Abstract: This paper proposes an algorithm for mass and micro-calcification detection by manual thresholding and prewitt detector. This algorithm has been tested using mammography images of different densities from multiple databases of a health clinic and images taken from the internet (40 images in total). The results are very accurate, allowing better detection of breast pathologies (mass and micro-calcification). Finally, the detection of breast pathologies was performed using as input a detection algorithm specially designed for this purpose. After segmentation by manual thresholding, morphological opening, morphological dilatation and Prewitt contour detection we have a demarcation of the masses and breast micro-calcification. The results obtained show the robustness of the proposed manual thresholding method. In order to evaluate the efficiency of our pathology detector, we compared our results with those in the literature and performed a qualitative evaluation with a rate of 98.04% for the detection of breast pathologies. A radiologist from the health clinic evaluated the results and considers them acceptable to the CAD.

Keywords: manual thresholding; qualitative evaluation; detecting pathologies; morphological opening; demarcation.

1- Introduction

Breast density, a measure of the extent of radio-dense fibrous glandular tissue in the breast, has the potential to be used as a predictor of breast cancer risk, it is a measure of how much tissue can be seen on mammography [1]. Some tissues, such as the mammary gland, are dense and appear white on a mammogram. This density makes it difficult for doctors to see tumors, which also appear white. Adipose tissue is less dense and appears clear on mammography, allowing better detection of tumors. In 1976, Wolfe published an article showing a relationship between breast density and breast cancer risk. He showed that women with dense breasts were four to six times more likely to develop breast cancer [2]. Dense tissue in more than 50% of the breast could account for about one-third of breast cancers [3]. Mammographic density has also been associated with breast cancer tumor characteristics, including tumor size, lymph node status, and lymphatic or vascular invasion in screened cancers [4]. A three-fold increased risk of second breast cancer has also been observed in women diagnosed in situ with ductal carcinoma who have very dense breasts [5, 6].

Mammography poses major problems because of the high breast density that obscures the mammographic image. The main disadvantage of mammography today is that it is difficult to differentiate between normal, dense tissue and cancerous tissue when searching for small tumors surrounded by glandular tissue. A woman's breasts are naturally denser, or more glandular when they are young, making it difficult for the radiologist to analyze the mammographic image.
Breast cancer detection technology is evolving rapidly, with newcomers in the field such as digital mammography and computer-assisted detection. Image enhancement by manipulating fine intensity differences using image processing algorithms is the basis of any computer-aided detection system. For these reasons, our objective in this paper is to apply a combination of methods (pre-processing, segmentation, edge detection, etc.) to digital mammography images in order to detect breast pathologies (masses and micro-calcifications) on mammographic images.

1.1 Paper contour

In this article, we present in section 2 a review of the literature on associated computational methods applied to breast pathologies. Section 3 describes the presentation of the method and protocol used as well as the database used in this article. The results and discussions are described in section 4. The comparison of the results with the state of the art is described in section 5 and finally the conclusion/perspective is presented in the following section.

2- State of the art

In Diagnostic Assistance Systems (DAS), the segmentation of breast masses and micro-calcifications are an important and delicate task since the subsequent description, classification and resetting treatments are strictly linked to the segmentation result. In this section we will focus on the detection of masses and micro-calcifications.

2.1- Mass detection

A good detection of the contour of the lesion produces a description faithful to its characteristics. This ensures a classification that minimizes the rate of false positives and maximizes the rate of true negatives.

However, it has been shown that detection of masses is more difficult than detection of cystic fibrosis [7]. Indeed, it is difficult to distinguish masses from normal regions due to their low contrast and ambiguous edges partially masked by tissue.

Threshold methods have been used extensively for the segmentation of breast masses [8, 9]. Kom et al. [8] detected suspicious masses from mammographic images by adaptive thresholding. The proposed algorithm was tested on a database of 61 digital mammography images on which masses had been previously marked by experienced radiologists. The results show that the proposed method has a sensitivity of 95.91% for mass detection. Kurt B et al. [9], proposes a new segmentation algorithm to identify candidate mass regions in mammograms. The proposed system consists of three parts: breast region and pectoral muscle segmentation, image enhancement and identification of suspicious mass regions. In this study, we focused on the identification of suspicious mass regions using a combination of the Havrda and Charvat entropy method and the Otsu thresholding method. An open access database of the Mammographic Image Analysis Society (MAM), which contains 59 masses, was used for the study. The proposed system achieved a 93% sensitivity rate for the identification of suspicious mass regions in 56 abnormal and 40 normal images.
Mudigonda et al. [10] used multi-level thresholding to detect closed contours. The major disadvantage of this approach is the fact that the masses are considered to have a uniform density with respect to the background of the image, which is not always verified [11].

Another work in the same context is [12] where the authors performed a two-stage adaptive thresholding (DuSAT). A global thresholding which deals with the analysis of the histogram peaks (HPA) of the whole image, the threshold is obtained by maximizing the proposed thresholding criterion. Then a local thresholding is performed for each pixel in a defined neighborhood window to provide accurate segmentation results.

Very recently, Anitha and her team [13] proceeded in the same way as Kai et al. [12]. Other methods have been proposed based on wavelet transformations to improve the contrast of mammographic images [14] before applying an adaptive thresholding technique. To extract the tumor region Elmoufidi et al. in [15] used Local Binary Patterns (LBPs) which compare the luminance level of a pixel with the levels of its neighbors. This gives a texture information. They [15] present a mass detection method using a linear region operator in combination with wavelet processing from mammography images. The method includes the removal of artifacts and the elimination of pectoral muscles according to morphological operations. Finally, mass segmentation for detection using an adaptive threshold technique is performed to separate the mass from the background. Testing them on 45 images from the MIAS database and 85 DDSM images gives a result rate of 91.1%.

In the case of contour approaches, Lu et al. in [16] used active contours for mass detection. First, they proceeded with a pre-processing step to remove artifacts from the film and improve image contrast. Then they used the circular Hough transform to detect the contour of the mass. This serves as the initial contour to start the active contouring process. The algorithms were tested on 17 digitized mammograms.

With respect to cooperative segmentation, several researchers agree that this type of approach offers a promising avenue of research. It contributes to a better taking into account of the characteristics of the entities of the mammography image, thus enhancing the quality and reliability of segmentation.

This line of thought has been adopted by [17], whose approximate EHL results have been used to initialize level sets.

However, in the region growth approach, the most critical step is the selection of the starting points, which is usually done manually, so human interaction is introduced [18].

To conclude, we observe in this part of the review that the work has been conducted at the direction of mass detection, they show that [12, 13, 14] use adaptive thresholding while [15] use a combination of a linear operator per part in combination with a wavelet processing for mass detection. In this section the references use respective MIAS or MAM databases while we use in our paper a local and internet database. These methods will allow us to discuss our methodology as well as the results we will have obtained in this study.
2.2- Detection of micro-calcifications

The detection of micro-calcifications are a precursor sign of the appearance of mammary pathologies, and this is why many scientists study it like [18, 19].

Abdul Malek et al. [18] have combined the growth of seed-based regions and contour segmentation in a sequential order. The first process of growth of the region consists in identifying an initial starting point. Most region growth methods manually identify the seed point that involves human interaction, and is thus considered as the initial starting point (seed) for our region growth algorithm. The contour segmentation method is implemented in order to improve segmentation results. The method is tested on 50 mammography images confirmed by a radiologist as micro-calcifications. Experimental results show that the algorithm successfully segments the micro-calcifications with a precision of 0.94.

Stojic et al. [19], when they developed a system that acts on the contrast enhancement of the images to detect MCS, using two approaches. The first method is based on multi-fractal theory, and the second on mathematical morphology. In their morphological approach, they thought in a logical way: if we are going to subtract from the image enhanced by a white top hat, the image enhanced by a black top hat, the luminous details will be strongly accentuated. Moreover, if we add an original image to this difference a high-pass filter is made. Consequently, to better enhance light details and equalize the uneven background, the above mentioned process can be iteratively repeated: the final segmentation of the Mcs is obtained by the thresholding applied to the output image after several iterations. The contours of the segmented light objects can be extracted and superimposed on the original, indicating the Mcs.

To conclude, it is observed in this section that the work has been carried out in the direction of the detection of micro-calcifications, an early sign for the detection of mammary pathologies.

They show that the work of [18] uses the region growth method while [19] uses a combination of methods, namely contrast enhancement, multi-fractal and mathematical morphology for the detection of Mcs. These methods will allow us to discuss our methodology as well as the results we will have obtained in this study.

3- Presentation of the method

Following is the flowchart of the proposed approach:
Acquiring Original (mammography)

Median filter

Manual Threshold

Morphological opening

Morphological dilatation

contour detection

**Figure 1**: General structure of the algorithm for detection of masses and Mcs

**Step 1**: Acquisition of digital mammograms

Approximately 40 pairs of conventional lateral and craniocausal mammograms were selected as part of the screening program for the detection of breast pathologies currently underway in a health clinic and on images taken on the Internet. Each pair of mammograms being successive to contain a single malignant mass that the radiologist recommends to biopsy in the left or right mammogram. The mammograms were digitized with a commercially available Senographe T800 laser scanner. After scanning, each digital mammogram was subsampled to obtain an image of 600 * 887 pixels. These reduced versions were used in the initial stages of the method for calculation purposes.

**Step 2: median filters**

The images in our database are likely to contain noise as they were obtained from several sources (internet and a health clinic). Reason why we perform a pre-treatment. Moreover we use the median filter because it eliminates the noise while the average filter widens the noise to the rest of the image.

The median filter [20, 21] is a non-linear digital filter, often used to combine both effective suppression of impulse noise and preservation of sufficiently important details in the image. It works by replacing the value of one pixel with the median value of all the pixels in its neighborhood. This filter also removes isolated pixels. Thus we can observe the results of the filtering in figures 2 b), 3 b), 4 b) and 5 b).

**Step 3: Manual thresholding**

After contrast enhancement of the filtered mammograms to bring out all the clear details, candidate regions that may contain Mcs or masses are isolated from the rest of the background by a simple manual thresholding operation.

The implementation of manual image thresholding is very simple, but the problem lies in the choice of threshold. Depending on the threshold value, all pixels below the threshold are
classified as background and the recall pixels are horizontal or vice versa [22]. Thus we can observe the results of manual thresholding in figures 2 c), 3 c), 4 c) and 5 c).

**Step 4: morphological opening operator**

Parasitic details generated by the thresholding step are eliminated using a morphological opening operator [20, 23] with a circular structuring element of size 3. The effects of the morphological aperture are:

- The disappearance of small particles (whose size is smaller than the size of the structuring element)
- To separate coarse particles where they are finer.
Thus we can observe the results of the morphological opening operator in figures 2 d), 3 d), 4 d) and 5 d).

**Stage 5: Morphological dilatation**

Let X be a figure, namely a set of pixels. For a structuring element B, the expansion of X by B is the set obtained by replacing each pixel p of X by its window Bp:

\[ \text{Dil}_B (X) = \bigcup \{Bp \mid p \in X\} \]

Similarly, by applying DилB (X) to the whole image we notice that the homogeneous regions of the expanded image also remain unchanged, but when the input pixel is an object (pixel to 1), it remains unchanged regardless of its neighbors. It is rather the zero pixels that change to 1 when their neighborhood contains at least one 1-pixel. It is said that this pixel expands [20, 23] to increase the size of the object. Thus we can observe the results of the morphological dilation operator in figures 2 e), 3 e), 4 e) and 5 e). In addition, for all the images in this article we use a structuring element of size 3.

**Step 6: Prewitt detection**

The Sobel and Prewitt detectors produce thicker and therefore less localized contours than Roberts' operator, but are more resistant to noise. These two operators give identical results in terms of thickness and contour localization, but differ slightly in terms of noise resistance (different low-pass filtering). The Prewitt operator [20, 22] is used, especially in edge detection.

The operator calculates the gradient of the image intensity at each point, giving the direction of the largest possible increase of light to dark and the rate of change in this direction. The result thus shows how "abruptly" or "smoothly" the image changes at that moment - there, and thus the probability is that part of the image represents an edge, and how that edge is likely to be oriented. Thus we can see the results of Prewitt's Detection in figures 2 f), 3 f), 4 f) and 5 f).

**4- Results and discussion**
In the example shown in the figures below, we applied the process described in Figure 1. Mammography techniques were applied to the images in our database of the parts of the light region of interest and then we determined the threshold values. It is worth mentioning that the radiologists at the health center diagnosed two images in Figure 2a), and Figure 3a) as probable breast cancers. In Figure 3a) they recommend in particular that a biopsy be performed to confirm the diagnosis.

The results of the segmentation procedure were analyzed according to the protocol described in section 3. A total of 40 images was analyzed; three examples of results are shown in Figures 2 to 4. It should be noted, from the example in Figure 2, that the procedure works well and gives satisfactory results. It even allows the detection of masses and micro-calcifications also to identify the pectoral muscle region.

**Figure 2**: Results of the detection of masses and Mcs. (a) Original image (b) image after median filter (c) image after manual thresholding operator (d) image after application of the morphological opening operator (e) image after application of the morphological dilatation operator (f) image after Prewitt’s operator

**Figure 3**: Results of the detection of masses and Mcs. (a) Original image (b) image after median filter (c) image after manual thresholding operator (d) image after application of the morphological opening operator (e) image after application of the morphological dilatation operator (f) image after Prewitt’s operator
Figure 4: Results of the detection of masses and Mcs. (a) Original image (b) image after median filter (c) image after manual thresholding operator (d) image after application of the morphological opening operator (e) image after application of the morphological dilatation operator (f) image after Prewitt's operator

Figure 2a) 3a) 4a) shows the mammography images to be used for various transformations as shown in Figure 1. The application of the Median filter on all images is used here to reduce noise while maintaining the contours of the image, it also removes isolated pixels. It is used here with a filter width =7 and the percentage of the retained pixel is equal to 50%, we can observe the result on figures 2 b), 3 b), and 4 b).

Figures 2c), 3c), and 4c) present the manual thresholding operation which consists in extracting a region from an original image with a threshold $S=154$ for figure 2c), a threshold $S=128$ for figure 3c), and a threshold $S=139$ for figure 4c) with an RGB channel. The binary morphological opening operation eliminates parasitic details generated by the thresholding step with a size or disk of 3 and dark outer edge. We can observe the results of the morphological opening on figures 2 d), 3 d), and 4 d).

After the morphological opening operation, we apply the binary morphological dilation operation which consists in moving the structuring element on each pixel of the image, and to see if the structuring element touches the structure of interest generated by the thresholding step with a size or disk of 3 and dark outer edge. We can observe the results of the morphological dilatation on figures 2 e), 3 e), and 4 e). Finally the Prewitt operator is used for edge and contour detection. We can observe the results of the Prewitt detector on figures 2f), 3f), and 4f). It highlights masses and micro-calcifications. We illustrate it on figure 5 presented below.

Figure 5: applications of the prewitt operator on thresholding with identification of masses, Mcs and pectoral muscle.

Comparison of the results with the state of the art

The segmentation of the masses, Mcs by our process defined in figure 1 on a set of 40 images (health clinic, internet). According to the result obtained, the algorithm succeeds in identifying exactly the foci of the Mcs, masses on the digital mammography images.

To demonstrate the robustness of the algorithm, it has been tested on mammograms with different breast tissue densities. Overall, for the mammograms evaluated, the mean values of
the quality measurement for breast pathology detection were 0.98, meaning that the algorithm appears extremely robust with regard to density types.

After evaluating our manual segmentation approach, we want to evaluate it again but against the literature cited in (section 2) and summarized in the table (Table 1).

<table>
<thead>
<tr>
<th>Name author</th>
<th>Method used</th>
<th>Number of images</th>
<th>Accuracy, sensitivity</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kom et al.</td>
<td>thresholding</td>
<td>61 clinical images</td>
<td>95.91%(sensitivity)</td>
<td>Mass detection</td>
</tr>
<tr>
<td>Abdul Malek et al.</td>
<td>Region growth+ contour segmentation</td>
<td>50 clinical images</td>
<td>94%(accuracy)</td>
<td>Mcs detection</td>
</tr>
<tr>
<td><strong>Our approach</strong></td>
<td>Manual thresholding + mathematical morphology</td>
<td>40 (internet + health clinic)</td>
<td>0.98 for masses and 0.95 for my Mcs (qualitative)</td>
<td>Detection of masses, Mcs</td>
</tr>
</tbody>
</table>

**Table 1** Comparison of the results obtained by our approach with those of the literature

**Conclusion/Perspectives**

In this paper, our objective was to apply a combination of methods described in Figure 1 to digital mammography images, with the aim of allowing the detection of masses and micro-calcifications on digital mammography images to help radiologists detect different breast pathologies (masses and Mcs).

Thanks to the low-pass filter (median filter), artifacts have been eliminated on the original image. A manual thresholding is used to highlight areas presenting masses, pectoral muscle, breast and Mcs. Morphological opening and dilatation is used to eliminate parasitic details generated by the thresholding step and finally a Prewitt detection method is used to detect the contours of the masses and micro-calcifications.

This contribution is very useful in the early detection of breast cancer. It then gives a great opportunity to patients to treat breast cancer at an early stage.

Our experiences have shown that the results obtained are satisfactory, especially since we work on clinically acquired images (with some taken from the internet), which are therefore particularly complex. We are convinced that the proposed approach, in addition to its originality, is adequate to the problem. Nevertheless, various aspects of this work undoubtedly deserve to be deepened. Among others, taking into account three-dimensional information can only improve the result. The resolution of the problem in 2D was imposed to us by the medical team of the clinic as well as images obtained on the Internet, in order to use this approach in an image search system. It is intended for the detection of various pathologies and for the training of trainees. The results were deemed conclusive by the medical team. We plan to test our algorithm (see improvement) on all breast pathologies to help detect different pathologies and to perform a more thorough quantitative evaluation (confusion matrix).
References


