Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Combination of Nonlinear Filters and ANN for Detection of Microcalcifications in Digitized Mammography

J. Quintanilla-Domínguez, M.G. Cortina-Januchs, Aleksandar Jevtić and D. Andina
Technical University of Madrid
Group for Automation in Signals and Communications
Madrid, Spain
{joelq, januchs}@salamanca.ugto.mx
{aleksandar.jevtic, d.andina}@upm.es

J.M. Barrón-Adame and A. Vega-Corona
University of Guanajuato
Laboratorio de Inteligencia Computacional
Guanajuato, Mexico
{badamem, tono}@salamanca.ugto.mx

***Abstract*—Breast cancer is one of the leading causes to women mortality in the world. Cluster of Microcalcifications (MCCs) in mammograms can be an important early sign of breast cancer, the detection is important to prevent and treat the disease. In this paper, we present a novel method for the detection of MCCs in mammograms which consists of image enhancement by histogram adaptive equalization technique, MCCs edge detection by Coordinate Logic Filters (CLF), generation, clustering and labelling of suboptimal features vectors by Self Organizing Map (SOM) Neural Network. The experiment results show that the proposed method can locate MCCs in an efficient way.**

***Index Terms*—Cluster of Microcalcifications, Image Enhancement, Edge Detection, Coordinate Logic Filters, SOM.**

I. Introduction

Breast cancer is one of the most dangerous types of cancer among women around the world. Early detection of breast cancer is essential in reducing life loss. Currently the most effective method for early detection and screening of breast cancers is mammography. However, achieving this early cancer detection is not an easy task. Although the most accurate detection method in medical environment is biopsy, it is an aggressive, invasive procedure that involves some risks, patient's discomfort and high cost. MCCs can be an important early sign of breast cancer, they appear as bright spots of calcium deposits. Individual microcalcification (MC) is sometimes difficult to detect because of the surrounding breast tissue, their variation in shape, orientation, brightness and diameter size [1]. MCCs are potential primary indicators of malignant types of breast cancer, therefore their detection can be important to prevent and treat the disease. But it is still a hard task to detect all the MCCs in mammograms, because of the poor contrast with the tissue that surrounds them. However, many techniques have been proposed to detect the presence of MCCs in mammograms: Image enhancement techniques, Artificial Neural Networks (ANN), Wavelets, Support Vector Machines (SVM), Mathematical Morphology, Bayesian image analysis models, high order statistic, Fuzzy Logic, etc.

Image enhancement algorithms have been utilized for the improvement of contrast features and the suppression of noise. A. Papadopoulos *et al.* [2] proposed five image enhancement algorithms for the detection of MCCs in mammograms. The Contrast-Limited Adaptive Histogram Equalization (CLAHE), the Local Range Modification (LRM), 2-D Redundant Dyadic Wavelet Transform (RDWT), RDWT Linear Stretching (WLST) and Wavelet Shrinkage (WSRK) techniques.

Wavelets have been also employed in MCCs detection providing high spatial frequency features in mammograms. Gholamali Rezai-rad and Sepehr Jamarani [3] present an approach for detecting MCCs in mammograms employing combination of ANN and Wavelet-based subband image decomposition. Y. Sung-Nien *et al.* [4] developed a Computer-Aided Diagnosis (CAD) system for detection of MCs in mammograms. Their work was divided in two stages. First, all suspicious MCs were preserved by thresholding a filtered mammogram via a wavelet filter according to the Mean Pixel Value (MPV) of that image. Secondly, Markov random field parameters based on the Derin-Elliott model were extracted from the neighborhood of every suspicious MCs as the primary texture features. Both Bayes classifier and backpropagation neural network were used for computer experiments. A. Vega-Corona *et al.* [5]. proposed and tested a method to the detect MCs in digital mammography. The method combines selections of Region of Interest (ROI) where MCs were diagnosed, enhancing the image by histogram adaptive techniques, processing by multiscale wavelet and gray level statistical techniques, clustering and labelling of suboptimal feature vectors applying an unsupervised statistical method based on improved K-means algorithm and a neural feature selector based in a GRNN and detector based on a MLP to finally classify the MCs. M. Bhattacharya *et al.* [6]. proposed a method based on discrete wavelet transform due to its multiresolution properties with the goal to segment MCs

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

in digital mammograms. Morphological Tophat algorithm was applied for contrast enhancement of the MCCs. Finally fuzzy c-means clustering (FCM) algorithm was implemented for intensity-based segmentation. L. Song *et al.* [7] applied mathematical morphology and wavelet transform to locate the MCCs in digital mammogram.

G. Veni *et al.* [8] proposed a method based in SUSAN edge detector and adaptive contrast thresholding technique and spatial filters for detection of MCs. L. Wei *et al.* [1] proposed an adaptive classification scheme in the context of SVM learning, which demonstrated to out perform several methods in breast cancer classification.

In this paper, we present a method for detection of MCCs in mammograms. This paper is organized as follows: In section II; model and theoretical background is presented; experimental results are presented in section III; in the last section the conclusions are presented.

II. Model and Theoretical Background

In this section, we will give an overview of the proposed method for detection of MCCs in mammograms. Fig. 1 shows a block diagram of our method. In the first stage, many ROIs were selected from the mammograms, next, in the second stage we apply an image enhancement algorithm to the ROI images by means of Adaptive Histogram Equalization (AHE). The high bright values in the ROI image are enhanced and the low bright values are diminished. In the next block we obtain a binary ROI image by means of K-means clustering algorithm.In the next block, we apply an edge detector involving Coordinate Logic Filters (CLF) to the ROI image. In the next block we build a subotimal feature vector by means of the AHE and by the edges of the MCCs detected by the CLF by pixel as $S_s = \mathbf{x}^{(q_s)} : q_s = 1, \ldots, Q_s$, where $\mathbf{x}^{(q_s)} \in \mathbb{R}^D$ is a $D - dimensional$ vector and Q_s is the number of pixels into ROI image. The feature vector set by pixel in S_s, is then clustered using Self-organizing Map (SOM) ANN to determine two classes. One class represents background and healthy tissue (S_0) and the other one represent MC (S_1). The subsets feature vectors obtained are $S_0 = \{\mathbf{x}^{(q_0)} : q_0 = 1, \ldots, Q_0$ and $Q_0 \leqslant Q_s\}$ and $S_1 = \{\mathbf{x}^{(q_1)} : q_1 = 1, \ldots, Q_1$ and $Q_1 \leqslant Q_s\}$, where $\mathbf{x}^{(q_1)}$, $\mathbf{x}^{(q_0)} \in \mathbb{R}^D$ and $Q_0 + Q_1 = Q_s$.

A. Database

The mammograms used in this work were extracted from the Mammographic Image Analysis Society (MIAS) database which contains 322 digitized mammograms [9]. The images in the database are digitized at 50-micron pixel edge, which are then reduced to 200-micron pixel edge and clipped or padded so that every image has 1024×1024 pixels. The images from this database have detailed information, including the characteristics of background tissue (fatty, fatty-glandular, or dense-glandular), class of abnormality (calcification, masses and speculated masses) and severity of abnormality (benign or malignant).

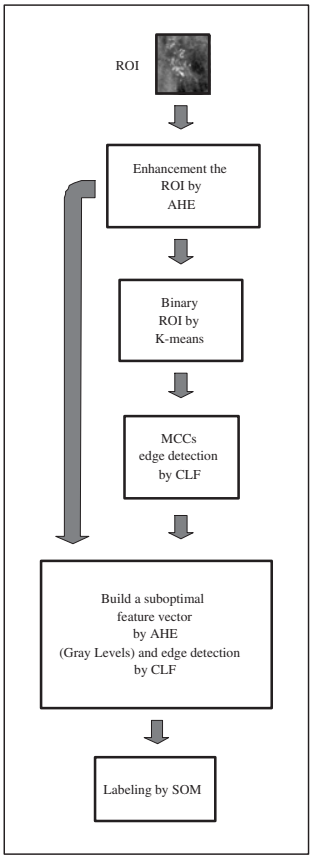

Fig. 1. Block diagram to edge detection of MCCs.

B. Enhancement the Region of Interest

We analyze ROI images because the relevant information of MCCs is concentrated in this area. MCCs are relatively high-frequency components buried in the background of low-frequency components and very high-frequency noise in the mammograms. Image enhancement algorithms have been used for the improvement of contrast features and the suppression of noise. In this work we applied AHE based on the modification of the gray levels of the ROI image with nonlinear transformation function. The AHE, it is a technique widely used and well-established for the medical images enhancement [2] [10]. This technique suppresses pixel values of very small amplitude and enhance only those pixels that are larger than determined threshold within each level of transform space. This is formulated with the following equation:

$$T(G) = \alpha[sigm(k(G - \beta)) - sigm(-k(G + \beta))] \qquad (1)$$

where $G = G(i, j)$ is the gray value of a pixel at (i, j) of the input image and $T(G)$, is the nonlinear transformation function, where α is defined by:

$$\alpha = \frac{1}{sigm(k(1 - \beta)) - sigm(-k(1 + \beta))} \qquad 0 < \beta < 1 \quad (2)$$

and $sigm(x) = \frac{1}{1+e^{-x}}$, $\beta \in \mathbb{R}$ and $k \in \mathbb{N}$ are the threshold control and enhancement rate, respectively.

C. Binary Image by K-means Clustering Algorithm

Previously process, we applied an image enhancement technique with the goal of improve the visibility of MCs by increasing their pixel intensity relative to the background. So, to apply CLF is necessary to convert the ROI image from gray scale to a binary image, to make this conversion we applied a well-established unsupervised statistical method base on K-means clustering algorithm.

D. Feature Extraction

Feature extraction is of key relevance in this work. The features can be calculated from the ROI characteristics such as: the size, shape, density, and smoothness of borders. The feature space is very large and complex due to the wide diversity of the normal tissues and the variety of the abnormalities. Only some of them are significant. Using excessive number of features may degrade the performance of the algorithm and increase the complexity of the classifier. Some redundant features should be removed to improve the performance of the classifier. According to that features are selected, the feature space can be divided into three sub-spaces: intensity features, geometric features, and texture features [11]. In this work, we extract only two kinds of features from the intensity gray levels and geometric features respectively. The first one is obtained after applying the image enhancement by AHE. The other one is obtained by edge detection using CLF.

1) Edge Detection by CLF.: CLF constitute a class of nonlinear digital filters that are based on the execution of Coordinate Logic Operations (CLO). CLF are very efficient in digital signal processing applications, such as: noise removal, magnification, skeletonization, coding, as well as in edge detection, feature extraction, fractal modeling and can also execute the morphological operations (erosion, dilation, opening and closing). The main feature of CLF is the direct execution of CLO between the pixel values without keeping any carry bit [12], [13]. The CLO are the basic logic operations (NOT, AND, OR, and XOR, and their combinations) applied to corresponding individual binary values or pixels found within 2D signals (images). CNOT, CAND, COR, and CXOR represent the coordinate equivalents for each respectively as applied to multi-bit digital data [14]. Given a binary image *B* defined by:

$$B = \{b(i,j); i = 1, 2 \ldots, M, j = 1, 2, \ldots, N\} \tag{3}$$

the evaluation of a CLO (i.e. CXOR) between two images (here: *B1* and *B2*) is performed on a pixel-by-pixel basis and results in the output image *F*:

$$\begin{gathered} F = B_1 \text{ CXOR } B_2 = \{b_1(i,j) \text{ CXOR } b_2(i,j)\} \\ i = 1, 2 \ldots, M \;\; j = 1, 2, \ldots, N \end{gathered} \tag{4}$$

CLF are the application of the CLO to a single image as dictated by a binary structuring element *E*. Since the dimensions of *E* are often much smaller in size than the input image *B*, the resulting output represents local neighborhood characteristics of the image. A configuration for *E*, used in this work as in [14], is shown in (5),

$$E = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{5}$$

Given structuring element *E* from (5) centered on input pixel *b(i,j)* (in image *B*), the output pixel *f(i,j)* (in image *F*) is calculated by:

$$\begin{gathered} f(i,j) = b(i-1,j) \,\mathbf{o}\, b(i,j-1) \,\mathbf{o}\, b(i,j) \\ \mathbf{o}\, b(i+1,j) \,\mathbf{o}\, b(i,j+1) \end{gathered} \tag{6}$$

where **o** represents any of the CLOs. Since the new state of each pixel depends only on the present state of that pixel and those of its neighbors, the new state for each pixel in the filtered image can be computed independently and simultaneously. Mertzios and Tsirikolias have presented the idea of using CLF for the purpose of edge extraction [12]. Danahy *et al.* in [14] introduces an alternative method for calculating CLF using Coordinate Logic Transforms (CLT). Moreover, a new edge detection technique is introduced, enhancing the capabilities of the basic CLT for the application of detecting edges within 2D signals. In this work we applied the edge detector proposed in [12]. Edge detection of MCCs in a ROI image with CLF can be achieved by:

$$F = [(B_E^{\text{CAND}} \text{CXOR } B) - (B_E^{\text{COR}} \text{CXOR } B)] \tag{7}$$

Where F is the final ROI image with edge detection, B is the resulting binary ROI image after applying eq. (1) and an unsupervised K-means clustering algorithm. B_E^{CAND} and B_E^{COR} represent the erosion and dilation of the ROI image B respectively. Using the filter structure of eq. (6), the erosion of the ROI image B is given by:

$$\begin{gathered} f(i,j) = b(i-1,j) \text{ CAND } b(i,j-1) \text{ CAND } b(i,j) \\ \text{CAND } b(i+1,j) \text{ CAND } b(i,j+1) \end{gathered} \tag{8}$$

and the dilation is given by:

$$\begin{gathered} f(i,j) = b(i-1,j) \text{ COR } b(i,j-1) \text{ COR } b(i,j) \\ \text{COR } b(i+1,j) \text{ COR } b(i,j+1) \end{gathered} \tag{9}$$

B_E^{CAND} CXOR B, represents the evaluation of a CLO (CXOR) between two images performed on a pixel-by-pixel.

We build a suboptimal feature vector, with two features. The first one is the gray level intensity from the ROI image enhancement by AHE and the other one is obtained by edge detection using CLF, as follow:

$$S_s = \{\mathbf{x}^{(q_s)} = [x_{T_G}^{(q_s)}, x_F^{(q_s)}] : \; q_s = 1, \cdots, Q_s\} \tag{10}$$

where

$$T_G(i,j) \rightarrow \mathbf{x}_{T_G} = \{x_{T_G}^{(q_s)}\}_{q_s=1,\ldots,n\times m} \tag{11}$$

and

$$F(i,j) \rightarrow \mathbf{x}_F = \{x_F^{(q_s)}\}_{q_s=1,\ldots,n\times m} \tag{12}$$

E. Labeling by SOM

Artificial Intelligence (AI) is the study of how computer systems can simulate intelligent processes [15]. A SOM ANN is a kind of networks without teaching. It has the function of self-organizing. Through training by itself, it can classify the input samples automatically. The idea is that S_s may be clustered in two possible classes, and build two sets (S_0) and (S_1) around of the prototypes of the class centres $\mathbf{z^{(0)}}$ and $\mathbf{z^{(1)}}$. So, we applied a SOM Neural Network to classify the features into S_s in two classes, S_0 and S_1, and defined as $S_{0/1}$ respectively in (13), where,

$$S_s = \{\mathbf{x}^{(q_{0/1})} : q_{0/1} = 1,\ldots,Q_{0/1}, \mathbf{x}^{(q_{0/1})} \in \mathbb{R}^D\} \tag{13}$$

In SOM structure each neuron is connected to input vector through a synaptic weight vector $w_i = [w_{i,1},\ldots,w_{i,m}]$. When an input pattern is presented to the network, the best-matching (winning) neuron v is determined by minimizing the following cost function:

$$v(x^{(q)}) = \min_i \|x^{(q)} - w_i\|, i = 1,\ldots,l \tag{14}$$

where $x^{(q)}$, belongs to m-dimensional input space, $\|\cdot\|$, denoting the Euclidean distance. Then, the synaptic weight vectors are updated as follow:

$$w_i^{(q+i)} = w_i^{(q)} + \eta(q)h_{i,v(x)}(q) \mid x^{(q)} - w_i^{(q)} \mid, \quad i = 1,\ldots,l \tag{15}$$

where $\eta(q)$ and $h_{i,v(x)}(q)$ are the learning rate and neighborhood function centered on the winner, respectively. Although the algorithm is simple, its convergence and accuracy depend on the selection of neighborhood function, the topology of the output space, a scheme for decreasing the learning rate parameter, and the total number of neuronal units [16] [17].

III. Experimental Results

This section presents some preliminary results of our method. To test our method, we selected several ROIs images from mammograms with dense pattern tissue and the presence of MCCs, of size of 256×256 pixels. The ROI images are extracted out of the database with an overlay image previously marked by an expert. ROI images are selected to test our method and shows our results. The Fig. 2 shows, the original ROI images with MCCs.

The visibility of microcalcifications is improved using image enhancement by histogram adaptive equalization technique, the goal is to improve the visibility of MCCs by increasing their pixel intensity relative to the background.

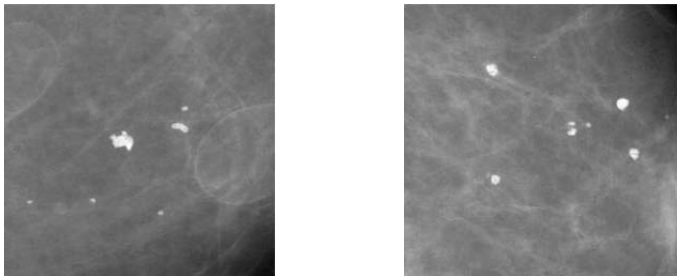

Fig. 2. Original ROI images with MCCs.

The Fig.3 shows, the resulting ROI images after having applied eq. (1) with parameters control of threshold $\beta = 0.85$ and rate of enhancement $k = 15$. This parameters must be predetermined manually and produce good results in image enhancement step.

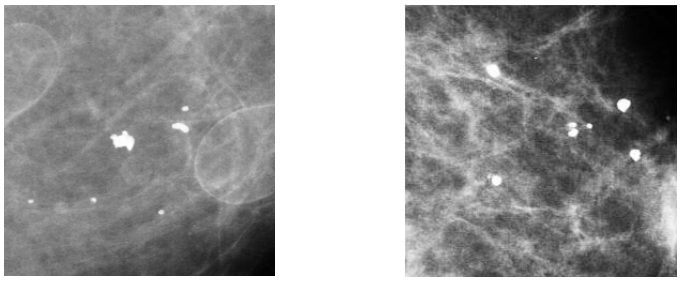

Fig. 3. Enhanced ROI images

The Fig. 4 shows the binary ROI images by K-means algorithm.

Fig. 4. Binary ROI image by K-means

The Fig. 5 shows the results of MCCs edge detection by CLF applying eq. (7). After we built the suboptimal features vector of the set S_s obtained in image processing. We clustered and labeled the feature vectors into set S_s by SOM Neural Network. We obtained the labeled sets S_0 and S_1. We obtained the labeled sets S_0 and S_1 with 65219 and 317 samples respectively for the first ROI image and 65194 and 342 samples for S_0 and S_1 respectively for the second one.

The Fig. 6 shows the results of detection of MCCs on the ROI images using our proposed method. In this work we consider the individual clustering feature vector in each ROI to verify

Fig. 5. MCCs edge detection by CLF

Fig. 6. Detection of MCCs by our method

the result of each partition visually. This consideration allows us to reconstruct the ROIs image from the allocated labels. Due the obtained results, the proposed method is a good alternative in order to detect MCCs, which can be compared with others works [5], [17].

IV. Conclusions

It is well known that mammogram interpretation is a very difficult task even for experienced radiologists. The fundamental enhancement needed in mammograms is an increase in contrast, especially for dense breasts. Contrast between malignant tissue and normal dense tissue may be present on a mammogram but below the threshold of human perception. Our proposed method aims at the improvement of systems performance in the detection of the MCCs in ROI images that could be missed or misinterpreted by medical experts. Edge detection is a fundamental and essential pre-processing step in applications such as image segmentation and computer vision, because edges represent important contour features in the corresponding image. CLF are efficient in image-processing tasks, for example, edge detection. In case of mammograms CLF present a good performance for MCCs edge detection. The edge of MCs is a very important feature to determine the malignancy of MCs. In medical image, ANN have been applied to a variety of data-classification and pattern recognition tasks, such as the differential diagnosis of interstitial diseases and have been shown to be a potentially powerful classification tool. For these reasons, in this paper, we have proposed a novel method for the detection of MCCs using image enhancement, CLF and ANN. Finally, computer simulations demonstrated, that our method can locate MCCs in mammograms satisfactorily and evaluated of visual way.

Acknowledgment

This research has been partially supported by National Ministry Spanish (MICINN) institution under the following project: Technical University of Madrid AL09-P(I+D)-12 , PTFNN (MCINN ref: AGL2006-12689/AGR); Consejo Nacional de Ciencia y Tecnologia (CONACyT) and Laboratorio de Inteligencia Computacional (LABINCO) of Guanajuato University, of Mexico

References

[1] L.Wei, Y.Yang, and R.M.Nishikawa. Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis. *Pattern Recognition*, 42(6):1126 – 1132, 2009.

[2] A. Papadopoulos, D.I. Fotiadis, and L. Costaridou. Improvement of microcalcification cluster detection in mammography utilizing image enhancement techniques. *Computers in Biology and Medicine*, 38(10):1045 – 1055, 2008.

[3] G. Rezai-rad and S. Jamarani. Detecting microcalcification clusters in digital mammograms using combination of wavelet and neural network. In *CGIV '05: Proceedings of the International Conference on Computer Graphics, Imaging and Visualization*, pages 197–201, 2005.

[4] Y. Sung-Nien, L. Kuan-Yuei, and H. Yu-Kun. Detection of microcalcifications in digital mammograms using wavelet filter and markov random field model. *Computerized Medical Imaging and Graphics,*, 30(3):163–173, 4 2006.

[5] A. Vega-Corona, A. Álvarez, and D. Andina. Feature vectors generation for detection of microcalcifications in digitized mammography using neural networks. *Lecture notes in computer science*, 2687:583, 2003.

[6] M. Bhattacharya and A. Das. Fuzzy logic based segmentation of microcalcification in breast using digital mammograms considering multiresolution. *Machine Vision and Image Processing Conference, 2007.*, pages 98–105, 2007.

[7] L. Song, Q. Wang, and J. Gao. Microcalcification detection using combination of wavelet transform and morphology. In *Signal Processing, 2006 8th International Conference on*, volume 4, 2006.

[8] G. Veni, E.E. Regentova, and L. Zhang. Detection of clustered microcalcifications with susan edge detector, adaptive contrast thresholding and spatial filters. *Lecture notes in computer science*, 5112:837–843, 2008.

[9] University of Essex. Mammographic image analysis society. http://peipa.essex.ac.uk/ipa/pix/mias/. 2008.

[10] A.F. Laine, Schuler, S.J. Fan, and W. Huda. Mammographic feature enhancement by multiscale analysis. *Medical Imaging, IEEE Transactions on*, 13(4):725–740, 1994.

[11] H. D. Cheng, Xiaopeng Cai, Xiaowei Chen, Liming Hu, and Xueling Lou. Computer-aided detection and classification of microcalcifications in mammograms: a survey. *Pattern Recognition*, 36(12):2967–2991, December 2003.

[12] B.G. Mertzios and K. Tsirikolias. Applications of coordinate logic filters in image analysis and pattern recognition. *Image and Signal Processing and Analysis, 2001. ISPA 2001. Proceedings of the 2nd International Symposium on*, pages 125–130, 2001.

[13] Y. S. Boutalis, K. Tsirikolias, B. G. Metrtzios, and I. T. Andreadis. Implementation of morphological filters using coordinate logic operations. *Pattern Recognition*, 35(1):187 – 198, 2002.

[14] E. E. Danahy, K. A. Panetta, and S. S. Agaian. Coordinate logic transforms and their use in the detection of edges within binary and grayscale images. *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, 3:III – 53–III – 56, 2007.

[15] D. Andina and P. Duc Truong. *Computational intelligence for engineering and manufacturing*. Springer Verlag, 1st. edition, 2007.

[16] T.Kohonen. The self organization map. *Proceedings. IEEE*, 78(9):1464–1480, September 1990.

[17] A. Vega-Corona, M. Sánchez-García, M González-Romo, J. Quintanilla-Domínguez, and J.M. Barrón-Adame. Contextual and non-contextual features extraction and selection method for microcalcifications detection. In *Proceedings of the World Automation Congress, 2006.*, volume 5, July 2006.