-Parker Index [23] is the numerical importance of the most abundant species, defined by:

$$Bp = \frac{max(n_i)}{N} \qquad (3)$$

where $max(n_i)$ is the amount of the most abundant specie.

The J Index [21] compares the observed Shannon-Wiener Index against the distribution of individuals between the observed species which would maximize diversity, and is defined by:

$$J = \frac{H}{H'} \qquad (4)$$

where H is the Shannon-Wiener Index calculated by $H = -\sum_{i=1}^{S} p_i \ln p_i$ and his maximum value H' is given by $H' = logS$.

The ED [20] Index compares the dominance of Simpson with the known species that maximize the diversity. It is a relation between Simpson index D and it's maximum:

$$Ed = \frac{Ds}{Ds'} \qquad (5)$$

where Ds is the Simpson Index [24] given by $Ds = \frac{\sum_{i=1}^{S} n_i(n_i-1)}{N(N-1)}$ and his maximum dominance Ds' calculated by $Ds' = (\frac{S-1}{S}) * (\frac{N}{N-1})$.

The Hill Index [25] calculates the equally abundance distribution of the species, given by:

$$Hill = \frac{\frac{1}{Ds-1}}{e^h - 1} \qquad (6)$$

where Ds and H are respectively the Simpson and Shannon-Wienner's Indexes previously explained. In this index, if the species are equally abundant then the index should take the maximum value. This value will decrease as the relative abundance differ between species.

The Buzas-Gibson Index [26] indicates the degree of evenness trough Shannon index H and is defined by:

$$Bg = \frac{e^H}{S} \qquad (7)$$

where H is given by Shannon-Wienner's Index.

III. PROPOSED METHODOLOGY

The methodology proposed in this work intends to study the uses of diversity indexes in a spatial approach to describe breast segmented tissues and allow a discrimination of then into mass and non-mass. The goal is to perform a false positive reduction of a detection methodology previously published in [27]. The methodology has five steps as illustrated in Figure 1: ROI Acquisition, Enhancement, Spatial Diversity Analysis (using the indexes previously described), Recognition and Validation. Each step is explained in the following text.

In order to test the proposed methodology, this work use a database generated by a detection methodology published in [27]. The ROI's were segmented using a circular template-matching technique. It detect the suspicious regions of randomly extracted 603 mammograms obtained in Digital Database Screening Mammography [28] each one containing only one mass. It were generated a total number of 2679 suspicious regions, including all 603 masses and other 2076
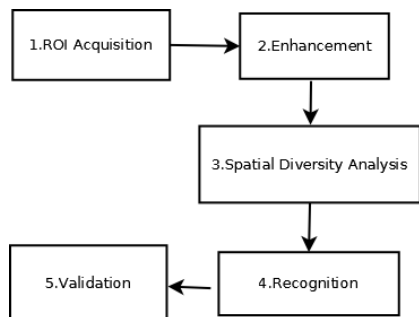
Figure 1.   Steps of the proposed methodology

non-masses. Figure 2 shows some examples of these regions. We could note that mass regions are generally rounded, has high and homogeneous pixel values at their core. Non-mass regions has random contour. Their pixel distribution values is less accentuated than a mass region. So they share similar tonalities but mass regions are more concentrated on high values whereas non-mass regions are more equally distributed over many tonalities.
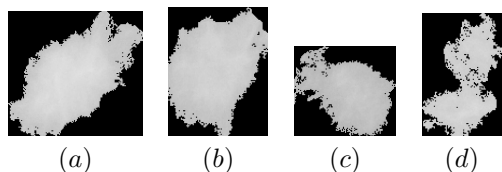

Figure 2.   (a-b) Database mass examples and (c-d) non-mass examples

A. Enhancement

To produce an image with higher quality for the next stages, it was applied for each region an histogram equalization [29]. This step aims to make equally important tones of gray to represent spatial and visual image. Also, after the equalization of the image, six versions of each sample were made through the scalar quantization using six numbers of gray levels (2^8, 2^7, 2^6, 2^5, 2^4 and 2^3). This aims to clustering similar tones of gray in order to enhance their participation in the sample.

B. Spatial Diversity Texture Analysis

Each sample generated after the enhancement step was submitted to a spatial texture analysis using the Diversity Indexes presented in Section II. The spatial analysis measures the autocorrelation between the points over a certain localization or distance and even direction. Our goal is to measure also the autocorrelation information besides their diversity and use this information to create a pattern. For this goal, we adopt three approaches to make the spatial association: Rectangular, Circular and Directional.

1) Rectangular (HVDW): In this, the sample is divided into tiles, using horizontal (H), vertical (V), diagonal (D) and windows (W) like exemplified in Figure 3. The image was divided in 4 horizontal, vertical, diagonal tiles and 9 windows. Each resulting region is used separately.
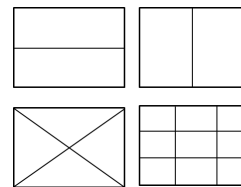

Figure 3.   Rectangular scheme for split image region

2) Circular (CIRC-RING): Here the sample is analyzed by circles of different radius $r_1, r_2, ... r_n$ and also rings formed by the subtraction of two circles of different radius, where $r_1 > r_2$, as exemplified in Figure 4. It was used 5 different equidistant radius. They were obtained by dividing the largest radius into five parts. The largest radius was determined as the radius of the larger circle circumscribed in the region. The use of five radius was defined after empirical tests.
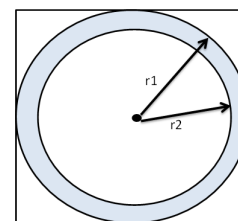

Figure 4.   Circular scheme for split image region

3) Directional (DIR): This analysis [30] is done for each pixel of the sample, that will become a Head point in certain moment (See Figure 5a). In addition, to improve the capacity of describing texture patterns, we have conducted a directional analysis. For it, we must define a direction vector, that is an azimuth which corresponds to an angle in x axis, assuming a lag increasing rate. We assume tolerances for lag and azimuth to better capture neighborhood in a matrix arrangement of pixels. These restrictions are exemplified in Figure 5b.

In practice, for a specified direction, the features may be computed for a number of lags and directions take in consideration tolerances of lag and directions. We also split the set of species in 4 sets without intersection, in order of occurrence, to analyze the contribution of separate species to form a pattern and increase the importance of rare species that could have been obscured by the high occurrence of other common specie. For each one, the initial lag separation distance h was 1 pixel. After, this lag was incremented in 1 unity. The maximum number of lags chosen was 10. The tolerance lag adopted was ± 0.45.
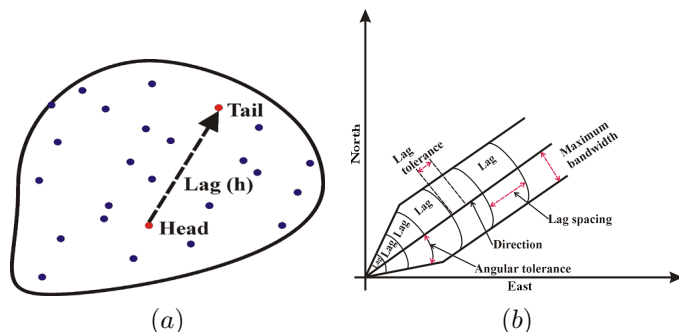
Figure 5. (a) Spatial association between points Head (origin) and Tail (destination) (b) Parameters used for indexes of spatial autocorrelation calculations.

C. Pattern Recognition using Support Vector Machines

In all spatial approaches described in Section III-B, each set of spatial diversity features vectors generated for each samples was classified using Support Vector Machine (SVM) [31] as previous used on [5]. It is a machine learning technique based on the creation of a high dimensional hyperplane of separation optimizing the limits of generalization. It uses radial kernel to mapping characteristics to a high dimensional space where non linear features could be separated. This kernel is defined by:

$$K(x,y) = e^{-\gamma \|x-y\|^2} \qquad (8)$$

where $\gamma > 0$ is a parameter that also is defined by the user.

The process of classifying starts with a definition of a set of training samples and test samples. For this study, the training set for all approaches have the same samples to make possible the approaches comparison. The training set was formed by 400 samples of known mass and 400 samples of known non-mass samples. The balancing is important because of the large difference in number of samples non-mass relative to the mass. The rest of samples was used as a test set. Therefore, to prevent the classifier tends to a certain group, was held the balance amount of the samples.

SVM has the penalty parameter C that regulating the classification function for the best overall accuracy. This parameter was calculated for each set of samples during the training stage as the parameter γ for radial function. Also, with the objective to give more importance classification to mass group, was adjusted weight penalty imposed on the group represented to the value 15 in all classifications tests.

D. Validation

In order to evaluate the classifier with respect to its discrimination ability, this work were validated through following metrics [32]:

- True positive rate (Tpr): percentage of masses correctly classified;
- True negative rate (Tnr): percentage of non-masses correctly classified, meaning false positive reduction;

- Mean false positive per image (Fp/i): average number of false positive per image that remain after the reduction process;
- Means false negative rate per image (Fn/i): average number of incorrectly classified masses.

IV. RESULTS

This section intends to present and discuss the results obtained from the proposed methodology. Table I presents results of all executed tests for each approach. The best outcome for this analysis happens when you have a large reduction of false positives while the hit rate remains high for mass group. Reduce false positives, but reduce the mass hit rate implies to produce results that do not detect cancer.

The best result found reaches 75.81% reduction of false positives (Tnr) with a rate of 93.53% of Tpr using DIR approach and Berger-Parker index. This results means that for the image database evaluated were generated 0.833 false positives in average and 13.195 masses were lost during the step of reducing false positives, generating a rate of 0.065 per image. Buzas-Gibson and Total Diversity also shows similar results with the CIRC, RING and HVDW approaches completing the best results found by the proposed methodology. The differential of the approach using directional Berger-Parker may be the division of species into subgroups appropriate for this index that quantifies the importance of the most abundant species in relation to other species.

Analyzing the other results we conclude that there is a low variation in the settling of masses, represented by Tpr averaging of 93.11%. This demonstrates the effectiveness of the diversity indexes jointly spatial approaches to correctly represent the mass group. The same can not be said of the hit rate of non-mass which obtained average variation of 6.7 percentage points. Analyzing how the indexes are obtained, we conclude that the evenness indexes that represent the abundance of the species over the population had difficulty in presenting notable differences between mass and non-mass patterns. One likely reason is the segmentation process, executed in previous methodology. The resulting regions are internally homogeneous, even in their borders. Even though there are different species within each homogeneous region. We observed that only 5 species occur only in regions of mass and all others are shared. Also, the species distribution in quantitative terms are very similar. Therefore, if the population of the samples do not differ enough and its distribution is similar, so evenness indexes will no be able to effectively translate the patterns found. Unless we use the information of the spatial distribution of species to characterize then. In this sense, the spatial approach becomes important to quantify mutual associations between species and include this information in the process of pattern recognition.

Table I
THE PROPOSED METHODOLOGY RESULTS ORDERED BY OVERALL ACCURACY, TPR AND TNR. BEING THE BEST RESULT SHOWN IN THE UPPER LEFT

Approach	Diversity Index	Tpr	Tnr	Fp/i	Fn/i	Approach	Diversity Index	Tpr	Tnr	Fp/i	Fn/i
DIR	Berger-Parker	93.53	75.81	0.833	0.065	CIRC	Brillouin	90.05	68.40	1.088	0.100
CIRC	Total Diversity	92.03	74.56	0.876	0.080	HVDW	Brillouin	94.19	63.58	1.254	0.058
RING	Total Diversity	92.03	74.55	0.876	0.080	DIR	Total Diversity	88.88	68.20	1.095	0.111
HVDW	Total Diversity	94.52	71.91	0.967	0.055	CIRC	Berger-Parker	92.70	64.02	1.239	0.073
RING	Buzas-Gibson	93.20	73.21	0.922	0.068	RING	Berger-Parker	92.70	64.01	1.239	0.073
CIRC	Buzas-Gibson	93.20	73.21	0.922	0.068	HVDW	Hill	93.36	62.47	1.292	0.066
CIRC	Ed	96.68	69.55	1.048	0.033	RING	J	94.36	60.59	1.357	0.056
DIR	J	89.88	76.30	0.816	0.101	CIRC	J	94.36	60.59	1.357	0.056
HVDW	Berger-Parker	88.39	77.45	0.776	0.116	DIR	Brillouin	96.94	53.94	1.586	0.031
RING	Ed	95.68	69.55	1.048	0.043	HVDW	J	94.02	56.02	1.514	0.060
HVDW	Ed	94.69	70.37	1.020	0.053	DIR	Ed	94.02	55.34	1.538	0.060
DIR	Buzas-Gibson	92.37	70.52	1.015	0.076	DIR	Hill	95.02	53.42	1.604	0.050
HVDW	Buzas-Gibson	91.54	70.32	1.022	0.085	RING	Hill	94.36	52.98	1.619	0.056
RING	Brillouin	90.05	68.40	1.088	0.100	CIRC	Hill	94.36	52.98	1.619	0.056

The Table II show the comparison of the best results found by this work with the previous work from which was obtained the image database. In previously work, the best result was found using jointly geometry and texture features, with the feature selection using a Genetic Algorithm. This result was a false positive reduction of a 83.24%. Although the reduction has been greater, we need to analyze the amount of not detected mass drastically increases in 3 times. So, in comparison with the other work, our methodology ensures that the proposed approach maintains a substantial reduction of false positives while maintaining a high hit rate of mass.

V. CONCLUSION

This paper presented a research on the application of diversity indexes and spatial approaches for the reduction of false positives in mass detection methodologies under digitized mammography's regions. It has been shown that these indexes can quantify with the aid of spatial approaches, texture characteristics able to differentiate masses and non-masses extracted from a image database generated by a previous mass detection methodology.

The best result obtained, decrease at most of 75.81% of false positives without significant loss of mass accuracy rate. Even the result of the lower accuracy attains a reduction of 52.98% of false positives. Compared with previous work, this approach reduces false positive without interfering significantly in detecting masses. We can conclude that the results are promising. But is still necessary to lower the false negative rate, that corresponds a mass regions that are not detected and harmful for a patient. In this methodology we used SVM penalty functions to optimize the rate of mass detection but this was insufficient.

So, although we have obtained good results with this study of diversity indexes applied in a spatial manner, we believe that the reduction of false positives remains to be larger to help permanently specialist. Therefore, as future works, we will extend it using morphological approaches under species to improve the pattern recognition step, also apply feature selection and make more tests in different images databases. Also implementing tools using this methodology to be directly applied by physicians and test it's performance in real situations.

REFERENCES

[1] A. C. S. (ACS), "Learn about breast cancer," 2011, available at http://www.cancer.org, last accessed February 4, 2012.

[2] N. Karssemeijer, J. Frieling, and J. Hendriks, "Spatial resolution in digital mammography," *Investigative radiology*, vol. 28, no. 5, p. 413, 1993.

[3] N. Houssami, S. Ciatto, L. Irwig, J. Simpson, and P. Macaskill, "The comparative sensitivity of mammography and ultrasound in women with breast symptoms: an age-specific analysis," *The Breast*, vol. 11, no. 2, pp. 125–130, 2002.

[4] V. da Silva, A. de Paiva, A. Silva, and A. de Oliveira, "Semivariogram applied for classification of benign and malignant tissues in mammography," *Image Analysis and Recognition*, pp. 570–579, 2006.

[5] G. Braz Junior, A. Cardoso de Paiva, A. Corrêa Silva, and A. Cesar Muniz de Oliveira, "Classification of breast tissues using moran's index and geary's coefficient as texture signatures and svm," *Computers in Biology and Medicine*, vol. 39, no. 12, pp. 1063–1072, 2009.

[6] G. Junior, A. Paiva, A. Silva, and A. de Oliveira, "Classification of breast tissues using getis-ord statistics and support vector machine," *Intelligent Decision Technologies*, vol. 3, no. 4, pp. 197–205, 2009.

[7] J. Bozek, M. Mustra, K. Delac, and M. Grgic, "A survey of image processing algorithms in digital mammography," *Recent Advances in Multimedia Signal Processing and Communications*, pp. 631–657, 2009.

Table II
COMPARISON OF THE RESULTS OBTAINED BY THE PROPOSED METHODOLOGY WITH PREVIOUS WORK

Work	Approach	Index	Tpr	Tnr	Fp/i	Fn/i
Our	DIR	Berger-Parker	93.53	75.81	0.833	0.065
[27]	CIRC+RING	Simpson Index	78.33	79.9	0.69	0.22
[27]	CIRC+RING	Simpson + Geometry	83.24	84.14	0.55	0.17

[8] H. Cheng, X. Shi, R. Min, L. Hu, X. Cai, and H. Du, "Approaches for automated detection and classification of masses in mammograms," *Pattern recognition*, vol. 39, no. 4, pp. 646–668, 2006.

[9] R. Conceicao, M. O'Halloran, M. Glavin, and E. Jones, "Support vector machines for the classification of early-stage breast cancer based on radar target signatures," *Progress In Electromagnetics Research*, vol. 23, pp. 311–327, 2010.

[10] F. Moayedi, Z. Azimifar, R. Boostani, and S. Katebi, "Contourlet-based mammography mass classification using the svm family," *Computers in Biology and Medicine*, vol. 40, no. 4, pp. 373–383, 2010.

[11] M. M. Eltoukhy, I. Faye, and B. B. Samir, "Breast cancer diagnosis in digital mammogram using multiscale curvelet transform," *Computerized Medical Imaging and Graphics*, vol. 4, pp. 296–276, 2010.

[12] D. Tralic, J. Bozek, and S. Grgic, "Shape analysis and classification of masses in mammographic images using neural networks," in *Systems, Signals and Image Processing (IWSSIP), 2011 18th International Conference on*. IEEE, 2011, pp. 1–5.

[13] I. Cheikhrouhou, K. Djemal, and H. Maaref, "Protuberance selection descriptor for breast cancer diagnosis," in *Visual Information Processing (EUVIP), 2011 3rd European Workshop on*. IEEE, 2011, pp. 280–285.

[14] B. Surendiran and A. Vadivel, "Feature selection using stepwise anova, discriminant analysis for mammogram mass classification," *Int. J. Recent Trends in Engineering*, vol. 3, no. 2, pp. 55–57, 2010.

[15] C. Varela, S. Timp, and N. Karssemeijer, "Use of border information in the classification of mammographic masses," *Physics in medicine and biology*, vol. 51, no. 2, pp. 425–442, 2006.

[16] X. Liu, J. Liu, D. Zhou, and J. Tang, "A benign and malignant mass classification algorithm based on an improved level set segmentation and texture feature analysis," in *Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on*. IEEE, 2010, pp. 1–4.

[17] J. Shi, B. Sahiner, H. Chan, J. Ge, L. Hadjiiski, M. Helvie, A. Nees, Y. Wu, J. Wei, C. Zhou *et al.*, "Characterization of mammographic masses based on level set segmentation with new image features and patient information," *Medical physics*, vol. 35, no. 1, p. 280, 2008.

[18] B. Verma, "Novel network architecture and learning algorithm for the classification of mass abnormalities in digitized mammograms," *Artificial Intelligence in Medicine*, vol. 42, no. 1, pp. 67–79, 2008.

[19] A. Magurran, *Measuring biological diversity*. Taylor & Francis, 2004.

[20] A. Magurran and A. Magurran, *Ecological diversity and its measurement*. Princeton university press Princeton, NJ, 1988, vol. 179.

[21] E. Pielou, *Ecological diversity*. Wiley New York, 1975.

[22] C. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.

[23] R. May, "Patterns of species abundance and diversity," *Ecology and evolution of communities*, pp. 81–120, 1975.

[24] E. Simpson, "Measurement of diversity," *Nature; Nature*, 1949.

[25] L. Jost, "The relation between evenness and diversity," *Diversity*, vol. 2, no. 2, pp. 207–232, 2010.

[26] M. Buzas and L. Hayek, "She analysis for biofacies identification," *The Journal of Foraminiferal Research*, vol. 28, no. 3, pp. 233–239, 1998.

[27] A. P. Nunes, A. Paiva, and A. Silva, "Detection of masses in mammographic images using geometry, simpsons diversity index and svm," *Int. J. Signal and Imaging Systems Engineering*, vol. 3, no. 1, pp. 40–51, 2010.

[28] M. Heath, K. Bowyer, and D. Kopans, "Current status of the digital database for screening mammography," *Digital Mammography (Kluwer Academic Publishers)*, pp. 457–460, 1998.

[29] R. Gonzalez and E. Richard, "Woods, digital image processing," *ed: Prentice Hall Press, ISBN 0-201-18075-8*, 2002.

[30] A. Silva, P. Carvalho, and M. Gattass, "Diagnosis of lung nodule using semivariogram and geometric measures in computerized tomography images," *Computer Methods and Programs in Biomedicine*, vol. 79, no. 1, pp. 31–38, 2005.

[31] V. Vapnik, *The nature of statistical learning theory*. Springer Verlag, 2000.

[32] P. Baldi, S. Brunak, Y. Chauvin, C. Andersen, and H. Nielsen, "Assessing the accuracy of prediction algorithms for classification: an overview," *Bioinformatics*, vol. 16, no. 5, p. 412, 2000.