

Automatic Diagnosis of Breast Tissue

Atef Boujelben¹, Hedi Tmar², Mohamed Abid² and Jameleddine Mnif¹

University of Sfax

¹*ANIM "Departement of Radiology and Medical Imaging- Habib Bourguiba Hospital-Sfax. Faculty of Medicine",*

²*CES "National Engineers School of Sfax" Tunisia*

1. Introduction

The Breast cancer whose region is difficult to be visually detected is a major cause of death among women (Nishikawa, 2007). So, the quality of radiologist judgment of whether the suspected region is malignant or benign will not be guaranteed. So far, screening mammography has been the best available radiological technique for an early detection of breast cancer (Siddiqui et al., 2005). However, because of the large number of mammograms to be analysed, radiologists can make false detections. Thus, there are new solutions of automatic detection pertaining to the problems of analysis that can be explored. In this context, Computer Aided Diagnosis (CADi) and Computer Aided Detection (CADe) are two systems that can solve these problems (Rangayyan et al., 2007). In fact, CADe System is based on the detection of Region Of Interest (ROI) and decision. As for, CADi build on good isolation of ROI, analysis and classification to have a decision and/or aid for decision.

This paper proposes a CADi System based on Texture/Shape characterization to reduce the load of radiologists work. In fact, in the past several years the mammography process has seen tremendous evolution. In processing and analysis techniques, many methods based on shape characterization are adopted. So, breast tumours and masses appear in mammograms with different shapes and characteristics. There are two kinds of tumours: malignant ones which usually have rough, microlobulated, or spiculated contours, and benign tumours that have commonly smooth, round, macrolobulated, or oval contours (Reston, 1998). It is true that this type of characterization is efficient and allows good mammogram exploration, but the quality of results is women-old dependent: if the woman is younger, it is too hard to analyse his mammogram. For this reason, we include a texture description of the region to cope with this problem. So, the density of the region can discriminate the malignity or benignity of ROI by analysing the texture. Thus, the technique adopted in this work takes into consideration both texture and shape characterisation. In general, the quality of analysis is dependent on the quality of segmentation. However, current approaches do not guarantee a good quality of segmentation. The majority of segmentation-methods take only edge (inter-area) aspects into account to delimit the ROI. In this context, the manual segmentation and semi-automatic method are widely used. In (Boujelben et al., 2009a) (Boujelben et al., 2009b), the threshold-based segmentation is carried out by fixing a rectangular box around the suspicious tumour area and then using Sobel filter in order to avoid noise. However, this can affect noise and discontinuities in the border of ROI. In addition, active contours,

depending on edge criterion, solve the problems of segmentation as noise and discontinuities (Osher & Fedkiw, 2002). However, breast quality makes segmentation effective only by taking both intra-area and inter-area aspects into account. To attain our objective which is the ROI segmentation in mammographic images, we apply the Level Set method based on external function (convergence function) that represents area and contour criteria as much as possible. In this paper, we include the texture/shape detection in the process of mammograms diagnosis. The main purpose of this work is the elaboration of a CADi to reach a good identification of ROI and contribute to a better quality of analysis. This work, is integrate within the MIPAX (Medical Image Processing and Analysis eXchange) project which was defined as the object of CES (Computer, Electronic And Smart engineering systems design Laboratory in National School of Engineers of Sfax) and ANIM(Numeric Archiving and Medical Imaging in National School of Medicine) collaboration. So, this project was split in three parts; Numeric Archiving PACS (Picture Archiving and Communication System), Data Base of User Environment and Automatic Analysis of Medical Images. This present work articulate around the last part. To attain our objective (CADi), we firstly show why and how to adapt Level Set-based approach in case of pseudo-detection, which is a semi-automatic detection by using level-set technique; and secondly, we study the performance of boundary, region and texture features in a mammogram diagnosis process. The remainder of this paper is organised as follows. Section 2 presents the state of the art of shape/texture analysis; without loss of generality, we outline the most original and important work addressing mammogram analysis. Section 3 describes the proposed block diagram for mass diagnosis. Section 4 illustrates the deformable model, namely, Level Set approach adopted in segmentation and its adaptation in case of breast cancer detection. Afterwards, section 5 presents the adopted method for analysis and shows how a combination of shape and texture features could be advantageous for a good diagnosis. As for section 6, it presents the results obtained by the proposed scheme. Lastly, section 7 gives some concluding remarks and draws some future work.

2. Context of state of the art

The medical imaging is an active domain that embraces various topics like image processing, mass segmentation or detection, mass analysis and decision or aid for decision. The results of our research can be viewed in the context of two areas of related work: the detection of breast cancers, and the analysis of detected breast cancers. The purpose of this paper is to examine how to differentiate the malignant tumours from the benign ones. So, the analysis steps are related to pseudo-detection results. In this context, we attempt to present an adaptation of Level Set technique for pseudo-detection and investigate two approaches for shape and texture analysis. Therefore, the analysis of texture is used to qualify the density of ROI or to have an idea about the space distribution of micro-calcification (Dheeba & Wiselin, 2010) (Wiesmiller & Chandy, 2010)(Boujelben et al., 2011). Generally, texture feature extraction methods can be classified into three major categories; namely statistical, structural and spectral. In a biomedical image like mammogram, the characteristics of the pixels in the texture pattern are not similar everywhere. To cope with this specificity, statistical approaches for texture analysis such as the moments of gray-level histogram, based on a Gray-Level Co-occurrence Matrix (GLCM), is used. It is computed to discriminate different textures in mammographic images (Oliver et al., 2007) (Zwiggelaar & .R.Denton, 2004) (Lambrou et al., 2002) (Masala et al., 2007) (Ahirwar & Jadon, 2011). In this context, Zwiggelaar et al.(Zwiggelaar & .R.Denton, 2004) include some mathematic operators like translation and transportation in order to select a sub-set of features from GLCM to have a

decision about tumour characterisation. Quite close from ours, Lambrou *et al.* (Lambrou *et al.*, 2002) studied the effectiveness of GLCM and the higher-order-statistic based on twenty features. In (Masala *et al.*, 2007), ROI is characterised by eight features extracted from GLCM. Thereafter, four classifiers were evaluated: Multilayer Perception (MLP), Probabilistic Neural Network (PNN), Radial Basis Function Network (RBF) and k-Nearest Neighbours (KNN). In opposition to (Masala *et al.*, 2007), we evaluate the effectiveness of two classifiers: MLP and KNN classifiers. Using statistical approaches, we extract six characteristics from Co-occurrence matrix. Therefore, we compute the average value of each characteristic over each orientation(0, 45, 90 and 135). The second criterion in medical image analysis process is the analysis of shape which is built over two phases, namely boundary and region analysis. In case of boundary, many work focused on the Radial Distance Measure (RDM)(Rangayyan *et al.*, 2006)Boujelben *et al.* (2009a)(Alvarenga *et al.*, 2006), Convexity (CVX), Fourier Fraction (FF) (Rangayyan *et al.*, 1997), Fractal Dimension (FD) (Nguyen & Rangayyan, 2005) (Nguyen & Rangayyan, 2006) and the angular measure (Yang *et al.*, 2005) (Denise *et al.*, 2008) (Rangayyan *et al.*, 2006). However, methods defined in the context of angular measures provides so far either of the two categories: Radial Angle (RA) (Yang *et al.*, 2005) or Turning Angle(TA) (Denise *et al.*, 2008) (Rangayyan *et al.*, 2006). In this context, Sheng Chih *et al.* (Yang *et al.*, 2005) used the RA, which is the smallest angle included between the gradient direction and the radial direction of the edge. The RA forms a good feature for a malignant/benign discrimination. In fact, if the angle increases towards 180 degree, then it is a benign mass. In contrast, if the angle decreases towards zero degree, then it is a malignant mass. Nevertheless, the computation of RA takes many times. His temporel complexity increases because a large number of points in the perimeter of the region are taken into account through the calculus. Unlike RA, TA tackle the problem of temporel complexity. However, the calculus is limited to the convex points that forms the perimeter of the region(Rangayyan *et al.*, 2006) with preserving the result quality. Quite close from (Rangayyan *et al.*, 2006), we propose a novel measure denoted Index Angle (IA)(Boujelben *et al.*, 2009b). Its based on the external and internal angle concepts. We will show in section 5 the technical details of the IA calculus. We will also show how the IA can be efficient to differentiate malignant mass from benign ones. On the other hand, the RDM descriptor (Alvarenga *et al.*, 2006) (Delogu *et al.*, 2008) was taken a great importance in medical imaging litterature. It is based on the computation of the distances between contour points and gravity center of the region. However, it provides a complete knowledge concerning circularity of the region. The RDM technique presents any advantages like normalized computation and insensibility to affine transformations. From the RDM, Alvarenga *et al.*(Alvarenga *et al.*, 2006) and Delogu *et al.* (Delogu *et al.*, 2008) extracted many features like Roughness (R), Standard DEViation (SDEV), etc. They combined the RDM and the region features to improve mass description. In (Alvarenga *et al.*, 2006), the performance and relevance of a set of shape features extracted from the RDM method and the Convex-Hull are evaluated. In contrast, Delogu *et al.* (Delogu *et al.*, 2008) evaluated the combination of some features extracted from RDM and others like Convexity and Circularit. The RDM is a method that can differentiate between the malignant and benign cases, but, it can cause noise in the calculation of each boundary point. Furthermore, it can also cause a long time of calculation. To deal with these problems, we propose to calculate the features only in the concave and convex points. In fact, the extended RDM that we propose, XRDM (Boujelben *et al.*, 2009a), is shown in section 5. Like contour descriptor explained above, region descriptor was taken a great importance in image description. It is used to describe the regularity of the mammogram masses. However, simple morphologic features like Circularit and Eccentricity are extracted through this descriptor (Yang *et al.*, 2005)(Boujelben *et al.*, 2009b)(Delogu *et al.*, 2008). Alvarenga *et al.* (Alvarenga *et al.*, 2006) evaluated the

performance and relevance of seven shape features; namely Perimeter (P), Normalized Radial Length (NRL), SDEV, Area Ratio (AR), contour roughness (R), Circularity and Mshape. These characteristics are enriched by adding some other ones in (Retico et al., 2007). The most important added characteristics are Zero Crossing (ZC) (i.e. a count of the number of times the radial distance plot crosses the average radial distance) and Convexity, which allows the representation of the studied shape better than the characteristics cited above. In this chapter, we present an approach of shape analysis in our diagnosis process of mammograms. From region criteria we use characteristics like Circularity, which can be useful in this direction and can give an indication on the regularity of a given mammogram mass, Internal/External Circle (IEC), which can be used to measure the elongation of shape, and NRV. Added to that, the features combination based on shape and texture is used in the CADi systems. In (Maglogiannis et al., 2007), Maglogiannis et al. proposed an intelligent system for automatic breast cancer diagnosis using Support-Vector-Machines based classifiers (SVM). The features used are based on texture and shape criteria like radius (means of distances from centre to points on the perimeter), SDEV of grey-scale values, Perimeter, Area, Smoothness (local variation in radius lengths), Compactness, Concavity (severity of concave portions of the contour), concaves points (number of concave portions of the contour), symmetry and FD. Furthermore, 22 features based on edge-sharpness, shape and texture are extracted by Nandi et al. (Nandi et al., 2006). They adopted Genetic Programming (GP) for features classification. This method handles implicit feature selection. The GP is also used in (Zadeh et al., 2001) to compare the performance of four different texture and shape feature extraction methods which are conventional shape quantifiers, co-occurrence-based method of Haralick, wavelet transformations and multi-wavelet transformations. Zadeh et al. (Zadeh et al., 2004) began again their work done in (Zadeh et al., 2001) by considering 17 shape and 44 texture features. They selected the best feature using Genetic Algorithm (GA).

In summary, the medical imaging literature is split so far either of two orientations: many methods independently process on contour, region and texture indexes while others attempt a combination of all these indexes and a selection of the more important features with GA. Yet, it is clear that the characteristic of masses has information based on these properties (region, contour and texture). The subject matter is that the combination of the characteristics of different properties can lead or not to a good quality of diagnosis.

However, the identification of breast region is important to improve the analysis process. So, breast tumours and masses appear in mammograms with different shape characteristics. Detecting the region can give an idea about the nature of diagnosis. However, in the past several years there has been tremendous evolution in mammography process. In this context, two approaches are used in the literature: automatic detection and region segmentation. Concerning detection, Torrent et al. (Torrent et al., 2008) presents a comparison of two clustering based algorithms and one region based algorithm for segmenting fatty and dense tissue in mammographic images. The first algorithm is a multiple thresholding algorithm based on the excess entropy, the second one is based on the Fuzzy C-Means clustering algorithm, and the third one is based on a statistical analysis of the breast. In addition, method based on multiresolution approach to the computer aided detection of clustered micro-calcifications in digitized mammograms based on Gabor elementary functions is illustrated in (Catanzariti et al., 2003). So, a bank of Gabor functions with varying spatial extent and tuned to different spatial frequencies is used for the extraction of micro-calcifications characteristics. Firstly, results show that most micro-calcifications, isolated or clustered, are detected and secondly the classification is illustrated by an Artificial Neural Network with supervised learning. On the other hand, Thangavel et al. (Thangavel

& Karnan, 2005) present an Ant Colony Optimization (ACO) and Genetic Algorithm (GA) for the identification of suspicious regions in mammograms. The proposed method uses the asymmetry principle (bilateral subtraction): strong structural asymmetries between the corresponding regions in the left and right breasts are taken as evidence for the possible presence of micro-calcifications in that region. Bilateral subtraction is achieved in two steps. First, the mammogram images are enhanced using median filter, then pectoral muscle region is removed and the border of the mammogram is detected for both left and right images from the binary image. Further GA is applied to enhance the detected border. So, the nipple position is identified for both left and right images using GA and ACO, and their performance is studied. Second, using the border points and nipple position as the reference of mammogram images are aligned and subtracted to extract the suspicious region. In the context of detection ROI, Schiabel *et al.* (Schiabel *et al.*, 2008) proposed a methodology based on the Watershed transformation, which is combined with two other procedures; histogram equalization, working as pre-processing for enhancing images contrast, and a labelling procedure intended to reduce noise. But, Jadhav *et al.* (Jadhav & Thorat, 2009) used statistical feature extraction method by using a sliding window analysis, for detecting circumscribed masses in mammograms. This procedure is implemented by taking into account the multi-scale statistical properties of the breast tissue, and succeeds in finding the exact tumour position by performing the mammographic analysis using first few moments of each window. We have demonstrated that fast implementation in both feature extraction and neural classification module can be achieved. Nevertheless, a system processes for the mammograms in several steps is adopted in (Arodz *et al.*, 2006). First, we filter the original picture with a filter that is sensitive to micro-calcification contrast shape. Then, authors enhance the mammogram contrast by using wavelet-based sharpening algorithm. Afterwards, present to radiologist, for visual analysis, such a contrast-enhanced mammogram with suggested positions of micro-calcification clusters. However, a multi-resolution representation of the original mammogram is obtained using a linear phase non-separable 2-D wavelet transform which is adopted in (Liu & Delp, 1997). This is chosen for two reasons. First, it does not introduce phase distortions in the decomposed images. Second, no bias is introduced in the horizontal and vertical directions as a separable transform would. Authors used coefficients of the analysis low pass filter. A set of features are then extracted at each resolution for every pixel. Detection is performed from the coarsest resolution using binary tree classifiers. This top-down approach requires less computation by starting with the least amount of data and propagating detection results to finer resolutions. In addition, wavelet coefficients describe the local geometry of an image in terms of scale and orientation apart from being flexible and robust with respect to image resolution and quality (Oliver *et al.*, 2007). In addition, Marti *et al.* (Marti *et al.*, 2003) propose a supervised method for the segmentation of masses in mammographic images. Based on the active region approach, an energy function which integrates texture, contour and shape information is defined. Then, pixels are aggregated or eliminated to the region by optimizing this function allowing the obtention of an accurate segmentation. The algorithm starts with a selected pixel inside the mass, which has been manually selected by an expert radiologist. Recently, explicit and implicit methods of deformable model are used in different applications (Brox *et al.*, 2009). In this context, for breast cancer detection, Ferrari *et al.* (Ferrari *et al.*, 2004) used a traditional active deformable contour model (Snake) to limit the breast in the image. To injure the problem of initialisation, they used an adaptative thresholding. For the elimination of the pectoral muscle, Boucher *et al.* (Boucher *et al.*, 2009) used the snake and Ball *et al.* (Ball & Bruce, 2007) used the Narrow Band level set methodology with an adaptative segmentation threshold controlled by a border complexity term. An overview of the literature shows that many methods of

segmentation and identification are used to detect ROI. In this paper, we propose method based on Level Set approach which includes edge and region proprieties. So, in the Level Set approach, two major problems are usually discussed in the bibliographies: initialisation and evolution function which is the point of interest in section 4. In the next section, we describe the proposed block diagram for mass diagnosis.

3. Overview

The proposed approach consists of three subtasks. Firstly, the identification of ROI is done using a level-set-based approach which includes edge and region criteria. Secondly, features are extracted, using shape/texture descriptors. Finally, in order to take decision about diagnosis, we are interested in MLP and KNN classifiers. Figure 1 shows the bloc diagram of the proposed scheme. The first point of this workflow is the segmentation which will be detailed in details in next section.

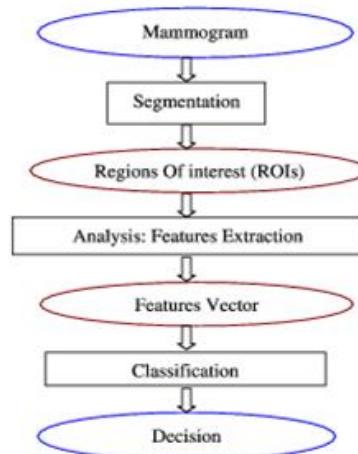


Fig. 1. Proposed Flow

4. Segmentation with deformable models

Image Segmentation is an important step to handle a good analysis. However, it is based on homogeneity and/or edge criteria of the region. It consists of ROI extraction. So, the choice of a segmentation technique depends on regularity/irregularity of the edge of the ROI. In addition, the noise can affect the segmentation quality. However, we are interested in noise avoiding. Generally, to tackle this problem in medical imaging, explicit and implicit methods of deformable model are used.

4.1 Explicit deformable models

The active contour model, or snake, is used to detect region of interest (ROI) or contour in image. It is an energy-minimising-spline technique. The result of this minimisation is guided by two terms; the first term controls the aspect of the curve: it is often called internal energy. The second term attracts the curve C towards object which one seeks the borders: it is often

called external energy. The detail of this method is illustrated in (Kass et al., 1988). The main concepts are:

- The snake is parametrically defined as:

$$v(s) = (x(s), y(s)), \text{ with } s \in [0, 1]$$

- The energy is defined by:

$$E = \int_0^1 \alpha \|V(s)\| + \beta \|V'(s)\|^2 ds - \lambda \int_0^1 (\|\nabla I(V(s))\|^2) ds \quad (1)$$

Where:

- α, β and λ are real constants, respectively coefficients of elasticity, rigidity and contraction (or dilation) from the curve.
- $\nabla I(V(s))$: is the gradient of image in s.

4.2 Implicit deformable models

Contrary to the explicit models, the implicit deformable models or Level Set approach (Casselles et al., 1997) uses a dense contour in which the implicit evolution avoids the needs to track surface markers in relation to each other. So, the Level Set is a method which studies the evolution of the curve and surfaces (Osher & Fedkiw, 2002). The points defining this interface will move towards the normal at a speed F according to the following equation:

$$\frac{\partial C(t)}{\partial t} = F \vec{N} \quad (2)$$

\vec{N} : Normal with the curve.

F: speed term depends on the curve.

The parametric curve C(t) is improved by the detection of the level zero and the function F evolves and moves according to:

$$\frac{\partial \phi}{\partial t} = F \|\nabla \Phi\| \quad (3)$$

The Evolution of this function depends on an initial curve Φ_0 . In this case, there are two aspects of research which are the initialization and the function F. The former is adopted in the present paper with a special focus on the region to be detected. In general (Sethian, 1998), speed F depends on three terms: first, on the local curve in each point (pondered with ϵ), second, on the term which is dependent on the image (pondered with β), and third, on a constant term (pondered with ν). The evolution of the interface is given by the following equation:

$$\frac{\partial \phi}{\partial t} = \epsilon * g(I) \|\nabla \Phi\| * \text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - \beta * \nabla g(I) \|\nabla \Phi\| + \nu * g(I) \|\nabla \Phi\| \quad (4)$$

where: I is the point (i,j) of image matrix

$$\epsilon, \beta, \nu \in [0, 1]$$

$$g(I) = \frac{1}{1 + \|\nabla \Phi\|}$$

To minimize the temporal complexity of this equation, we adopt the Narrow Band and Fast Marching method in the implementation algorithm of Level Set. Narrow band consists of computing Level Set on evolution from contour for early inside and outside near the Level Set zero (Osher & Fedkiw, 2002). The reasons for using this approach are twofold. The first reason is to optimize time computation efficiency for a numerical calculation Level set method. The second one is the fact that, in general, regions in breast are difficult to be detected. In fact, we should focus locally near to the zero Level Set and its neighbouring Level Set because the local contour has more information significance than distant ones. Still, to accelerate the convergence of Level Set approach, we adopt a monotonically advancing front based on Fast Marching approach (Osher & Fedkiw, 2002). Its idea is that if $T(x,y)$ is the time at which the curve crosses the point (x,y) then the surface $T(x,y)$ satisfies the equation:

$$\|\nabla T\| \cdot F = 1 \quad (5)$$

where:

$$F = \frac{1}{\exp(-\alpha \nabla I)} \quad (6)$$

This equation allows us to make a good implementation of deformable contours. Indeed, the changes of topology are automatically managed. Thus, if the image contains several objects, the contour is divided during its evolution including each object separately. Contour can also become deformed in order to be adjusted with complex forms, which cannot do explicit active contours (Snakes). Another positive point is that this method does not depend on initialization. Nevertheless, in the case of textured images, the criteria of gradient (edge properties) on which depends this equation (uniformity inter-region) affected a over-segmentation. So, the presence of textures in a mammographic image generates bad results because the small areas are privileged. But one can resort to a measurement of containing area in order to improve the quality of detection.

4.3 Adaptation of a Level Set approach

To solve the problem of mammographic images, which also depends on the information of the intra-region included in the information of inter-region, we added the criterion of the area. So, the region property is adopted firstly with the notion of image and secondly with the notion of propagation (addition of a fourth term). The evolution of the interface which has indeed ameliorated eq 4 is given by the following equation:

$$\frac{\partial C(t)}{\partial t} = \epsilon * g(I) \|\nabla \Phi\| * \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} - \beta * (\nabla g(I) \|\nabla \Phi\| + \frac{Moy(I)}{Max(I)}) \right) - \nu * g(I) \|\nabla \Phi\| - \Theta * SkewCN(I) \quad (7)$$

where: $\Theta \in [0,1]$

$\operatorname{Max}(I)$ =maximum of gray-level in image.

$\operatorname{Moy}(I)$ =average of gray-level 3×3 centred in (x,y)

$\operatorname{SkewCN}(I) = \frac{\operatorname{SkewnessCentred}}{\operatorname{Max}(\operatorname{Skewness})}$

The SkewnessCentred corresponds to the moment around the average. It measures the deviation of the distribution of the gray-level compared to a symmetrical distribution. For a deviation towards raised values, the Skewness-Centred is positive; whereas for a deviation towards low values, it is negative. Then SkewCentred can be calculated as follows:

$$\operatorname{SkewnessCentred} = \frac{1}{9} \sum_x \sum_y (I(x,y) - Moy(x,y)) \quad (8)$$

Figure 2 show the result of segmentation in the proposed scheme. In the first line, the first

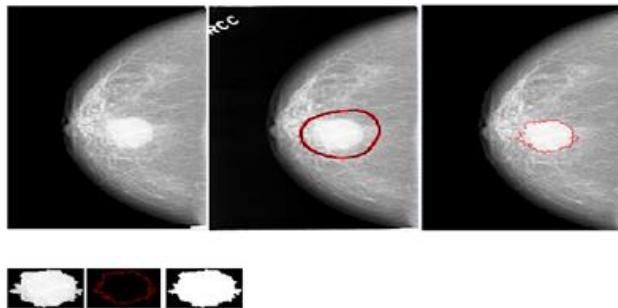


Fig. 2. ROI segmentation with mammogram; The first line showing from left to right: 1- the original image without contour, 2- the original image in DDSM database detoured by an experimented radiologic (red line), 3-the image segmented in this approach (red line). The second line showing the ROI isolated region in three cases: texture, boundary and area.

image is the original image without contour, the second image is the original image in DDSM database detoured by a radiologic (red line), and the third image is segmented in the context of this approach (red line) and in the second line, in every raw, we have region isolated ROI in the original texture, boundary and region, respectively. These three sub-images will be like entries for the vector which is based on texture, boundary and region. Likewise, it is clear that our results about segmentation are quite close to the manual segmentation results obtained by the radiologist. So, what remains is to see the quality of the segmentation that can be proven by the analysis step. Moreover, breast tumours and masses appear in mammograms with different shape characteristics. In this context, the method can reflect the irregularity or regularity of region and more precisely if compared to the manual process. Hence, the performance is illustrated according to two standpoints: precision of ROI segmentation in diagnostic relevance and computation time of optimization. After the ROI segmentation, the extraction of features is adopted in ROI: this is the point of interest that we will focus on, in the next section of the paper.

5. Analysis: Features extraction

In the ROI segmentation, we use an adaptation of a Level Set Approach with an edge and a region criterion. In this section, the features extraction is illustrated on any ROI. The mass have different shape characteristics, particularly, in breast cancer. In this framework, we can introduce a method based on shape analysis, basically, boundary analysis. Figure 3 shows the overall shape of benign and malignant mass. Firstly, we start with boundary information.

5.1 Boundary descriptor

Boundary analysis is often referred to in order to help to define regions according to any criteria. To differentiate between microlobulated region from macrolobulated ones, we can measure the convexity which is based on boundary. In fact, Retico *et al.* (Retico *et al.*, 2007) used CVX in which feature depends on region (ratio of region-air detected by perimeter/or Air for his Convex envelop). In this subsection, we attempt to improve the importance of the

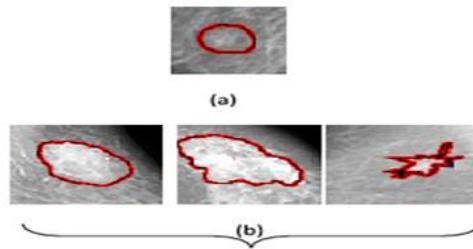


Fig. 3. Types of ROI: (a) benign, (b) malignant

analysis of boundary shape. Besides, CVX is the ratio of Convex envelop by perimeter for its perimeter of the region detected. When the mass tends to be round, its CVX tend to be near the 1. Conversely, a mass with speculated edge will have a CVX smaller than 0.5. Then, the CVX can be calculated as follows:

$$CVX = \frac{Perimeter(\text{ConvexEnvelop})}{Perimeter(\text{region})} \quad (9)$$

The advantage of this feature is that it is standardized and it is invariant with any affine transformation. After that we can differentiate between the macrolobulated and microlobulated region. Subsequently, to describe the rough ones, any researchers use RDM (Boujelben et al., 2009a) and Turning angle (Denise et al., 2008) (Rangayyan et al., 2006). In fact, RDM is a well-known method used in shape analysis. However, Euclidian distances $d(i)$ are calculated between the gravity centre in the region and all the points in boundary region (Figure 4). where:

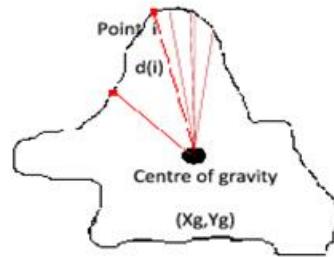


Fig. 4. Illustrative figure of RDM

$$d(i) = \sqrt{(X_i - X_g)^2 + (Y_i - Y_g)^2} \quad (10)$$

To eliminate large calculations from characteristics, all radial distances were normalized by using the maximum value of the radial distances.

$$dn(i) = \frac{d(i)}{\max[d(i)]} \quad (11)$$

where n : is the number of points (pixels) of the region boundary (the perimeter of region).

$$Xg = \frac{\sum_N X}{N} \quad (12)$$

$$Yg = \frac{\sum_N Y}{N} \quad (13)$$

where N : is the number of points (pixels) in the region area.

$$dmoy = \frac{1}{n} \sum_N dn(i) \quad (14)$$

The features extracted in the RDM are cited below. We will only give the expressions of the RDM features related to this work.

- The Standard Deviation of the Normalized Radial Distance Measure (SDEV) is defined as the variance of the distances around the ray (the average distance $dmoy$ previously defined) of a circle. This characteristic gives a good quality of information concerning the irregularity of contour. Indeed, when it is about a malignant tumour, the value of SDEV tends to move towards 0.5, and when it is in the case of benign tumour, the SDEV tends to move towards 0.

$$SDEV = \sqrt{\frac{1}{N} \sum_N (dn(i) - dmoy)^2} \quad (15)$$

- Rugosity (R): treats angular contours (contours which contain concave segments). It is given by the following equation:

$$R = \frac{1}{N} \sum_N \|dn(i) - dn(i+1)\| \quad (16)$$

- Area ratio (Ar): this characteristic differentiates between stellar contours and smooth contours. It is illustrated in the following equation:

$$Ar = \frac{1}{N * dmoy} * \sum_N (dn(i) - dmoy) \quad (17)$$

where $Ar=0$, if($dn(i) \leq dmoy$)

In practice, the computation of these features increases the complexity of calculation. To deal with the problem of complexity, we propose to calculate the features only in the concave and convex points. The RDM is a method that can differentiate between the malignant and benign cases, but, it can cause noise in the computation of each point of boundary. Furthermore, it can also cause a long-time calculation. To solve the problem of noise and computing time, one can improve the method of RDM. This can be done by implementing the idea of the eXtended Radial Distance Measure (for more details see (Boujelben et al., 2009a)). In fact, the eXtended RDM "XRDM" is adopted. So, to solve the problem of complexity, we propose to calculate the features only in the local concave and convex points as in Figure 5. These points are defined as follows:

- The concave point (Pconcave (i)) of the contour is a point which have a radial distance $d(i)$ lower than the radial distance $d(i-1)$ and the radial distance $d(i+1)$.

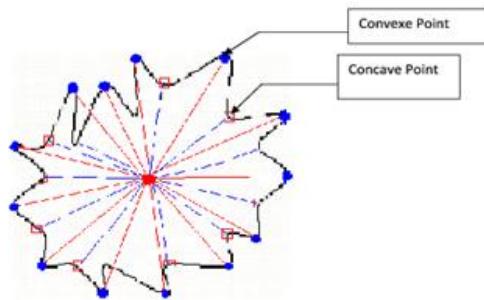


Fig. 5. Illustrative figure of eXtended RDM

- The convex point ($P_{\text{convexe}}(i)$) of the contour is a point whose radial distance $d(i)$ is higher than the radial distance $d(i-1)$ and the radial distance $d(i+1)$.

More formally:

$$P_{\text{concave}}(i) = (i; d(i) \leq d(i-1) \text{ et } d(i) \leq d(i+1))$$

$$P_{\text{convexe}}(i) = (i; d(i) \geq d(i-1) \text{ et } d(i) \geq d(i+1))$$

In the speculated region, the feature like turning angle cannot be applied because of the problem of segment tangents (Figure 5). However, RDM and XRDM can be used with regions that have an elliptic shape. In fact the major problem is when there is a speculated region.

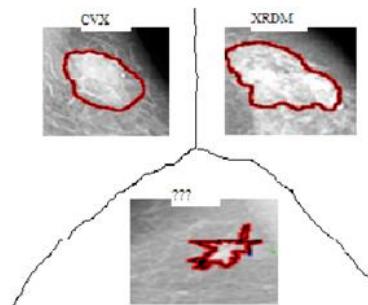


Fig. 6. Problem space convergence in malignant boundary

To solve this problem, we introduce a new angular characteristic named Index Angle (IA).

- The IA is the ratio of all the internal angles by external ones: the external angle is the angle between central convex points (Convex point p_i) and their "next-least" convexes points (Convex point p_{i-1}, p_{i+1}). In fact, the internal angle is the angle between a central convex point and its "next-least" concave points (Figure 7). The IA is applied to make a distinction between the edge shapes of the mass as being speculated or as being round. When the mass tends to be round, its IA tends to be near the 1. In opposition, a mass with speculated edge will have an IA smaller than 0.5. Then, the IA can be calculated as follows:

$$IA = \frac{\sum_i \phi(i)}{\sum_i \theta(i)} \quad (18)$$

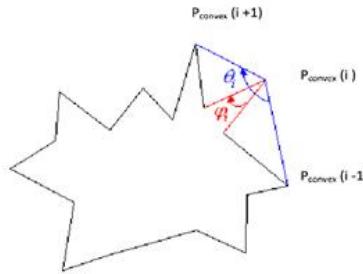


Fig. 7. An example of Index Angle computation

So, the IA is used only in the concave and convex points and not to any other points. Our objective is to minimize the temporal complexity differently from the radial angle used in (Rangayyan et al., 2006). Added to that, the advantage of this characteristic is that it is standardized and invariant with any affine transformation. In this subsection, we discussed boundary information of opacity (boundary vector) composing on CVX, XRDM and IA, in the next subsection, we will put our interest on the region criteria.

5.2 Region descriptor

We use Region Features to describe the mammographic masses through features extracted region. For this reason, we illustrate a method based on Circularity (C), Internal/External Circle (IEC) and Normalized Residual Value (NRV)

- Circularity (C): it describes the areas that can be circular. It can be useful in this direction and can give an indication of the regularity of a given mammogram mass. This feature is given by the following equation:

$$C = \frac{4 * \pi * Aire}{Perimeter * Perimeter} \quad (19)$$

where: P is the perimeter and A is the area of the segmented mass.

- Internal External Circle (IEC): This feature can be used to measure the shape elongation used by Chettaoui et al. (Chettaoui et al., 2005). In our work, we exploit this feature to describe mass region. The IEC feature is given by the following equation:

$$IEC = \frac{Inf - Radius}{Sup - Radius} \quad (20)$$

where: Sup-Radius represents the largest internal circle and Inf-Radius represents the smallest external circle (Figure 8).

For a round mass, the value of IEC is close to 1 since the value of Inf-Radius is very close to the value of Sup-Radius, whereas for a lengthened mass the value of IEC becomes close to 0 since the value of Inf-Radius is far from the value of Sup-Radius.

The advantage of this characteristic is that it is invariant with affine transformation and it is adequate to our work. In fact, its calculation is slow, since for each form, we should pass through all the points to determine the circle inscribed in the object which contains this point.

- Normalized Residual Value (NRV): This feature is extracted from the convex-hull by using the residual region. NRV gives best performances compared to the characteristics that can

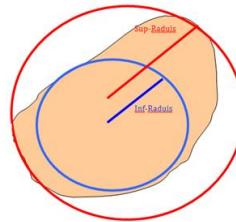


Fig. 8. Illustrative figure of internal/external circle computation

be extracted from the convex-hull, and can be useful in the distinction between the regular and irregular area. It is given by the following equation:

$$NRV = \frac{Aire(Residual - Region)}{Perimeter(Convex - Envelop)} \quad (21)$$

where: Perimeter is the perimeter of the convex-hull and Aire is the area of the residual region(Figure 9).

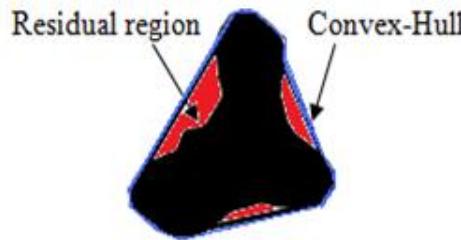


Fig. 9. Example of breast tumour and its respective Convex-Hull

In the two last subsections, we focused on shape features which are boundary and region. In the next subsection we will put emphasis on texture criteria.

5.3 Texture descriptor

The density (Figure 10) of breast region is an important property in ROI. To determine this quality, we adopt the texture method. In approach to texture feature extraction, which is frequently cited in the literature, is based on the use of CGLM. Co-occurrence matrix is a second-order statistical measure of image variation. In this subsection, we detail the feature of co-occurrence approach.



Fig. 10. The density of ROI opacity

We represent our analysis by texture statistical. From this approach, we extract six characteristics which are defined as follows:

$$mean = \frac{1}{N^2} \sum_x \sum_y p(x, y) \quad (22)$$

where: $p(x, y)$ denotes the gray-level in the co-occurrences matrix.

N : denote the width and the height of co-occurrences matrix. In order to reduce calcul of co-occurrence characteristics we adopted the original matrix of region, where Gray Level value are between 0 and maximum of Gray Level value, to adopt a region with size 32X32(i.e. Gray Level value between 0 and 31)

$$Variance = \sum_x \sum_y (x - moy)^2 p(x, y) \quad (23)$$

$$Energy = \sum_x \sum_y p(x, y)^2 \quad (24)$$

$$Contrast = \sum_x \sum_y (x - y)^2 p(x, y) \quad (25)$$

$$Entropy = - \sum_x \sum_y p(x, y) \log p(x, y) \quad (26)$$

$$Homogeneity = \sum_x \sum_y \frac{1}{1 + (x - y)^2} p(x, y) \quad (27)$$

The algorithm evaluates the properties of the region of the mammographic image. We investigate the performance of feature in texture from GLCM in diagnosis by using four orientations 0, 45, 90, 135. From each one, we inspect six features (then we take the average of one feature of the four orientations). In the next section, we will show the performance of the textural and shape vector in analyzing ROI in terms of diagnosis relevance by using kNN and MLP classifier.

6. Results and discussion

The terminology which is used to determine the performance of a CADi System is defined as follows:

- Sensitivity: percentage of pathological ROIs which is correctly classified.
- Specificity: percentage of non-pathological ROIs which is correctly classified.
- Accuracy: percentage of correctly classified pathological and non-pathological cases.

Because of the variation in the types of breast cancer, a large number of cases can reduce the dependency of analysis techniques versus image sets. The performance of an algorithm is affected by the characteristics of a database like the digitization techniques which are namely pixel size, subtlety of cases, choice of training / testing subsets, etc.

6.1 The DDSM dataset

The establishment of the DDSM allows the possibility of the common training and testing Dataset. The DDSM is the largest publicly available database of mammographic data. It

contains approximately 2620 screening mammography cases. From the total number of mammographic images included in the DDSM database, we use 200 malignant images and 200 benign ones. To make a good evaluation, we use the remaining 400 images which are divided into 200 ground malignant regions together with 200 entirely benign ones. To classify the area segmented with Level Set Approach using DDSM dataset with vector illustrated in the least section, one will use two classifiers which are kNN (K=7) and MLP. In the next subsection, we illustrate the results of sensibility and specificity adopted with an analysis method.

6.2 Experimental results: Performance in terms of diagnosis quality

The basic classification is based on two methods of classification KNN (K=7) and MLP as shown in Table 1, Table 2 and Table 3. It represents the results from different analysis in boundary, region and texture vector respectively.

classifier	Kppv	MLP
Sensitivity	87 %	92%
Specificity	88%	90%
Exactitude	88%	91%

Table 1. Results from analysis based on boundary description vector

classifier	Kppv	MLP
Sensitivity	92%	89%
Specificity	89%	90%
Exactitude	90%	89%

Table 2. Results from analysis based on region description vector

classifier	Kppv	MLP
Sensitivity	93%	90%
Specificity	87%	87%
Exactitude	89%	89%

Table 3. Results from analysis based on texture description vector

The sensitivity result varies between 87% in boundary vector and 93% in texture feature using KNN (K=7) classifier. But the result of specificity varies between 87% using KNN classifier in texture vector and 90% in both MLP classifier region and boundary classifier. These results seem to be variable because of a variation of result: 6% in pertaining to sensibility and 3% relating to specificity. In fact, breast tumour sometimes depends on the region, or/and contour or/and texture criteria. To solve this problem, we propose to combine features vectors. The results are shown on Table 4, Table 5 and Table 6.

classifier	Kppv	MLP
Sensitivity	90%	89%
Specificity	86%	89%
Exactitude	88%	89%

Table 4. Results from the analysis based on shape (boundary and region) description vector

classifier	Kppv	MLP
Sensitivity	92%	87%
Specificity	92%	88%
Exactitude	92%	88%

Table 5. Results from analysis based on boundary-texture description vector

classifier	Kppv	MLP
Sensitivity	89%	89%
Specificity	93%	93%
Exactitude	91%	91%

Table 6. Results from the analysis based on region-texture description vector

Table 4, Table 5 and Table 6 show the results of combination of region/boundary features, region/texture feature and boundary/texture feature, respectively. Table 4 illustrates the importance of shape information in this analysis. The result in terms of sensitivity tends to move towards 90% in KNN classifier. The result in terms of specificity tends to move towards 89% using MLP classifier. So, we can assume that shape vector is a good feature in differentiating the benign from the malignant mass. In boundary texture-combination, we find a good result in terms of sensibility and specificity by using KNN classifier differently from MLP classifier. Table 6 shows the result of features region-texture. In this approach we can assume good specificity using both MLP and KNN classifiers differently from sensibility. So, the majority of classifications that are obtained are favourable, but, the problem is in the stability of result in terms of sensibility, specificity and accuracy. This is conducted by the combination of all features in Table 7. However, all result are about 90% of sensibility, specificity and accuracy. This result seems to be logical: ROIs in mammographic image depends on taking account of region, boundary and texture properties.

classifier	Kppv	MLP
Sensitivity	90%	89%
Specificity	90%	90%
Exactitude	90%	89%

Table 7. Results from the analysis based on texture-shape description vector

These results are not the best result compared to local works (Boujelben et al., 2009a) (Boujelben et al., 2009b). However, in (Boujelben et al., 2009b) the result is about 94% in boundary information; and in (Boujelben et al., 2009a) the result is between 90% and 92% using XRDM method. In such work, we used DDSM database but the ROI is selected from the image by fixing a rectangular box around the suspicious lesion area and the classical method of segmentation based on Sobel filter and thresholding. But, in this approach, we note the stability of result which is an important quality in analysis of medical imaging. Comparing these results with other related work, we notice important ameliorations. In fact, Alvarenga et al. (Alvarenga et al., 2006) obtained 88% of sensitivity and 90% of specificity. In their experiments, they used a local images dataset and LDA (Linear Discriminant Analysis) method for classification. Additionally, Rangayyan et al. (Rangayyan et al., 1997) have used the LDA classifier and their result reaches 95% in terms of classification accuracy. Conversely, the result of Retico et al. (Retico et al., 2007) using a MLP classifier can reach 78.1% and 79.1% in sensitivity and specificity, respectively. Using a SVM classifier, Chang et al. (Chang et al., 2005) obtained 88.89% and 92.5% in terms of sensitivity and specificity, respectively. Yet,

the characterization of mammographic masses and tumours and their classification as being benign or malignant is difficult. In spite of acceptable results found by our proposed features, we should not make an assumption that it is the best or the worst because we did not use the same Database. In fact, the digitization can reflect the final result. However, we can assume that by combining feature vector based on shape/texture, we attempt to have good stability of results in differentiating between the benign and the malignant masses.

7. Conclusion

In this work, we attempted to improve the classification performance of shape and texture in analyzing ROI in case of mammographic images. We introduced the adaptation of Level Set approach to detect ROI by combining edge and region criteria. We also presented cooperation of features to have stability of results because the ROI depends on both shape and texture properties. The results in terms of sensitivity and specificity tend to reach 90%. The results have been validated via two algorithms of classification: kNN (K=7) and MLP. These results seem to be sufficient and an automatic method of detection based on a Level Set Approach can ameliorate the CADi system to have a CADe one.

8. References

Ahirwar, A. & Jadon, R. (2011). Characterization of tumor region using som and neuro fuzzy techniques in digital mammography, *International Journal of Computer Science and Information Technology* 3(1): 199–211.

Alvarenga, A., Fernando, A. & Albuquerque, C. W. C. (2006). Assessing the performance of the normalised radial length and convex polygons in distinguishing breast tumours on ultrasound images, *Pereira Revista Brasileira de Engenharia Biomédica* 22: 181–189.

Arodz, T., Kurdziel, M., Popiela, T. J., Sevre, E. & Yuen, D. (2006). Detection of clustered microcalcifications in small field digital mammography, *computer methods and programs in biomedicine* 81: 56–65.

Ball, J. & Bruce, L. (2007). Digital mammogram spiculated mass detection and spicule segmentation using level set, *Conference of the IEEE EMBS* pp. 23–26.

Boucher, A., Jouve, P., F.Cloppet & Vincent, N. (2009). Pectoral muscle segmentation on a mammogram, *ORASIS* .

Boujelben, A., Chaabani, A., Tmar, H. & Abid, M. (2009a). An approach based on rdm for analysis in breast cancer detection, *International Conference on Applied Informatics*, Bordj Bou Arerig Algeria.

Boujelben, A., Chaabani, A., Tmar, H. & Abid, M. (2009b). Feature extraction from contours shape for tumor analyzing in mammographic, *IEEE Digital Image Computing: Techniques and Applications*, Melbourne, pp. 395–399.

Boujelben, A., Tmar, H., Mnif, J. & Abid, M. (2011). Automatic application level set approach in detection calcifications in mammographic image, *International Journal of Computer Science and Information Technology* 3(4): 1–14.

Brox, T., Rousson, M., Deriche, R. & Weickert, J. (2009). Colour, texture, and motion in level set based segmentation and tracking, *Image and Vision Computing: Elsevier* .

Casselles, V., Kimmel, R. & Sapiro, G. (1997). Geodesic active contours, *International Journal of Computer Vision* (22): 61–79.

Catanzariti, E., Ciminello, M. & Prevete, R. (2003). Computer aided detection of clustered microcalcifications in digitized mammograms using gabor functions, *IEEE Image Analysis and Processing* pp. 266–270.

Chang, R., W.Wu, Moon, W. & D.R.Chen (2005). Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors, *Springer* 89: 179–185.

Chettaoui, C., Djamel, K., Djouak, A. & Maaref, H. (2005). Etude de formes des globules drépanocytaires par traitement numérique des images, *International Conference SETIT Tunisia* .

Delogu, P., Fantacci, M., Kasae, P. & Retico, A. (2008). Characterization of mammographic masses using a gradient-based segmentation algorithm and a neural classifier, *Computer Biologie Medecal* .

Denise, G., Rangayyan, R., Carvalho, D. & Santiago, S. (2008). Polygonal modeling of contours of breast tumors with the preservation of spicules, *IEEE Transations In Biomedical Engineering* 55.

Dheeba, J. & Wiselin, G. (2010). Detection of microcalcification clusters in mammograms using neural network, *International Journal of Advanced Science and Technology* 19.

Ferrari, R., Frere, A., R.M.Rangayyan, desautels, J. & Borges, R. (2004). Identification of the breast boundary in mammograms using active contour models, *Medical Biological Engineering and Computing* 42.

Jadhav, R. & Thorat, R. (2009). Computer aided breast cancer analysis and detection using statistical features and neural networks, *International Conference on Advances in Computing, Communication and Control* (23-24): 283–290.

Kass, M., Witkin, A. & Terzopoulos, D. (1988). Snakes: Active contour models, *International Journal of Computer Vision*, 1: 321–332.

Lambrou, T., Linney, A. D., Speller, R. D. & Todd-Pokropek, A. (2002). Statistical classification of digital mammograms using features from the spatial and wavelet domains, *Medical Image Understanding and Analysis* 3.

Liu, S. & Delp, E. (1997). Multiresolution detection of stellate lesions in mammograms, *IEEE International Conference on Image Processing* pp. 109–112.

Maglogiannis, I., Zafiroopoulos, E. & Anagnostopoulos, I. (2007). An intelligent system for automated breast cancer diagnosis and prognosis using svm based classifiers, *Springer Science Business* .

Marti, J., Freixenet, J., Munoz, X. & Arnau, A. (2003). Active region segmentation of mammographic masses based on texture contour and shape features, *A Pattern Recognition and Image Analysis Springer* 2652: 478–485.

Masala, G., Tangaro, S., Golosio, B., Oliva, P., Stumbo, S., Bellotti, R., Carlo, F. D., Gargano, G., Cascio, D., Fauci, F., Magro, R., Raso, G., Bottigli, U., Chincarini, A., Mitri, I. D., Nunzio, G. D., Gori, I., Retico, A., Cerello, P., Cheran, S. C., Fulcheri, C. & Torres, E. L. (2007). Comparative study of feature classification methods for mass lesion recognition in digitized mammograms, *Maggio-Giugno nuovo cimento* 30(3).

Nandi, R., Nandi, A., Rangayyan, R. & Scutt, D. (2006). Classification of breast masses in mammograms using genetic programming and feature selection, *International Federation for Medical and Biological Engineering* .

Nguyen, T. & Rangayyan, R. (2005). Fractal analysis of contours of mammographic masses, *International Conference on Biomedical Engineering Innsbruck Austria* pp. 186–191.

Nguyen, T. & Rangayyan, R. (2006). Shape analysis of breast masses in mammograms via the fractal dimension, *In Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society Shanghai China* pp. 3210–3213.

Nishikawa, R. (2007). Current status and future directions of computer-aided diagnosis in mammography, *Computerized Medical Imaging and Graphics* 31: 224–235.

Oliver, A., Llado, X., Marti, R., Freixenet, J. & Zwigelaar, R. (2007). Classifying mammograms using texture information, *Medical Image Understanding and Analysis*, Aberystwyth and Wales and UK, pp. 223–227.

Osher, S. & Fedkiw, R. (2002). Level set methods and dynamic implicit surfaces, *Springer-Verlag New York, Applied Mathematical Sciences*.

Rangayyan, R., Ayres, F. & Desautels, J. (2007). A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs, *Journal of the Franklin Institute* 344(3-4): 312–348.

Rangayyan, R., Faramawy, N., Desautels, J. & Alim, O. (1997). Measures of acutance and shape for classification of breast tumors, *IEEE Transactions on Medical Imaging* 16: 799–810.

Rangayyan, R., Guliano, D., Carvalho, J. & Santiago, S. (2006). Feature extraction from the turning angle function for the classification of contours of breast tumors, *IEEE-The International Special Topics Conference on Information Technology in Biomedicine* pp. 1–6.

Reston, V. (1998). Illustrated breast imaging reporting and data system (bi-radstm), *American College of Radiology* 3.

Retico, A., Delogu, P., Fantacci, M. & Kasaei, P. (2007). An automatic system to discriminate malignant from benign massive lesions on mammograms, *Medical Physics* 14: 596–600.

Schiabel, H., Santos, V. & Angel, M. (2008). Segmentation technique for detecting suspect masses in dense breast digitized images as a tool for mammography cad schemes, *ACM Special Interest Group on Applied Computing* 20(16): 1333–1337.

Sethian, J. (1998). Adaptive fast marching and level set methods for propagating interfaces, *Cambridge Univ Press* pp. 3–15.

Siddiqui, M., Anand, M., mehrotra, P., Sarangi, R. & Muthur, N. (2005). Biomonitoring of organochlorines in women with benign an malignant breast disease, *Technical report*, 250-257, Environmental Research.

Thangavel, K. & Karnan, M. (2005). Computer aided diagnosis in digital mammograms: Detection of micro-calcifications by meta heuristic algorithms, *GVIP Journal* 7(5): 41–55.

Torrent, A., Oliver, A., Freixenet, J., Boada, I., Feixes, M., Marti, R., Llado, X., Pont, J., Perez, E., Pedraza, S., Marti, J., & Krupinski, E. (2008). Breast density segmenattion: A comparaison of clustering and region based techniques, *Springer-Verlag Berlin Heidelberg IWDM* pp. 9–16.

Wiesmiller, S. & Chandy, D. A. (2010). Content based mammogram retrieval using gray level aura matrix, *International Joint Journal Conference on Engineering and Technology* pp. 217–221.

Yang, S. C., Wang, C., Chung, Y., Hsu, G., Lee, S., Chung, P. & Chang, C. (2005). A computer-aided system for mass detection and classification in digitized mammograms, *In Biomed Eng Appl Basis Comm* pp. 215–228.

Zadeh, H., Radc, F. & Nejad, S. (2001). Shape-based and texture-based feature extraction for classification of microcalcification in mammograms, *Medical Imaging: Image Processing* pp. 301–310.

Zadeh, H., Radc, F. & S.P.Nejad (2004). Comparison of multiwavelet, wavelet, haralick, and shape features for micro-calcification classification in mammograms, *Pattern Recognition* pp. 1973–1986.

Zwigelaar, R. & R.Denton, E. (2004). Optimal segmentation of mammographic images, *International Workshop on Digital Mammography*, Norwich, UK.