Mammography Images Segmentation Based on Fuzzy Set and Thresholding

Ali Mohammed Salih, Mohammed Y. Kamil*
Department of Physics, College of Science, Mustansiriya University, IRAQ
*Correspondent author email: m80y98@uomustansiriya.edu.iq

Abstract
Breast cancer is the most widespread cancer that influences ladies about the world. Early recognition of breast tumor is a standout amongst the hugest variables influencing the probability of recuperation from the illness. Hence, mammography remains the most precise and best device for distinguishing breast malignancy.

This paper presents a method for segment the boundary of breast masses regions in mammograms via a proposed algorithm based on fuzzy set techniques. Firstly, it was used data set (mini-MIAS) for evaluate algorithm. it was preprocessing the data set to remove noise and propose a fuzzy set by using fuzzy inference system by generated two input parameters (employs image gradient), then used thresholding filter. Then it was evaluated this proposed method, qualitative and quantitative results were obtained to demonstrate the efficiency of this method and confirm the possibility of its use in improving the diagnosis.

Keywords: Mammography, Segmentation, Fuzzy logic, Mass detection, Thresholding

Introduction
Cancer is a set of diseases that grow out of control and cause cells in the body to change. Most kinds of cancer cells in the long run frame an irregularity or masses called a tumor. Breast cancer starts in breast tissue, which is comprised of glands for milk production that called lobules and the ducts that connect nipple to the lobules. Also the rest of the breast is made up of connective, fatty, and lymphatic tissue[1]. Breast cancer is the most common cancer affecting women worldwide. It is visible in two forms: A cancerous tumor is malignant, meaning it can grow and spread to other parts of the body. A benign tumor means the tumor can grow but will not spread. Early detection of breast cancer is one of the most important factors affecting the prospect of recovery from the disease. For that cause, the mammography remains the most accurate and best tool in detecting breast cancer [2]. mammography is a private kind of x-ray imaging used to make detailed images of the breast. It uses low dose x-ray, high contrast, high accurate films specifically for imaging the breasts[3]. Studies have shown a reduction in both sharp breast cancer and death-rate in women who go through the ordinary

Copyright © 2018 Authors and Al-Mustansiriyah Journal of Science. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.
mammographic screening [4, 5]. In the medical community, breast tissue density is an imperative hazard marker for growth of breast malignancy [6, 7]. The extent of fibroglandular and fatty tissue of the breast region is assessed by a radiologist for the understanding of mammographic images. The outcome is subjective and differs from one radiologist to another. Radiologists can fail to detect a significant quantity of abnormalities [8] and countless anomalies may end up being kind after biopsy [6]. with a specific end goal to improve the accuracy of explain the mammograms, an assortment of computer aided diagnosis (CAD) systems that perform automated mammographic examination have been proposed, as stated in refs. [9, 10], these systems are usually employed as supplement to the radiologists’ assessment. Thus, their role in modern medical practice is considered to be significant and important in the early detection of breast cancer. In recent years, some researchers have been using the Fuzzy set approach in the field of medical image processing [11, 12]. B. O. Alilja, et al. [13], proposed the use of Fuzzy Rough Set Method to select the most significant texture features from mammogram images. Selected features are using to build an easier and ideal learning model in order to perfect the classification goodness of mammogram analysis systems. J. Kaur and M. Mahajan [14], proposed combine the random walker with advantages of fuzzy logic to make resulting segmentation better in quality and texture. they used employed fuzzy rules to approximate boundaries in images which improved segmentation results. V. Ananthi, et al. [15], proposed a new fuzzy approach into the segmentation of images. Numbers of gray levels-interval-valued intuitionistic fuzzy sets are constructed from two numbers of gray levels-fuzzy sets that coincide to the background of an image and foreground. Here, L refers the number of gray levels in the image. Threshold for an image is select by finding an interval-valued intuitionistic fuzzy set with least entropy. S. Arora and Q. Kaur [16] adjust edge detection technique employed fuzzy inference system (FIS) based on conventional operators. A. Borkar and M. Atulkar tried to detect the edge of images based on Fuzzy Inference System without defining threshold value in MATLAB program [17, 18]. C. L. Chowdhary and D. Acharjya [19], selected a set of clusters to shape a final clustering solution with the employed of research in theory and implementation of fuzzy sets. P. Kaur, et al. [20] have been proposed method is useful in segmented several areas of the medical images and helps to find abnormalities cases in image. In (IFS), there are membership, non-membership and hesitation degree[21].

Materials and Methods
Zadeh introduced fuzzy set in 1965 [22], where probability theory was first invented to provide gradient implications for natural language data. The use of this theory has been expanded to various domains cooperation with inaccurate data [23].

\[ X = \{x_1, x_2, x_3, \ldots\} \]  

A fuzzy set \( B \) in a finite set can be represented mathematically like:

\[ B = \{(x, \mu_B(x)) | x \in B\} \]

where the function:

\[ \mu_B(x) : X \to [0,1] \]

is the membership function of an element \( x \) in the finite set \( X \) or the measure of the degree of belongingness. Also, the measure of non-belongingness is \( 1 - \mu_B(x) \) [24]. The input mammogram images are taken from (mini-MIAS) database that consists of 322 images. The images in the database consist of right and left breast orientation of dense glandular, fatty glandular and fatty. The images are digitized to 200-micron pixel edge and padded/clipped so that each image is 1024 \( \times \) 1024 pixels and 8-bit precision in Portable Gray Map (PGM) format.

Proposed Algorithm
This section of the manuscript talks about the proposed algorithm to determine tumor area in mammography images to facilitate the doctor's radiology diagnosis process (see fig. 1).
The first step is input original image, after that, it has been preprocessing to separate of breast profile region without artifacts and noise. Then, it employs image gradient by using Mean Filter (MF) to determine the disparity of the image to ensure the segmentation in relative low contrast areas, as well used High Pass Filter (HPF) to detect image pixels that belong to the input regions, are shown in fig. 2.

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1
\end{bmatrix}, \quad
\begin{bmatrix}
-1 & -2 & -1 \\
-2 & 12 & -2 \\
-1 & -2 & -1
\end{bmatrix}
\]

Figure 2: Mean Filter (MF) and High Pass Filter (HPF).

Then \(hpf\) and \(mf\) are separately defined as \(hpf = HPF \ast A\) and \(mf = MF \ast A\). Where \(\ast\) represent the convolution operator and \(A\) represent the original image.

Thereafter, the fuzzy system was built based on fuzzy inference system (FIS). Two fuzzy sets are consisting of underlying variables “Zero”, and “not Zero” in order to represent every variable’s intensities. The inputs for all membership functions are Gaussian and set the range between -1 and 1, as shown in Figure 3(a). The output membership functions for variable segmentation show in Figure 3(b) that also set the range between 0 and 1 and the fuzzy sets are Triangle function. Figure 4 show surface viewer rule in F-Set include the rules, the two fuzzy set rules that allow evaluating the input image are shown below:

1. If (hpf is zero) and (mf is zero) then (segmentation is white).
2. If (hpf is not zero) or (mf is not zero) then (segmentation is black).

After defuzzification image, we use region of interest (ROI) to determine the abnormality are, also used thresholding filter to detect the boundaries around the masses or calcifications. Finally output image and compare with ground truth image by specific the radiologist.
Results and Discussion
This section shows the results of the research conducted on the dataset (MIAS). The proposed methodology is applied on the database (mini-MIAS) in order to assess our method in using some of information given at the (MIAS) such as the radius of the circle enclosing the abnormality and image coordinates of abnormality center. The proposed algorithm provided good results when diagnosing the tumors as shown in figs. (5-8).

The quantitative analysis for performance images of the segmentation results is compared with the ground truth by using Dice Similarity Index (DSI), Correct Detection Ratio (CDR), Miss Rate, and Sensitivity as shown in table 1. The values (DSI and CDR) shown in table 1 are the highest possible value of the images: mbd028, mbd081, mbd134, the reason behind this is that tumors are masses; while on image mbd231, the algorithm failed to diagnose the tumor because the tumor are calcifications, such calcifications are dispersing on the images and have low difference; it is also difficult to identify them out of the surrounding area, such specifications make it difficult to segment properly. As for table2, the characteristics of the area ground truth and segmentation; one of the adopted characteristics is (Area, center x-axis (C.X), center y-axis (C.Y), Diameter, Minor Axis Length, Major Axis Length, Perimeter, and Solidity).

<table>
<thead>
<tr>
<th>Image</th>
<th>CDR</th>
<th>DSI</th>
<th>Sensitivity</th>
<th>Miss Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>mbd028</td>
<td>0.8321</td>
<td>0.9084</td>
<td>0.2848</td>
<td>16 %</td>
</tr>
<tr>
<td>mbd081</td>
<td>0.7125</td>
<td>0.8321</td>
<td>0.1493</td>
<td>28 %</td>
</tr>
<tr>
<td>mbd134</td>
<td>0.8293</td>
<td>0.9067</td>
<td>0.2072</td>
<td>17 %</td>
</tr>
<tr>
<td>mbd231</td>
<td>0.0020</td>
<td>0.0040</td>
<td>0.0080</td>
<td>99 %</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Image</th>
<th>C.x</th>
<th>C.y</th>
<th>Area</th>
<th>Diameter</th>
<th>Major Axis length</th>
<th>Minor Axis length</th>
<th>Perimeter</th>
<th>Solidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>mbd028</td>
<td>338</td>
<td>704</td>
<td>8142</td>
<td>101.8</td>
<td>107.6</td>
<td>97.8</td>
<td>346.2</td>
<td>0.93</td>
</tr>
<tr>
<td>mbd081</td>
<td>68</td>
<td>67</td>
<td>9604</td>
<td>110.6</td>
<td>119.1</td>
<td>107.4</td>
<td>427.3</td>
<td>0.88</td>
</tr>
<tr>
<td>mbd134</td>
<td>146</td>
<td>116</td>
<td>46906</td>
<td>244.4</td>
<td>269.7</td>
<td>228.3</td>
<td>1117.6</td>
<td>0.91</td>
</tr>
<tr>
<td>mbd231</td>
<td>71</td>
<td>131</td>
<td>1</td>
<td>1.1</td>
<td>1.3</td>
<td>1.2</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1: Measurements (CDR, DSI, Sensitivity, Miss Rate) for segmentations carried out on mammograms from the MIAS database.

Table 2: Region properties for ground truth and segmentation.
Figure 5: the result for mdb028 image. (a) Original Image. (b) Apply hpf. (c) Apply mf. (d) Final result Fuzzy set. (e) ROI. (f) Apply threshold. (g) Comparison.

Figure 6: the result for mdb081 image. (a) Original Image. (b) Apply hpf. (c) Apply mf. (d) Final result Fuzzy set. (e) ROI. (f) Apply threshold. (g) Comparison.
Conclusions

The outcomes for the proposed method shown higher (CDR and DSI) values with compared to Sensitivity. This leads to reduces false positive detection, which the marked pixel is the nearest to the real bounded. Fuzzy set introduced the accuracy of the segmentation based on IF-THEN rule, the boundaries of the selected set areas have been determined for more image analysis. This algorithm is convenient for applications in different areas of computer vision, like medical imaging.

References


