

Classification of mammographic lesion based in Completed Local Binary Pattern and using multiresolution representation

Y A S Duarte¹, M Z Nascimento² and D L L Oliveira¹

¹Center of Mathematics, Computing and Cognition, Federal University of ABC (UFABC), Santo Andre, São Paulo, Brazil

²Faculty of Computing (FACOM), Federal University of Uberlândia (UFU), Uberlândia, Minas Gerais, Brazil

E-mail: yan_asd@hotmail.com

Abstract. This paper presents a comparison of two methods for features extraction of mammograms based in completed local binary pattern (CLBP) and wavelet transform. In first part, CLBP was applied in digitized mammograms. In second part, we applied CLBP in the sub-bands obtained from the wavelet multi-resolution representation of the mammographies. In this study, we evaluated the CLBP in the image in the spatial domain and in the sub-bands obtained with wavelet transform. Then, the statistical technique of variance analysis (ANOVA) was used to reduce the number of features. Finally, the classifier Support Vector Machine (SVM) was applied in the samples. The proposed methods were tested on 720 mammographies which 240 was diagnosed as normal samples, 240 as benign lesion and 240 as malign lesion. The images were obtained randomly of the Digital Database for Screening Mammography (DDSM). The system effectiveness was evaluated using the area under the ROC curve (AUC). The experiments demonstrate that the textural feature extraction of the multi-resolution representation was more relevant with value of AUC=1.0. In our experiments, CLBP in the spatial domain resulted in value of AUC=0.89. The proposed method demonstrated promising results in the classification of different classes of mammographic lesions.

1. Introduction

The breast cancer is the second disease that most cause obits among women in all the world. In Brazil, the Instituto Nacional do Câncer (National Institute of Cancer - INCA) estimated more than 52,000 of new cases and more than 12,000 deaths occur due to this disease.

An early diagnostic is essential to increase the chance of a successfully treatment. The most common method used in diagnostic is the analysis of the screen/film mammography by a specialist. However, this method has some limitations. For example, the experience level of the specialist and the human vision limitation. A possible solution to assist the specialist in the breast cancer diagnostic is the used of a computer aided diagnosis (CAD) system. This kind of system can produces consistent results for the classification of disease. Several CAD systems have been employed in the breast lesions classification using the digitalized image of mammography [1, 2, 3]. However, there have been few investigations that utilize more recent techniques for textural feature extraction in mammographies and the obtained results can still be improved.



In this paper we present a method for extraction and classification of regions of interest (ROI) of mammograms. The method was based in the wavelet multi-resolution technique [4], completed local binary pattern (CLBP) operators [5], analysis of variance (ANOVA) [6] and Support Vector Machine (SVM) classifier. We used a subset of mammography cases from a publicly available database, the Digital Database for Screening Mammography.

2. Materials and methods

The method was organized in four steps: (1) calculation of Wavelet coefficients from the mammographies; (2) extraction of CLBP from wavelet coefficients and the original mammographies; (3) application of ANOVA statistical method and (4) classification of images using SVM. Fig. 1 presents the flowchart of the proposed method.

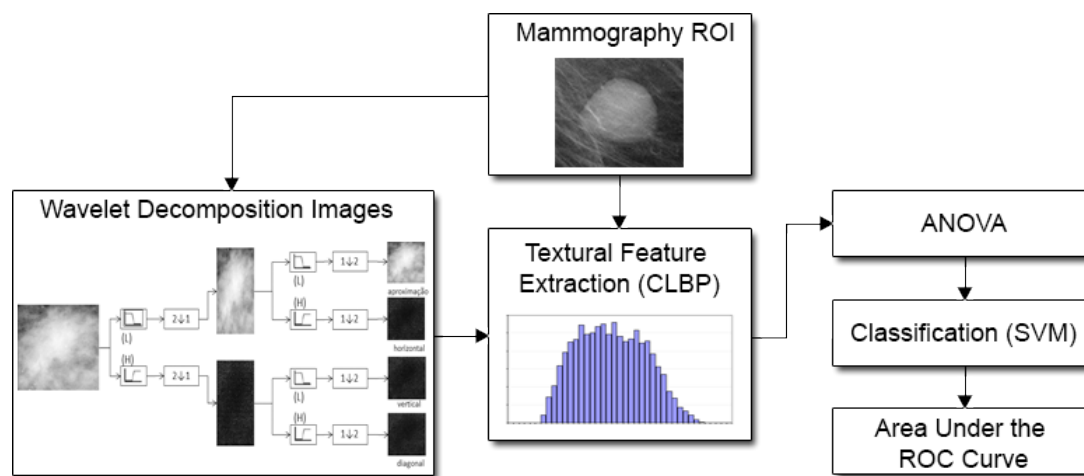


Figure 1: Flowchart of the proposed system.

2.1. Image database

In this experiments we used 720 mammographic images where 240 were diagnosed as normal samples, 240 as benign lesion and 240 as malign lesion. Each ROI was obtained of each mammograms of CC and MLO views get from the the database named Digital Database for Screening Mammography (DDSM). The dataset is composed of 360 mammographic ROI's in cranio-caudal (CC) view and 360 ROI's in mediolateral oblique (MLO) view randomly selected from the DDSM. Therefore, the ROI's were cuttings of size 512 x 512 pixels.

2.2. Wavelet coefficients

A multi-resolution technique allows an image data analysis that results in decomposition details at different resolution levels. Wavelet transform is a technique that represents an image in frequency and time domains [4]. The task is to decompose the original image into sub-bands that preserve high and low frequency information. The most used approach for discrete images is the wavelet transform with quadrature mirrors filter (QMF) banks. This process is made recursively by high-pass and low-pass filtering the image. Four sub-images are generated in this process: the approximation, which represents the original image with a smaller resolution, and the details, which represent the horizontal, vertical and diagonal directions, respectively. However, only the wavelet detail sub-band coefficients were used. For multi-resolution analysis, we used wavelet function Symlet 8 with 2 levels of decomposition.

2.3. Completed Local Binary Pattern

Completed local binary pattern provides a robust and discriminated way for the textural features extraction [5]. This technique has three distinct operators: (1) CLBPs that considers the signal component information; (2) CLBPm that considers the magnitude component information and (3) CLBPc that considers the intensity of the central pixel. This method consist in compare each point of a matrix with its neighbors and get information of signal, magnitude and the analyzed point (central). This procedure results in two histogram, one for signal and other for magnitude, and one matrix that contains the intensity of each analyzed pixel. In our experiments, we used the CLBP as much in the wavelet details sub-bands as in the digitalized image in order to compare the results. The parameters neighbors and radius were set to 8 and 1 respectively.

2.4. Attribute reduction and classification

ANOVA were applied to improve the performance of the classifier and eliminate outliers of data. This technique is a statistical model that compares the means of two or more experiments. This evaluation is performed to the extent that the observed differences between the means are significant for the comparison of two estimates. The ANOVA method was applied in the attributes obtained from the CLBP technique used in the ROI and we also investigated the CLBP operators obtained from the details wavelet sub-bands. This was done to determine the statistical significance of these values. Finally, the SVM classifier was applied. SVM is a machine learning technique based in the statistic learning theory. This theory find a hyperplane with maximum separation between the classes and assumes that the data is linearly separable [7]. In our experiments the SVM was used with linear kernel. For the training and the validation step, the 10-fold cross validation technique was applied with 90% of the data for training and the remaining 10% for test.

3. Results and discussion

The AUC results obtained of evaluation of each group lesion based in the CLBP are presented in Table 1.

Table 1: AUC results for each lesion group with CLBP and wavelet detail sub-band coefficients.

Classes	View	CLBP	CLBP on Coefficients
Benign versus Malign	CC	$0.71 \pm 0.17\%$	$1.0 \pm 0.0\%$
	MLO	$0.71 \pm 0.16\%$	$1.0 \pm 0.0\%$
Benign versus Normal	CC	$0.78 \pm 0.14\%$	$1.0 \pm 0.0\%$
	MLO	$0.78 \pm 0.15\%$	$1.0 \pm 0.0\%$
Normal versus Malign	CC	$0.89 \pm 0.11\%$	$1.0 \pm 0.0\%$
	MLO	$0.87 \pm 0.12\%$	$1.0 \pm 0.0\%$

Fig. 2(a), Fig. 2(b) and Fig. 2(c) show, respectively, ROC curves with the values of AUC more relevant for the classes of images.

The results showed that the technique CLBP applied in the wavelet detail sub-band coefficients were very effective reaching an AUC value of 1.0 in all the classes. The CLPB applied in the mammographic images without the use of wavelet transformed has similarly results in CC and MLO views of each group. In this case, the value of AUC was 0.71 for group of benign versus malign lesions. In the benign versus normal group AUC was 0.78. At last, the Normal versus Malign group reached an AUC value of 0.89. Using CLBP without wavelet detail sub-band coefficients the method provided results different groups evaluated. This experiments shows that the use of operators CLBP and wavelet improve AUC for different classes of mammographic images.

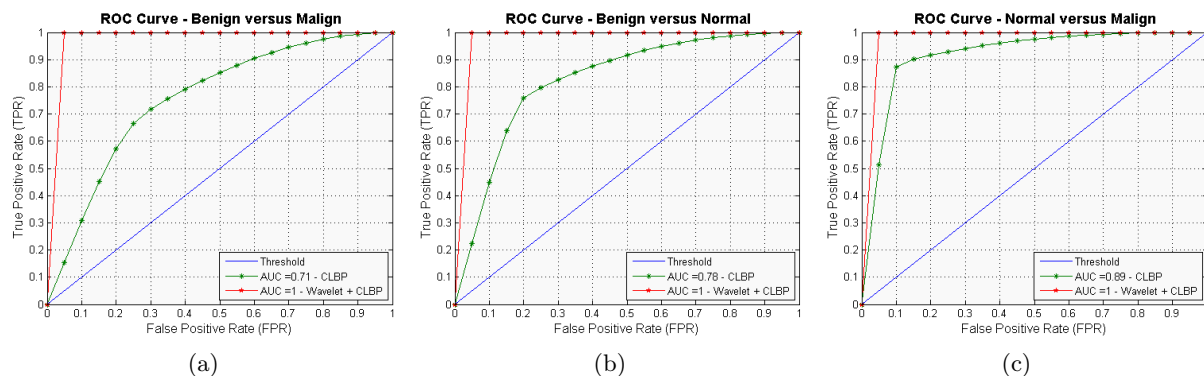


Figure 2: ROC curves of classes of images: Benign versus Malign (a), Benign versus Normal (b) and Normal versus Malign (c).

4. Conclusion

In this study we presented a method based in the wavelet multi-resolution technique, completed local binary pattern (CLBP) operators and analysis of variance (ANOVA) for features extraction. In general, the aggregation of CLBP and wavelet has reached better results in the classification of different kinds of abnormalities. Based on these relevant values of AUC, we can state that the proposed technique using this textural descriptors is an efficient tool for analysis of breast tissue on mammography. For further works we pretend to investigate the proposed method in another mammography database and analyse others techniques of texture feature extraction.

References

- [1] Llado X, Oliver A, Freixenet J, Marti R, Mari J 2009 A textural approach for mass false positive reduction in mammography *Comput Med Imaging Graph* **33** 415-422.
- [2] Dantas R D, Nascimento M Z, Jacomini R S, Pereira D C and Ramos R P 2012 Fusion of two-view information: SVD based modeling for computerized classification of breast lesions on mammograms *Mammography - Recent Advances* **1** 261-278.
- [3] Ramirez-Villegas J F, Ramirez-Moreno D F 2012 Wavelet packet energy, tsallis entropy and statistical parameterization for support vector-based and neural-based classification of mammographic regions *Neurocomputing* **77** 82-100.
- [4] Mallat S 1999 *A Wavelet Tour of Signal Processing* Access Online via Elsevier.
- [5] Guo Z, Zhang D and Zhang D 2010 A completed modeling of local binary pattern operator for texture classification *Image Processing, IEEE Transactions on* **19** 1657-1663.
- [6] Jacomini R S, Nascimento M Z, Dantas R D, Ramos R P 2013. Classification of mass in two views mammograms: use of analysis of variance (ANOVA) for reduction of the features *Recent Patents on Medical Imaging* **3** 80-88.
- [7] Osta H, Qahwaji R, Ipson S 2008 Comparisons of feature selection methods using discrete wavelet transform and support vector machines for mammogram images *In System Signals and Devices* 1-6