

# **Robust Textural Approach for False Positive Reduction in Mammographic Mass**

Norhène Gargouri Ben  
Ayed

University of Sfax, Sfax  
Engineering School  
Computers Imaging Electronics  
and Systems Group(CIELS)

Alima Damak Masmoudi

University of Sfax, Sfax  
Engineering School  
Computers Imaging Electronics  
and Systems Group(CIELS)

Dorra Sellami Masmoudi

University of Sfax, Sfax  
Engineering School  
Computers Imaging Electronics  
and Systems Group(CIELS)

## **ABSTRACT**

In the last few years and over the last decade, different algorithms have been introduced for mass detection in digital mammograms. However, approximately all the proposed methods present a high number corresponding to false positif. In this paper, we propose a novel approach dealing with the problem of false positif reduction.

This proposal is based on the use of a textural approach for the description of the masses texture properties.

The least square estimation is applied to minimize the local difference in order to obtain more stable directional features, this scheme is called adaptative LBP (ALBP) to improve the features extracted using LBP efficiently.

Artificial Neural Network (ANN) are used for the classification of the breast tissue. The evaluation of this approach is based on 1792 Regions Of Interest (ROIs) obtained from the Digital Database Of Screening Mammography (DDSM). The obtained results show that the proposed textural approach is effective in feature extraction and allows mass false positive reduction. Moreover, this proposal demonstrates better performance where compared to the methods of the state of the art.

## **General Terms**

Mass detection, classification.

## **Keywords**

Mass, Adaptative Local Binary Pattern, false positive reduction, classification.

## **1. INTRODUCTION**

Breast cancer is recognised as the major health problem in the united states and in the western countries. It is the most prevalent cancer among women [1]. In 2003, the American cancer Society estimates that between one in eight and one in twelve women in the United States can be affected by breast cancer [2]. According to the European Community, this disease represents 24% of cancer cases [3]. Although breast cancer incidence has increased, the rate of mortality has declined. This mortality reduction is affiliated to the widespread adoption of mammogram screening [4] and to the amelioration of the breast cancer treatment [5]. Mammography is the primary screening tool for early detection of abnormalities. However, this imaging breast modality can miss an important portion of abnormalities [6]. Moreover, a large number of mammogram abnormalities is proved to be benign after the biopsy [7].

The developpement of digital mammography allows the increase of the number of commercial Computer-Aided Detection (CAD) systems. The CAD helps the radiologists in the interpretation and diagnosis of mammograms. The main drawback of such system is the high number of false positives. The false positive reduction algorithms is a good solution. The aim of this paper is to classify the regions as mass or non-mass by using a novel textural approach for mass false positives reduction. The idea to use the textural information has been introduced in divers works [8, 9]. However, we investigate here the use of an extension of local binary patterns (LBP) [10] to mark the directional statistical features with the adaptative LBP features.

In this paper, we demonstrate that when we use the ALBP with the directional statistical features characterization and the ANN for the classification of the ROI can obtain better results on mass false positive reduction. We achieve experiments on a set of 1792 ROIs for DDSM database evaluating and examining the obtained results. The rest of the paper is organized as follows: Section 2 presents a background of the false positive reduction approach. In section 3, we explain our approach by introducing a review of LBP and presenting the proposed extension. Section 4 presents the conclusion.

## **2. Background**

All works trying to detect masses in digital mammography need the mass false positive reduction step to decrease the number of false positives. Figure 1 illustrates an example of suspicious masses. These masses are detected with a mass detection approach. Note that only the small region is a true mass, however, the rest of regions are false positives ones. In fact, the complexity of the texture in the breast tissue induces a detection of false positive cases.

Different techniques for false positive reduction have been developed during the last decade. We will present, in the following subsections, three different approaches to overcome the problem of false positive reduction: The first approach is based on the application of the Principal Component Analysis (PCA); the second approach is based on the proposed extension of the PCA approach noted as 2DPCA and the third approach is based on Local Binary Pattern (LBP).

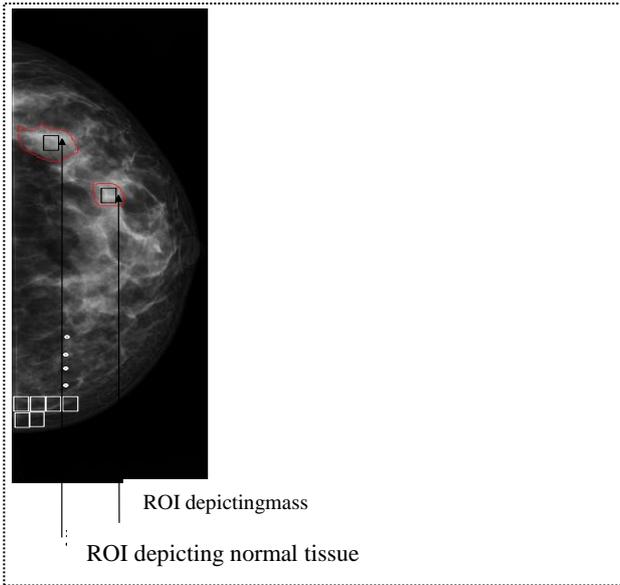


Fig 1: Example of suspicious masses

These algorithms extract image features from ROIs then train the classifiers.

### 2.1 PCA for false positive reduction:

Turk, et al.[11] developed an approach which is based on the eigenface approach to overcome the problem of mass positive reduction. They assume that a set of classified ROIs is available. The variability in the same class is due to the gray value, the texture and the size and shape of mass in the ROI. Karhunen-Loeve transform used the eigenvectors. Turk et al. Instead of eigenfaces, used the eigenrois. Each ROI is transformed into a new space by using the eigenrois. As a result, they obtain a vector describing the contribution of the eigenroi in the representation of the input image. They proposed to use the obtained vector in the construction of a model for the phase of training. The computation of the eigenrois allows 95% of the information.

Each tested ROI is classified in its most similar class. In this paper, Turk, et al.[11] used the K-Nearest Neighbors in the classification.

### 2.2 2D-PCA for false positive reduction:

This approach is an improvement of the eigenfaces approach. The 2D-PCA has two main advantages: firstly, it's simpl and secondly, it's easy in the evaluation of the covariance matrix.

In the approach based on eigenface, each image of size  $m \times n$  is converted to a vector of size  $m \times n$ . However for two dimensional data, it would be trained as a matrix. Oliver et al. [12], presented an extension of the PCA to mass positive reduction when using the 2DPCA approach. The main difference between the two methodologies is that the principal component of PCA is a vector, whereas that of 2DPCA is a set of vectors. In 2DPCA, the comparison of two images means the comparison of constructed features.

### 2.3 LBP for false positif reduction:

#### 2.3.1 Brief review of LBP:

LBP code [10] is computed for a given pixel in an image By comparing its value with those of its neighbours:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p, s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Where  $g_c$  corresponds to the gray value of the central pixel,  $g_p$  is the value of its neighbours, P presents the total number of neighbours and R corresponds to the radius of the neighborhood. If the coordinate of  $g_c$  is (0,0), the coordinates of  $g_p$  are  $(R \cos(2\pi p/P), R \sin(2\pi p/P))$ . For the neighbours that are not located in the grids center, the gray values may be estimated by interpolation.

For each LBP pattern, a histogram may be built. It represents the texture of the image. If U corresponds to the value of an LBP pattern, it is defined as the number of spatial transition between 0 and 1 in the pattern:

$$U(LBP_{P,R}) = |s(g_{p-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)|$$

For the uniform LBP patterns, the patterns have limited transitions ( $U \leq 2$ ) in a circular binary presentation [7]. To achieve the rotation invariance, a locally rotation invariant pattern may be defined as:

$$LBP_{uni,2}^{P,R} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP_{P,R}) \leq 2 \\ P+1 & \text{otherwise} \end{cases}$$

#### 2.3.2 Application of LBP to digital mammography:

Liadò et al [13], proposed to divide the ROI images into local regions of size  $5 \times 5$ . The descriptors of the texture are extracted using LBP and they are finally concatenated. The LBP histogram encodes the local region appearance as well as the spatial relations corresponding to the mass. As a final step, they classify the ROIs as mass or no-mass for false positive reduction by applying the SVM with polynomial Kernel.

### 3. Adaptative LBP for mass false positif reduction:

In order to reduce the estimation error of the local difference, we propose to use the adaptative LBP (ALBP); it is based on the application of the least square estimation (LSE). The application of the LSE as additional information in the ALBP may provide additional information for the stage of classification between mass and non-mass.

To minimize the directional difference, a parameter  $w_p$  is introduced:

$$w_p = \arg \min_w \left\{ \sum_{i=1}^N \sum_{j=1}^M |g_c(i,j) - w_p g_p(i,j)|^2 \right\}$$

The least square estimation (LSE) approach may be applied for optimization and  $w_p$  may be calculated as follows:

$$w_p = \frac{\bar{g}_p^T \bar{g}_c}{\bar{g}_p^T \bar{g}_p}$$

Where  $g_c$  is a column vector which contains all possible  $g_c(i, j)$  pixels and  $g_p$  corresponds to a vector which contains all  $g_p(i, j)$  pixels, these two vectors are defined as follows:

$$\vec{g}_c = [g_c(1,1); g_c(1,2); \dots; g_c(N,M)]$$

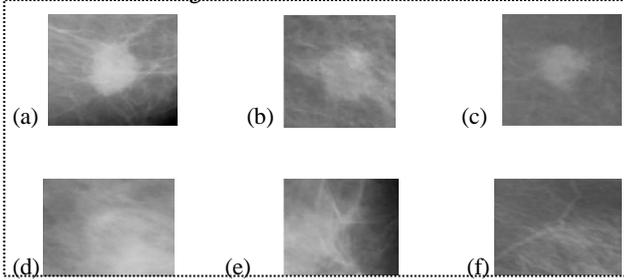
$$\vec{g}_p = [g_p(1,1); g_p(1,2); \dots; g_p(N,M)]$$

The ALBP is defined as follow:

$$ALBP_{P,R} = \sum_{p=0}^{P-1} s(g_p * w_p - g_c), 2^P$$

Where  $\vec{w} = [w_0, w_1, \dots, w_{P-1}]$  corresponds to the weight vector of the ALBP.

Textural approach has been widely used for medical imaging applications. They have proven their usefulness in class discrimination in mammographic images [14]. That's why, we propose in this paper the use of the extension of the LBP noted as ALBP describing the texture and the textons. The proposed methodology consists using ALBP as texture descriptor in order to built local descriptions of the ROIs, and then to combine them in a global description. Afterwards, this descriptor is used to reduce mass false positive in order to classify correctly masses from normal tissue. Figure 2 illustrates some examples of extracted ROIs images.



**Fig 2: Extracted ROIs: (a-c) ROI with mass (d-f) ROI without mass**

As a final step of the algorithm, we use the ANN for classification. This approach has been successfully used in many applications and especially when the expert is not defined [15]. The scheme of the ANN has been developed in the same manner as the biological nervous system and has been widely applied in medical image applications. This methodology adjust the weights between the proposed neurons for the input-output function approximation. So, ANN, has been used in the algorithms of mammographic images to imitate the perception capabilities of expert brain.

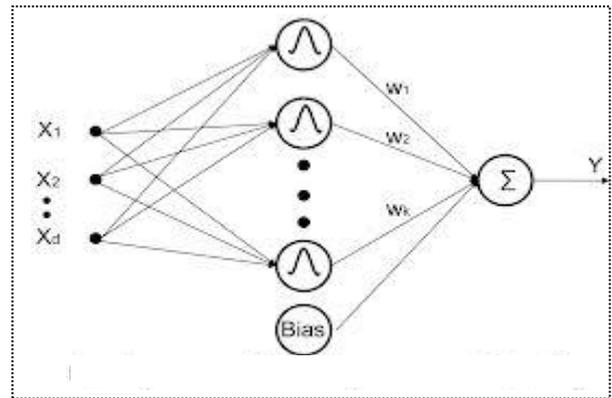
The multilayer perceptron (MLP) as well as the radial basis function network (RBF) are two basic types of ANN frequently applied in recent works.

This work intends to investigate the RBF to exploit the results using the ANN saw its advantages. Details of the used RBF parameter are presented in Table1 and the proposed architecture is presented in fig.3.

**Table 1. Details of RBF network parameter**

Specifications of RBF	parameters
Activation fuction	Radial fuction
Number of hidden layer	1
Number of hidden units	20
Number of input neurons	26
Number of input neurons	1
Maximum mean square error	
Number of iterations	2000

**Fig 3: RBF classifier architecture**



## 4. Results:

This section is composed of the following subsections. In the beginning, we will start with an introduction of the used database for the evaluation of the results is presented. Afterwards, the obtained results for different ratios of ROIs representing masses and ROIs representing normal tissue are shown. Finally, the obtained results for different sizes of ROIs are illustrated.

### 4.1. Database:

The evaluation of the proposed approach is based on 1160 ROIs from the DDSM database [16]. In this selected ROIs, 160 represent true masses and 1000 ROIs correspond to normal tissue. In the DDSM database, for the digitization of the mammograms, diverse scanners were used. The obtained mammographic images are 12 bits per pixel; they were re-scaled to present the same resolution.

### 4.2. Results for different ROIs ratio:

This subsection begins with comparing this work with recent one using DDSM database and the same number of ROIs. The first approach uses the standard PCA, the second approach used the 2DPCA method and the third is the LBP. Fig4.Shows the obtained area under curve (AUC) values for the different ROIs ratios.

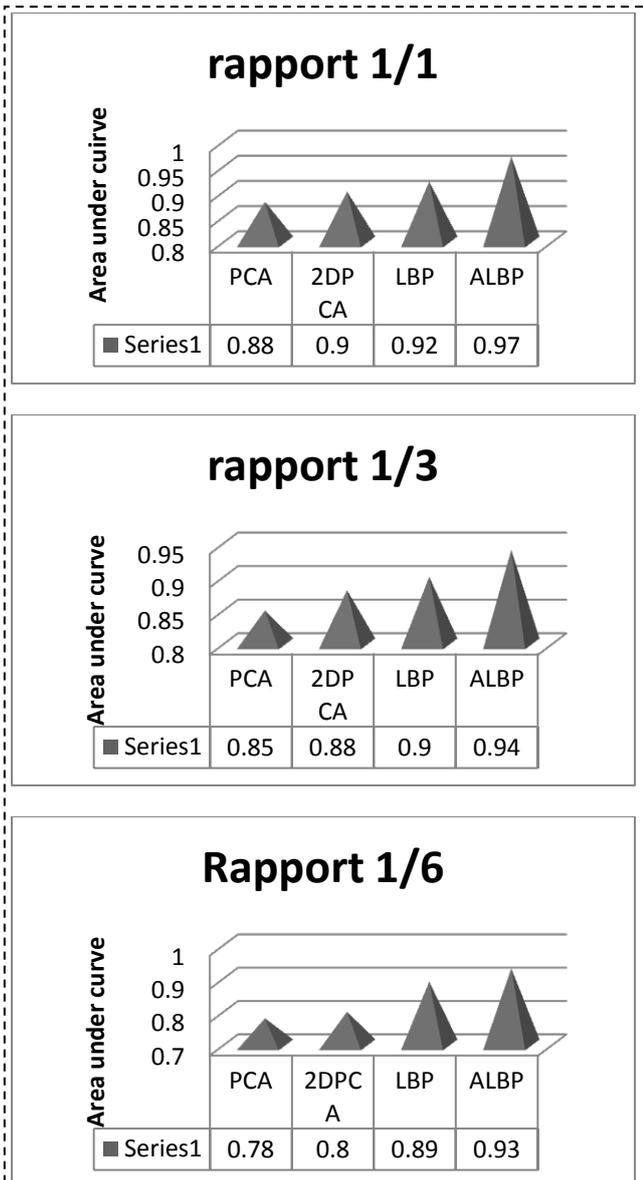


Fig 4: AUC of different image sizes

The performance of the proposed approach is better than PCA and 2DPCA approaches. This is clear through the obtained AUC value (0.83 when using PCA, 0.86 when using 2DPCA, 0.9 when using LBP and 0.95 when using ALBP), thus, the latter gives better results.

#### 4.3. Results for different ROI size:

The size of the lesion may be classified in six groups corresponding to ROI images, these ROIs size intervals are:  $<10\text{mm}^2$ , from 10 to  $60\text{mm}^2$ , from 60 to  $120\text{mm}^2$ , from 120 to  $190\text{mm}^2$ , from 190 to  $270\text{mm}^2$  and  $>270\text{mm}^2$ . Different number of masses is used in each interval; different lesion size is an important characteristic for correct classification of masses. Table 2 illustrates the obtained AUCs for different ROIs size of ROIs.

Table 2. AUC for different size of ROI

AUC							
	Size1	Size2	Size3	Size4	Size5	Size6	Mean
PCA	0.57	0.69	0.7	0.71	0.75	0.8	0.7
2DPCA	0.89	0.8	0.88	0.83	0.82	0.91	0.85
LBP	0.82	0.86	0.88	0.85	0.85	0.9	0.86
ALBP	0.85	0.87	0.9	0.91	0.93	0.96	0.9

#### 4.4. Quantitative comparison with the method of the state of the art:

Table 3 presents a comparison –in terms of AUC value- with different approaches dealing with false positive reduction. In this study, we use the same ratio of masses and the same database.

To conclude, the comparison of the proposed approach with different methods demonstrates the feasibility of ALBP descriptors for false positive reduction in mammographic masses.

Table 3. Comparison of different approaches developed for mass false positive reduction

	ROIs	Ratio	AUC
Oliver et al. [16]	672	1/3	0.9
Oliver et al. [17]	588	1/2	0.83
Sahiner et al. [18]	1792	1/2	0.83
Lladò et al.[12]	1024	1/3	0.91
ALBP+ANN	1160	1/2	0.96

#### 5. Conclusion:

CAD systems have been applied and have got greater value in the recent years as a second reader for medical imaging. It contributes to the reduction of false positive mammographic masses. This work presents a new CAD based on ALBP, which is a final texture feature descriptor. It is used to classify the ROIs to masses and non-masses. The ANN was used as classifier. Different image sizes were applied for improvement of detection rates. Finally, a comparative study with works in the state of the art was done for fairer evaluation.

Such comparison illustrates that the proposed method gives better results (AUC= 0.96).

## 6. Acknowledgments

The authors would like to thank DR. Abid Riadh, Radiologist at El Farabi Imaging center, Sfax, Tunisia, and at the Faculty of Medicine of Sfax for this helpful discussions and advices.

We want also to thank MVG and VGG for sharing their source codes of LBP.

## 7. REFERENCES

- [1] Eurostat, Health statistics Atlas on mortality in the European Union, Office for Official Publications of the European Union; 2002.
- [2] American Cancer Society Breast cancer: facts and figures. ACS; 2003–2004.
- [3] Esteve J, Kricger A, Ferlay J, Parkin D. Facts and figures of cancer in the European Community. In: Tech. rep., International Agency for Research on Cancer. 1993. for young women with breast carcinoma. *Cancer* 2003;97(2):352–8.
- [4] Sickles EA. Breast cancer screening outcomes in women ages 40–49: clinical experience with service screening using modern mammography. *J Natl Cancer Inst: Monogr* 1997;22:99–104.
- [5] De Koning HJ, Fracheboud J, Boer R, Verbeek AL, Collette HJ, Hendriks JHCL. Nation-wide breast cancer screening in the Netherlands: support for breast cancer mortality reduction. National evaluation team for breast cancer screening, *Int J Cancer* 1995;60(6):777–80.
- [6] Bird RE, Wallace TW, Yankaskas BC. Analysis of cancers missed at screening mammography. *Radiology* 1992;184:613–7.
- [7] Basset LW, Gold RH. Breast Cancer Detection: Mammograms and Other Methods in Breast Imaging. New York: Grune & Stratton; 1987.
- [8] Sahiner B, Chan HP, Wei D, Petrick N, Helvie MA, Adler DD. Image feature selection by a genetic algorithm: Application to classification of mass and normal breast tissue. *Med Phys* 1996;23:1671–84.
- [9] Varela C, Tahoces PG, Méndez AJ, Souto M, Vidal JJ. Computerized detection of breast masses in digitized mammograms. *Comput Biol Med* 2007;37(2):214–26.
- [10] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Machine Intell* 2002;24(7):971–87.
- [11] M. A. Turk & A. P. Pentland. “Eigenfaces for recognition.” *J. Cognitive Neurosci.* 3(1), pp. 71–86, 1991.
- [12] Oliver A, Martí J, Martí R, Bosch A, Freixenet J. A new approach to the classification of mammographic masses and normal breast tissue. *IAPR Int Conf Pattern Recognit* 2006;4:707–10.
- [13] Lladó, X., Oliver, A., Freixenet, J., Martí, R., Martí J. “A textural approach for mass false positive reduction in mammography,” *Computerized Medical Imaging and Graphics* 33, 415–422, 2009.
- [14] N. Gargouri Benayed, A. Damak Masmoudi, D. Sellami Masmoudi and R. Abid : A new GLLD operator for mass detection in digital mammograms, *International Journal of Biomedical Imaging*.
- [15] D. E. Rumelhart and G. E. Hinton, in *Learning Internal Representations By Error Propagation, Parallel Distributed Processing*, vol. 1, pp. 318–362, MIT Press, Cambridge, Mass, USA, 1986.
- [16] Heath M, Bowyer K, Kopans D, Moore R, Kegelmeyer PJ. The digital database for screening mammography. *Int Work Dig Mammogr* 2000:212–8.
- [17] Oliver, A., Lladó, X., Martí, J., Martí, R., Freixenet, J.: False positive reduction in breast mass detection using two-dimensional PCA. In: *Lect. Not. in Comp. Sc.*, vol. 4478, pp. 154–161 (2007)
- [18] Sahiner B, Chan HP, Wei D, Petrick N, Helvie MA, Adler DD. Image feature selection by a genetic algorithm: Application to classification of mass and normal breast tissue. *Med Phys* 1996;23:1671–84.