Tsallis Entropy-based mass and non-mass classification

Rafaela Souza Alcântara[1]
Departamento de Ciência da Computação, UFBA, Salvador, BA
Perfilino Eugênio Ferreira Junior[2]
Departamento de Matemática, UFBA, Salvador, BA
Aline Shayonara de Santana Ramos[3]
Departamento de Engenharia da Computação, UFBA, Salvador, BA

1   Introduction

Breast cancer is the most frequent disease among women. In 2012 about 1.7 millions of new cases were registered around the world [1]. Screening mammography provides a significant reduction on breast cancer mortality rate and computational methods are improving this exam efficiency. This paper presents a new approach for mass and non-mass classification based on Tsallis entropy feature extraction from singular value decomposition (SVD).

2   Methodology

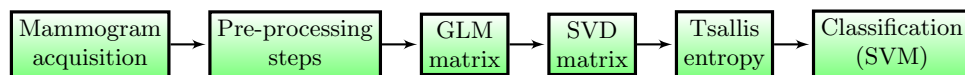Figure 1 shows the pipeline of the process of this work.

Figura 1: Flowchart

Based on [2], gray-level co-occurrence matrix (GLCM) was calculated for each quantized image on database fixing a distance d and varying the direction θ. From GLCM results, SVD was applied using Equation (1).

$$A = U_{(G \times G)} \ \Sigma_{(G \times G)} \ V^T_{(G \times G)} \tag{1}$$

where G corresponds to gray-level value ($G \in \{2^3, 2^4, 2^5, 2^6, 2^7, 2^8\}$), U and V matrices are composed by eigenvectors of AA^T and $A^T A$, respectively. Σ matrix represents the

---

[1]rafaela.alcantara@ufba.br
[2]perfeuge@ufba.br
[3]shaysantanar@gmail.com

2

diagonal matrix ($\Sigma = diag(\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_G)$) composed by all singular values. These values were used then to calculate Tsallis entropy [3] through Equation (2)

$$S = \frac{1 - \sum_{i=0}^{G}(\sigma_i)^q}{q - 1} \qquad (2)$$

where q is an real number that can be changed according to the application.

3 Results

Exhaustive tests were developed to refine q-index range and the best results were provided by $q \in [1.71, 1.79]$. The best results are described on Table 1.

Five metrics statistics were calculated over classification results: accuracy, sensibility, specificity, positive predict value and negative predict value.

Tabela 1: Best results for $q \in [1.71, 1.79]$.

q	d	Statistics (%)				
		Acc.	Sens.	Spec.	PPV	NPV
1.73	1	82.2	79.4	84.6	83.8	80.4
1.76	1	84.2	79.8	88.7	87.6	81.4
1.79	1	82.6	74.5	90.7	88.9	78.0

4 Conclusions

The main proposal was to provide a novel approach for mass and non-mass classification based on Tsallis entropy extraction from SVD of GLCM matrices. To improve accuracy rate, q-index value were adjusted and the best value was provided for pair $(q, d) = (1.76, 1)$ where accuracy was 84.2%.

Referências

[1] Globocan, Fact Sheets by Cancer, http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx, 2012. Accessed 03 July 2015.

[2] R. M. Haralick, K. Shanmugam, and I. Dinstein, Textural Features for Image Classification, *Systems, Man and Cybernetics, IEEE Transactions on*, 3:610-621, 2014. DOI: 10.1007/s40314-014-0163-6.

[3] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, *Journal of Statistical Physics*, 1988, 52:479-487. DOI: 10.1007/BF01016429