

Segmentation of Breast Masses in Digital Mammograms Using Adaptive Median Filtering and Texture Analysis

Nasseer M. Basheer, Mustafa H. Mohammed

Abstract: Breast cancer continues to be one of the major causes of death among women. Early detection is a key factor to the success of treatment process. X-ray mammography is one of the most common procedures for diagnosing breast cancer due to its simplicity, portability and cost effectiveness. Mass detection using Computer Aided Diagnosis (CAD) schemes was an active field of research in the past few years, and some of these studies showed a promising future. These CAD systems serve as a second decision tool to radiologists for discovering masses in the mammograms. In this paper, a breast mass segmentation method is presented based on adaptive median filtering and texture analysis. The algorithm is implemented using MATLAB environment. The program accepts a digital mammographic image (images taken from the Mammographic Image Analysis Society (MIAS) database). Adaptive median filtering is applied for contouring the image, then the best contour is chosen based on the texture properties of the resulting Region-of-Interest (ROI). The proposed CAD system produces (92.307%) mass sensitivity at 2.75 False Positive per Image (FPI) which is considered as a proper result in this field of research.

Index Terms: Adaptive Median Filtering, Digital Mammograms, Mass Detection, Texture Analysis.

I. INTRODUCTION

In the last two decades, breast cancer has been the second leading cause of cancer deaths among women in some countries, following lung cancer [1]. An estimated 226,870 new cases of invasive breast cancer are expected to occur among women in the US during 2012; about 2,190 new cases are expected in men. Out of the 226,870 cases, 39,920 breast cancer deaths are expected [2]. Survival from breast cancer is directly related to its stage at diagnosis. The earlier the detection, the higher chances of successful treatments [3]. Breast image analysis can be performed using X-rays, magnetic resonance or ultrasound. So far the most effective and economical breast imaging modality has been the X-ray mammography due to its simplicity, portability and cost effectiveness [4]. It is based on the difference in absorption of X-rays between the various tissue components of the breast such as fat, tumor tissue, and calcifications. Mammography has high sensitivity and specificity, even small tumors and micro calcifications can be detected on mammograms.

Projection of the breast can be made from different angles [5]. The interpretation and analysis of medical images represent an important and exciting part of computer vision and pattern recognition. Developing a computer aided diagnosis system for cancer diseases, such as breast cancer, to assist physicians in hospitals is becoming of high importance and priority for many researchers and clinical centers [6].

An extensive work had been made in the field of digital mammography, especially in the last decade. These methods varied in the diagnosis of masses in mammograms. Here are some of the most common approaches used in recent researches on this domain.

Martins et al. 2009 presented a methodology for masses detection on digitized mammograms using the K-means algorithm for image segmentation and co-occurrence matrix to describe the texture of segmented structures. Classification of these structures was accomplished through Support Vector Machines (SVM). The proposed method achieved 86% sensitivity and 1.2 False Positive per Image (FPI) [7].

Chan et al. 2005 designed a CAD system for breast mass detection on digital breast tomosynthesis (DBT) mammograms. Each mass candidate was segmented from the structured background, and its image features were extracted. A feature classifier was designed to differentiate true masses from normal tissues. The CAD system achieved a sensitivity of 85%, with 2.2 false-positive objects per case [8].

Petrick et al. 2002 proposed and evaluated the performance of a CAD algorithm in marking preoperative masses. First, Digitized mammograms were processed with an adaptive enhancement filter followed by a local border refinement stage. Test results showed that malignant masses were detected with the computer in 87% (135 of 156), 83% (130 of 156), and 77% (120 of 156) of the malignant cases at FPI rates of 1.5, 1.0, and 0.5 marks per mammogram, respectively [9].

Zwiggelaar et al. used a recursive median filtering technique that can be applied to images at a number of scales and orientations, giving a scale space description at pixel level. This technique is applied to mammography, in the detection of mass-like structures associated with speculated lesions. A sensitivity of 80% is achieved with 0.25 false positives per image [10]. Cascio et al. 2006 presented an algorithm for detecting massive lesions in mammographic images. A reduction of the surface under investigation is achieved, without loss of meaningful information, through segmentation of the whole image, by means of a ROI Hunter algorithm. Once the features are computed for each ROI, they are used as inputs to a supervised neural network with momentum. According to the authors, segmentation resulted in 91% sensitivity and 6.6 FPI [11].

Revised Manuscript Received on 30 March 2013.

* Correspondence Author

Dr. Nasseer M. Basheer: Computer Engineering Department, Technical College/ Mosul , Iraq.

Mr. Mustafa H. Mohammed: M.Tech Student, Computer Engineering Department, Technical College/ Mosul , Iraq.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

In this paper, a mass segmentation algorithm is presented based on adaptive median filtering and texture analysis. The proposed algorithm is implemented under MATLAB (v7.12) platform. The metrics used to report the performance of the detection algorithm are sensitivity and the number of false positives per image (FPI). A true positive mark (TP) is a mark made by the CAD system that corresponds to the location of a lesion. A false positive mark (FP) is a mark made by the CAD system that does not correspond to the location of a lesion [12].

$$\text{sensitivity} = \frac{\text{Number of True Positive Marks (TP)}}{\text{Number of Lesions}}$$

$$\text{FPI} = \frac{\text{Number of False Positive Marks (FP)}}{\text{Number of Images}}$$

Sensitivity is the fraction of the true positive masses discovered by the CAD system to the real number of masses, while FPI is the ratio of the total number of normal regions misdiagnosed by the CAD system as masses to the total number of images.

The paper is organized as follows: section 2 presents an overview on X-ray mammography, section 3 discusses the method proposed for mass detection including the pre-processing and segmentation stages, section 4 displays the results and their performance evaluation. Finally, section 5 concludes the article.

II. X-RAY MAMMOGRAPHY

X-Ray Mammography is commonly used in clinical practice for diagnostic and screening purposes. Screening mammography has been recommended as the most effective method for early detection of breast cancer. Mammography provides high sensitivity on fatty breast and excellent demonstration of micro-calcifications; it is highly indicative of an early malignancy. Due to its low cost, it is suitable for mass screening program [13].

X-ray mammography is the best current method for early detection of breast cancer, with an accuracy of between 85% and 95% [14]. Identifying abnormalities such as calcifications and masses often requires the eye of a trained radiologist. As a result, some abnormalities may be missed due to human error as a result of fatigue, etc. The development of CAD systems that assist the radiologist has thus become of prime interest, the aim is not to replace the radiologist but to offer a second opinion [14].

Test images were taken from the Mammographic Image Analysis Society (MIAS) database. MIAS is an organization of UK research groups interested in the understanding of mammograms. The database is arranged in pairs of films, where each pair represents the left and right mammograms of a single patient. All the images have the size of 1024 × 1024 pixels with a spatial resolution of 200µm and a gray scale level of 8 bits. The database also provide information about the characteristics of background tissue, class of abnormality and severity of abnormality. The (x, y) coordinates of the center of lesions are also provided by the database, along with the approximate radius (in pixels) of a circle enclosing the abnormality [15]. With all these information provided, the MIAS database is considered as one of the most common sources of test mammograms that can be used in scientific researches and CAD systems [1,3,5,6 etc...].

III. THE PROPOSED METHOD

The aim of this work is to create and implement an algorithm that accepts a raw mammographic image and mark any possible mass regions that may appear in the image based on intensity, shape, size and other texture measures.

A breast mass segmentation algorithm has been proposed and implemented under the MATLAB platform. First, a preprocessing stage is required to remove areas that are not related to the detection region (pectoral muscle and any artifact labels that can be applied on the mammogram). Preprocessing also intends to enhance the overall contrast of the image. Secondly, image segmentation is followed which is about extracting several, non-overlapping ROI candidates from the background tissue. Image segmentation in mammograms can be a complicated process due to the variability of the size, shape and uniformity of the masses and the diversity in intensity of the image background. In the following, these two stages will be investigated in detail.

A. Pre-processing

Preprocessing is an important issue in low-level image processing. The underlying principle of preprocessing is to enlarge the intensity difference between objects and background and to produce reliable representations of breast tissue structures [5]. In this work, the approaches applied in the preprocessing stage are listed below.

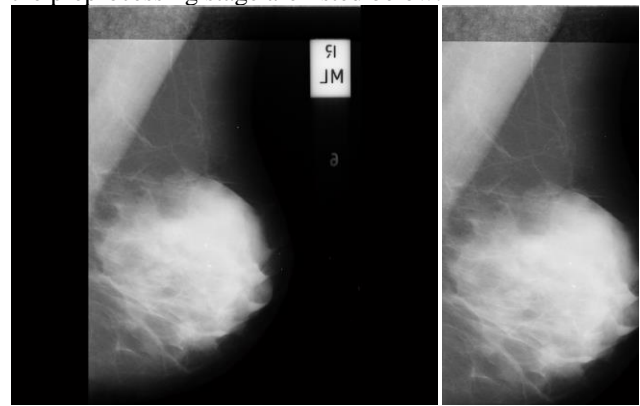


Fig.1: Initial cropping (a) a raw mammographic image, (b) image after initial cropping

1. Initial Cropping

First, a raw mammogram from the MIAS database is cropped to remove dark areas that are not related to the breast region. In the MIAS database, each image is clipped or padded so that it is 1024 × 1024 pixels, which is the size of all MIAS raw images. Therefore, the dark padding and regions outside the breast are excluded from participating in later stages to keep the areas of interest only for further processing. Figure 1 (a) and (b) shows a mammogram image before and after cropping.

2. Intensity Adjustment

Image enhancement techniques are used to improve an image, where "improve" is sometimes defined objectively (e.g., increase the signal-to-noise ratio), and sometimes subjectively (e.g., make certain features easier to see by modifying the colors or image's intensity values to a new range [16].

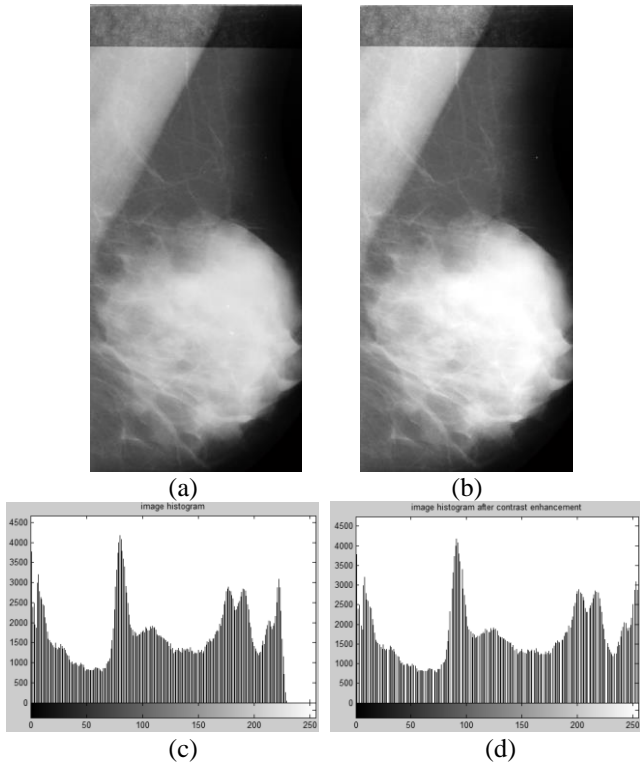


Fig.2: Comparison between test images and histograms before and after intensity adjustment (a) test image before enhancement, (b) test image after enhancement, (c) histogram of the image before enhancement, (d) histogram of the image after enhancement

The histogram of the cropped mammographic image is scaled so that it utilizes the entire histogram range (0 to 255). Applying intensity adjustment shows the enhanced contrast of the resulted image compared to the original one. This becomes clearer when comparing the histogram of the enhanced image to that of the original one. In figure (2), (a,b) show the image before and after intensity adjustment, while (c,d) compares the histogram of images before and after intensity adjustment.

3. Contrast Limited Adaptive Histogram Equalization (CLAHE)

Adaptive histogram equalization maximizes the contrast throughout an image by adaptively enhancing the contrast of each pixel relative to its local neighborhood. This process produces improved contrast for all levels of contrast (small and large) in the original image [5].

CLAHE is a modification of the original histogram equalization operating on small regions in the image, called *tiles*, rather than the entire image. Each tile's contrast is enhanced, then the neighboring tiles are combined using bilinear interpolation (directly implemented by MATLAB) to eliminate artificially induced boundaries. Test image after CLAHE is shown in figure (3a).

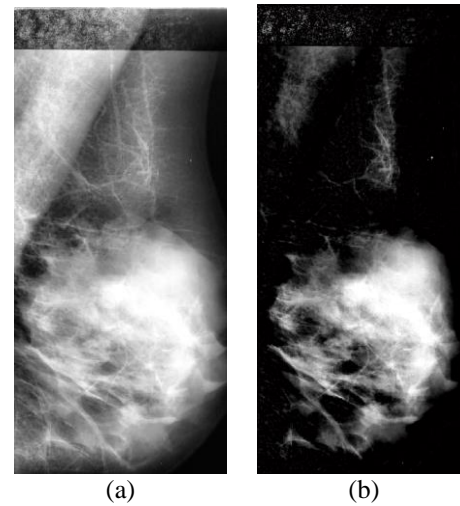


Fig.3: (a) test image after CLAHE, (b) test image after border clearing

4. Image Border Clearing

This is a supplemental stage to remove the high intensity areas that are known to be non-mass regions like the pectoral muscle and any artifact labels that are attached to the borders of the image. White neighborhoods all over the image that are connected to the image's borders are removed during this stage, resulting in a darker image, as shown in figure (3b).

B. Segmentation

The second stage of mass detection CAD schemes is to separate the suspicious regions that may contain masses from the background parenchyma [5], i.e., to partition the mammogram into several non-overlapping regions, then extract regions of interests (ROIs), and locate the suspicious mass candidates from ROIs. The suspicious area is an area that is brighter than its surroundings, has almost uniform density, has a regular shape with varying size, and has fuzzy boundaries [5]. Segmentation of masses from the rest of the image can be a complex process due to the diversity of mass characteristics from one image to the other. The following methods are applied in the segmentation stage.

1. Small Neighborhoods Removal

Test images after preprocessing stage may be distorted with pixel neighborhoods that are too small and not bright enough to be potential masses. These neighborhoods which are distributed throughout the image, may have a negative impact on the segmentation and therefore should be removed. Figure (4a) shows the resulted image after this stage.

2. Adaptive Median Filtering

Median filtering is commonly used in digital mammograms as a preprocessing stage for noise removal. Here, 2D adaptive median filtering is applied on the image for clarifying edges and making gradual changes in intensity more skew for expounding the borders of any possible masses as shown in figure (4b). Test image before median filtering has fuzzy, wide borders that are difficult to isolate. Adaptive median filtering makes the overall image fuzzier but turn the boundaries of any objects exist in the image into fine, crisp, straight lines that can be isolated directly later. A close up to the ROI in the image is shown in figure (5).

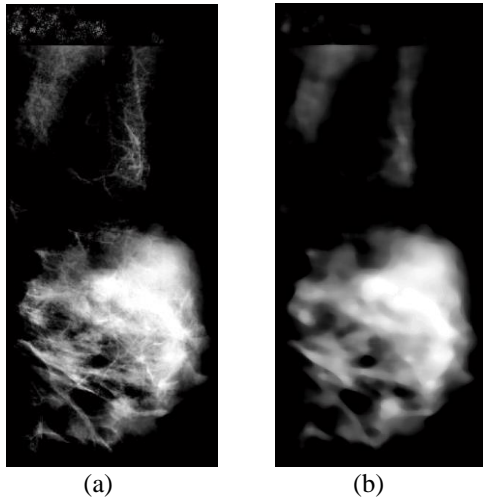


Fig. 4: (a) image after removing small objects. (b) image after Adaptive Median Filtering

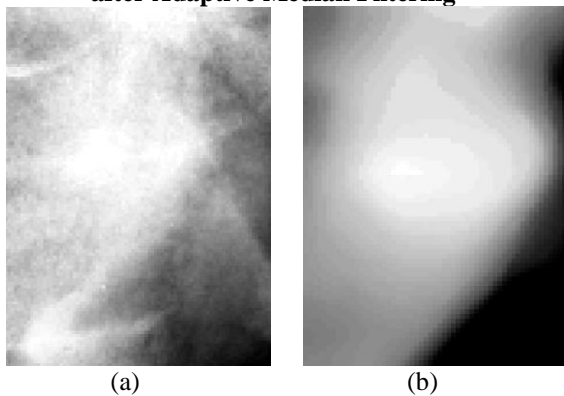


Fig.5: a potential ROI (a) before and (b) after adaptive median filtering

3. Border selection

The image resulted from median filtering has many edges and boundaries in the form of closed curves that are encircling each other. The last step for segmentation is to select one of these curves to be the edge of the potential mass region. The area inside the selected curve is considered as the mass. The criterion for selecting one of these curves is highly important for a good segmentation. This process is independent for each object in the image based on the texture features of that object. For each object, the process starts from the highest gray level (the smallest curve in the center of other curves) and goes through lower gray levels (larger curves to the outside). When the region enclosed by one of the curves (selected at a certain gray level) has a proper texture and geometric features, this region is considered as a potential ROI that is added to the resulted segmentation image as shown in figure (6).

IV. RESULTS

The proposed algorithm was applied on 36 images from the MIAS mammographic database. Some of these images were containing more than one possible mass region. The algorithm produces a binary image (segmentation mask) and uses its borders as markings on the same original image to indicate the abnormality. In the 36 images there were 39 possible mass regions. The algorithm was able to indicate abnormalities in 33 images out of 36, scoring (91.667%) image sensitivity.

In terms of masses, 36 out of 39 possible masses are indicated by the algorithm, scoring a (92.307%) mass

sensitivity. Both results are produced at 2.75 false positive per image (FPI) rate which is considered very challenging when compared to the results of the previously mentioned works listed in the introduction since its either better in terms of sensitivity or FPI which are inversely related to each other.



Fig.6: The potential ROI marked on the original image, showing the mass location on the breast region

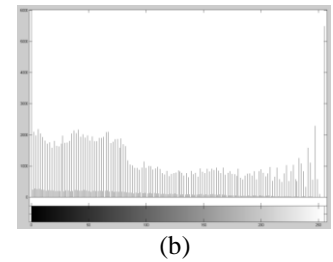


Fig.7. histogram of the image resulted from adaptive median filtering

V. CONCLUSION

In this paper, adaptive median filtering which is commonly used as a preprocessing step to remove noise is used here for reducing the possible number of gray levels for thresholding and producing crisp and clear boundaries in the image for easier and more practical segmentation. By viewing the image's histogram in figure (7), it can be noticed that the image resulted from adaptive median filtering has a fewer number of gray levels, which leads to a simpler and faster thresholding since the algorithm will scan the fewer number of gray levels of the filtered image instead of the full gray level range of the original one. Also, the use of intensity adjustment as a preprocessing step is very useful for enhancing the contrast and sharpening edges that may be found in the image due to the utilization of the entire intensity range of the image histogram.

REFERENCES

1. F. Moayed, R. Boostani, A. R. Kazemi, S. Katebi And E. Dashti; "Subclass Fuzzy-Svm Classifier As An Efficient Method To Enhance The Mass Detection In Mammograms", Iranian Journal Of Fuzzy Systems, Vol. 7, No. 1, Pp. 15-31, 2010.
2. CONE HEALTH, the network for exceptional care, Regional Cancer Statistics January 2012.
3. VALLIAPPAN Raman, PUTRA Sumari and MANDAVA Rajeswari; "A Theoretical Methodology and Prototype Implementation for Detection Segmentation Classification of Digital Mammogram Tumor by Machine Learning and Problem Solving Approach", International Journal of Computer Science Issues (IJCSI), Vol. 7, Issue 5, September 2010.
4. H.B.Kekre, Tanuja K. Sarode, Saylee M. Gharage; "Tumor Detection in Mammography Images using Vector Quantization Technique", International Journal of Intelligent Information Technology Application, Vol. 2, No. 5, pp. 237-242, 2009.

5. Valliappan Raman, Putra Sumari, H.H.Then, and Saleh Ali K. Al-Omari; "Review on Mammogram Mass Detection by Machine Learning Techniques", International Journal of Computer and Electrical Engineering, Vol. 3, No. 6, December 2011.
6. Sara Dehghani and Mashallah Abbasi Dezfooli; "Breast Cancer Diagnosis System Based on Contourlet Analysis and Support Vector Machine", World Applied Sciences Journal, Vol. 13, No. 5, pp. 1067-1076, 2011.
7. Leonardo O. Martins, Geraldo B. Junior, Aristófanés C. Silva, Anselmo C. de Paiva, and Marcelo Gattass; "Detection of Masses in Digital Mammograms using K-means and Support Vector Machine", Electronic Letters on Computer Vision and Image Analysis, Vol. 8, No. 2, pp. 39-50, 2009.
8. Heang-Ping Chan, Jun Wei, Tao Wu and Mark A. Helvie; "Computer-aided Detection System for Breast Masses on Digital Tomosynthesis Mammograms: Preliminary Experience", Radiology, Vol. 237, No. 3, December 2005.
9. Nicholas Petrick, Berkman Sahiner and Lubomir M. Hadjiiski; "Breast Cancer Detection: Evaluation of a Mass-Detection Algorithm for Computer-aided Diagnosis Experience in 263 Patients", Radiology, Vol. 224, No. 1, July 2002.
10. Reyer Zwiggelaar, James E. Schumm and Christopher J. Taylor; "The Detection of Abnormal Masses in Mammograms", Wolfson Image Analysis Unit, University of Manchester, Manchester, UK.
11. D. Cascio, F. Fauci, R. Magro and E. Lopez Torres; "Mammogram Segmentation by Contour Searching and Massive Lesion Classification with Neural Network", Institute of Electrical and Electronic Engineering (IEEE), 2006.
12. Mehul P. Sampat, Mia K. Markey and Alan C. Bovik; "Computer-Aided Detection and Diagnosis in Mammography".
13. Yao Yao; "SEGMENTATION OF BREAST CANCER MASS IN MAMMOGRAMS AND DETECTION USING MAGNETIC RESONANCE IMAGING", School of Electrical & Electronic Engineering, Nanyang Technological University.
14. Alfred M. Mutiso; "DATAMINING IN MEDICAL APPLICATIONS: COMPUTER AIDED DIAGNOSIS (CAD) IN MEDICAL IMAGING WITH AN EMPHASIS ON MAMMOGRAPHY".
15. The MAMMOGRAPHIC IMAGE ANALYSIS SOCIETY (MIAS), mini mammographic image database guide.
16. TRIPTY SINGH, M.NAGRAJA, SWARNALATA RAO, MAHUA BHATTACHARYA and SANTOSH BOMMANALLI; "ENHANCING IMAGE CONTRAST OF MAMMOGRAM & EQUALIZATION OF HISTOGRAMS", International Journal of Engineering Science and Technology (IJEST), Vol. 3, No. 1 Jan 2011.

AUTHOR PROFILE



Dr. Nasseer M. Basheer: Born in Mosul, Iraq in 1958. He obtained his BSc degree in Electrical Engineering in 1980, MSc in electronics and communication engineering in 1984 at Mosul university, and PhD in electronics engineering in 2008. He is interested in the subjects of digital image processing, microprocessors and FPGA techniques.



Mr. Mustafa H. Mohammed: Born in Mosul, Iraq in 1986. He obtained his BSc degree in Computer Engineering from Mosul university in 2008. He is currently preparing for the master degree in the Technical College/ Mosul. He is interested in digital image processing, computer networks and control systems.