Research Article Open Access

Mammography Images Segmentation Based on Fuzzy Set and Thresholding

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

Article Info

Received
10/10/2018

Accepted
25/10/2018

Published
10/03/2019

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

الخلاصة

يعد سرطان الثدي الأكثر انتشارًا والذي يؤثر على السيدات حول العالم. اذ يعتبر الكشف المبكر على ورم الثدي من أبرز المتغيرات التي تؤثر على احتمالية الشفاء من المرض. وبالتالي ، يبقى التصوير الشعاعي للثدي الأكثر دقة والأفضل جهازاً للتمييز بين الأورام الخبيثة في الثدي.

تقدم هذه الورقة طريقة لتجزئة وتقسيم حدود مناطق تكتلات الثدي في صور التصوير الشعاعي للثدي عن طريق خوارزمية مقترحة تعتمد على تقنيات المجاميع المضببة. أولاً ، تم استخدام مجموعة البيانات (mini-MIAS) لتقييم الخوارزمية. كانت عملية المعالجة الأولية لمجموعة البيانات لإزالة الضوضاء واقتراح مجموعة ضبابية باستخدام نظام الاستدلال الضبابي عن طريق إنشاء مدخلين متغيرين (يتضمن انحدار الصورة) ، ثم تم استخدام فلتر العتبة. ثم تم تقييم هذه الطريقة المقترحة ، وتم الحصول على نتائج نوعية وكمية تدل على كفاءة هذه الطريقة وتأكيد إمكانية استخدامها في تحسين التشخيص.

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 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 failure 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. Alijla, 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].

$$X = \{x_1, x_2, x_3, \dots\} \qquad (1)$$

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

$$B = \{(x, \mu_B(x)) | x \in B\} \qquad (2)$$

where the function:

$$\mu_B(x): X \to [0,1] \qquad (2)$$

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 × 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).

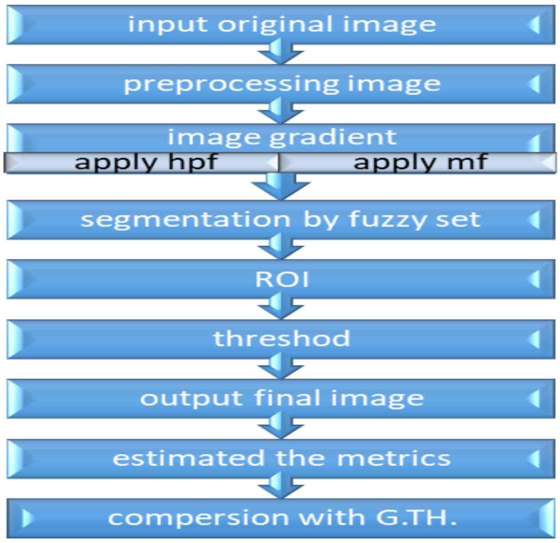


Figure. 1: Flowchart the proposed algorithm.

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.

$$1/9 \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \; , \; \begin{array}{|c|c|c|}\hline -1 & -2 & -1 \\ \hline -2 & 12 & -2 \\ \hline -1 & -2 & -1 \\ \hline \end{array}$$

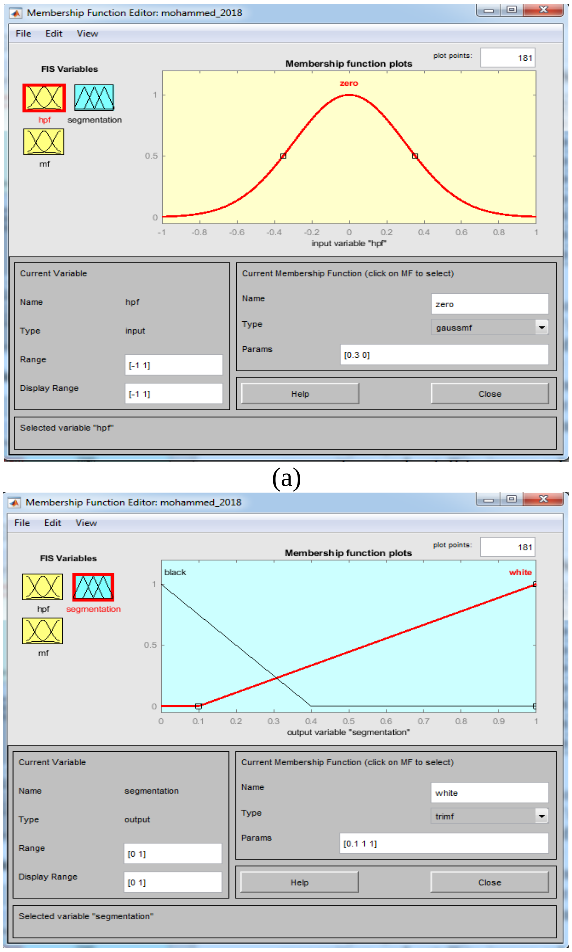
MF HPF

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

Then hpf and mf are separately defined as $hpf = HPF * A \text{ and } mf = MF * A$. Where $*$ 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 1and 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).

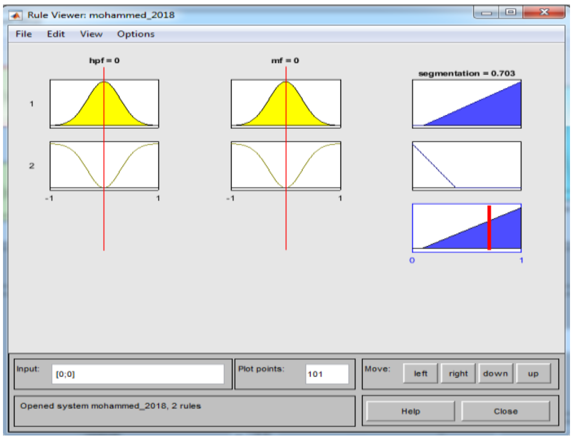


(a)

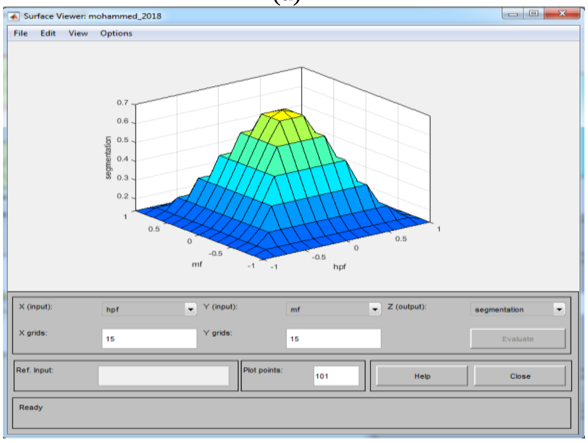
(b)

Figure 3: function editor, (a) Input membership. (b) Output membership

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.



(a)



(b)

Figure 4: viewer window for, (a) Rules. (b) Surface rules

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 asses 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).

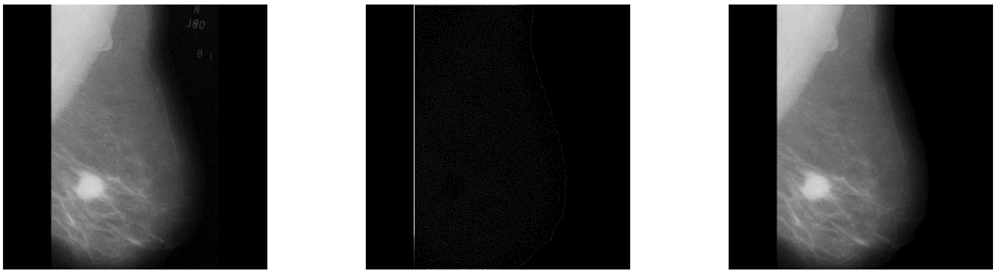
The quantitative analysis for performance images of the segmentation results is compared with the ground truth by using Dice Similarity Index (DSI), Correct Detection Ratio (CDR), Miss Rate, and Sensitivity as shown in table 1. The values (DSI and CDR) shown in table 1 are the highest possible value 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 1: Measurements (CDR, DSI, Sensitivity, Miss Rate) for segmentations carried out on mammograms from the MIAS database.

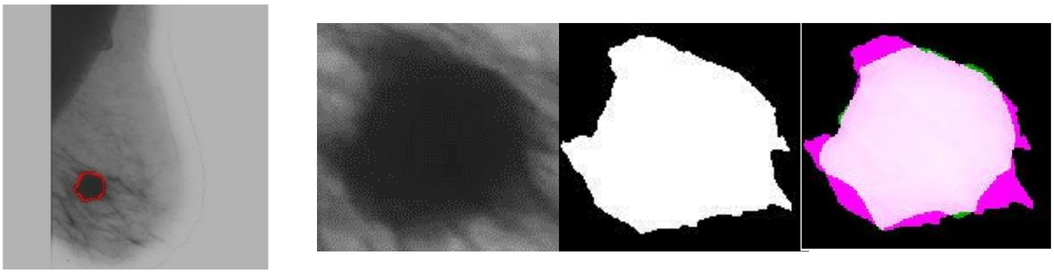
Image	CDR	DSI	Sensitivity	Miss Rate
mdb028	0.8321	0.9084	0.2848	16 %
mdb081	0.7125	0.8321	0.1493	28 %
mdb134	0.8293	0.9067	0.2728	17 %
mdb231	0.0020	0.0040	0.0080	99 %

Table 2: Region properties for ground truth and segmentation.

Image		C.x	C.y	Area	Diameter	Major Axis length	Minor Axis length	Perimeter	Solidity
mdb028	**G.T.**	338	704	8142	101.8	107.6	97.8	346.2	0.93
	Seg.	68	67	9604	110.6	119.1	107.4	427.3	0.88
mdb081	**G.T.**	488	549	35800	213.5	244.5	195.6	896.8	0.85
	Seg.	146	116	46906	244.4	269.7	228.3	1117.6	0.91
mdb134	**G.T.**	468	298	5179	81.2	94.5	71.4	277.3	0.95
	Seg.	50	63	5302	82.1	91.8	77.9	327.9	0.87
mdb231	**G.T.**	614	493	2621	57.8	82.7	57.3	346.9	0.66
	Seg.	71	131	1	1.1	1.3	1.2	0	1

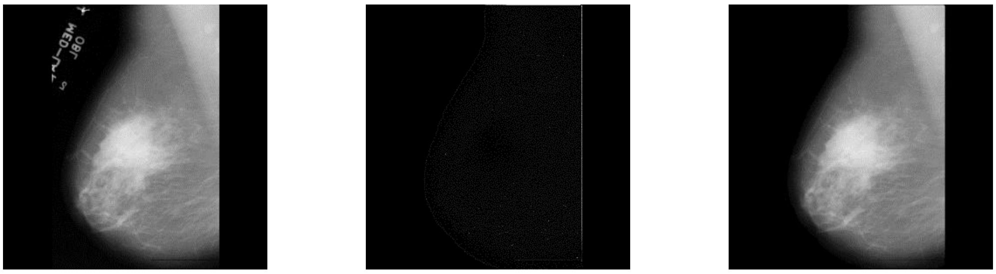


(a) (b) (c)

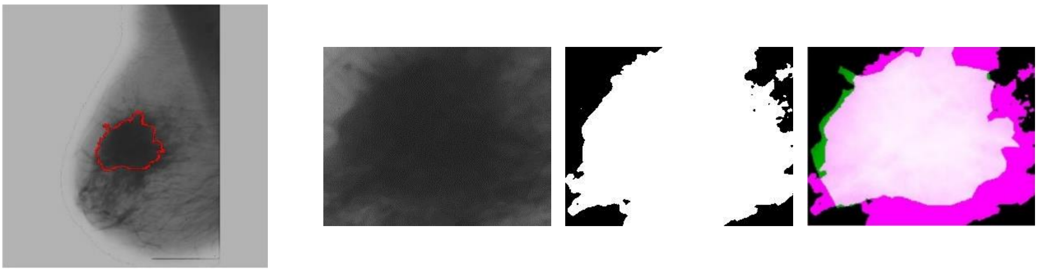


(d) (e) (f) (g)

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.

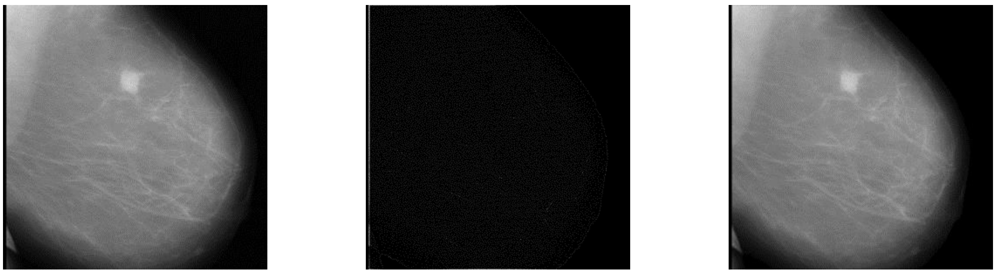


(a) (b) (c)

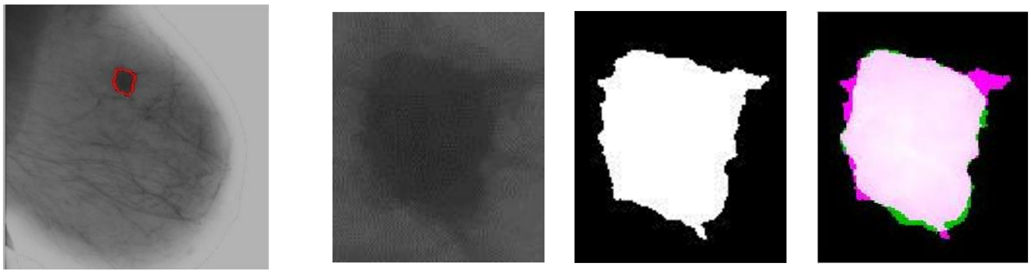


(d) (e) (f) (g)

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.

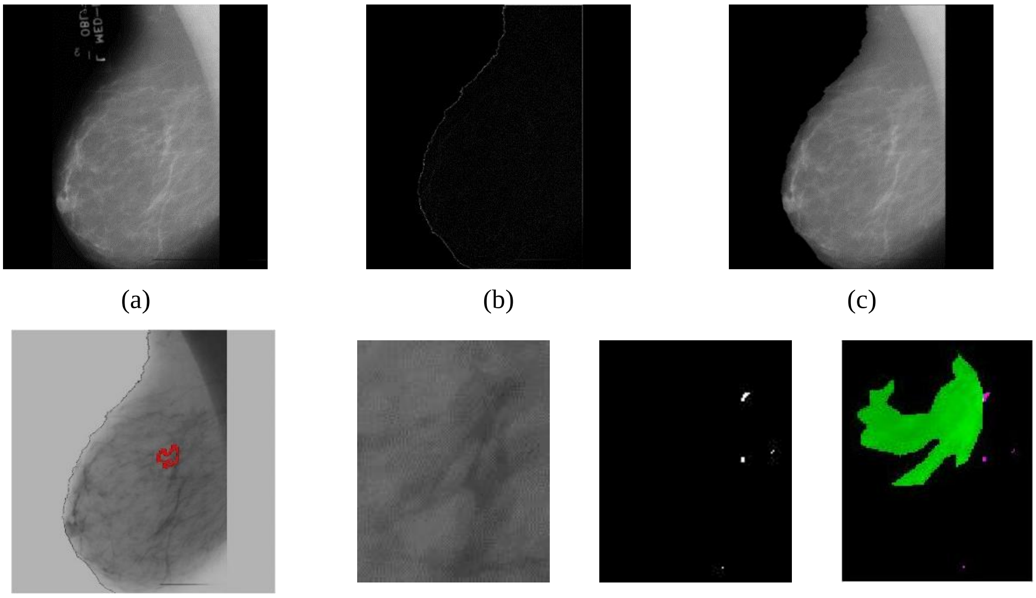


(a) (b) (c)



(d) (e) (f) (g)

Figure 7: the result for mdb134 image. (a) Original Image. (b) Apply hpf. (c) Apply mf. (d) Final result Fuzzy set. (e) ROI. (f) Apply threshold. (g) Comparison.



(a) (b) (c)

(d) (e) (f) (g)

Figure 8: the result for mdb231 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

[1] Senthilkumar and G. Umamaheswari, "Breast cancer detection using combined curvelet based enhancement and a novel segmentation methods," *BIOMEDICAL PAPERS-OLOMOUC*, vol. 159, pp. 83-86, 2015.

[2] A. Ferrero, S. Salicone, A. Mencattini, G. Rabottino, and M. Salmeri, "Uncertainty evaluation in a fuzzy classifier for microcalcifications in digital mammography," in *Instrumentation and Measurement Technology Conference (I2MTC), 2010 IEEE*, 2010, pp. 1250-1255.

[3] M. A. Duarte, A. V. Alvarenga, C. M. Azevedo, M. J. G. Calas, A. F. Infantosi, and W. C. Pereira, "Evaluating geodesic active contours in microcalcifications segmentation on mammograms,"

Computer methods and programs in biomedicine, vol. 122, pp. 304-315, 2015.

[4] M. Abdel-Nasser, A. Moreno, and D. Puig, "Towards cost reduction of breast cancer diagnosis using mammography texture analysis," *Journal of Experimental & Theoretical Artificial Intelligence,* vol. 28, pp. 385-402, 2016.

[5] N. Saidin, U. K. Ngah, H. A. M. Sakim, D. N. Siong, M. K. Hoe, and I. L. Shuaib, "Density based breast segmentation for mammograms using graph cut and seed based region growing techniques," in *Computer Research and Development, 2010 Second International Conference on*, 2010, pp. 246-250.

[6] N. Singh, A. G. Mohapatra, and G. Kanungo, "Breast cancer mass detection in mammograms using K-means and fuzzy C-means clustering," *International Journal of Computer Applications (0975–8887),* vol. 22, 2011.

[7] A. Touil and K. Kalti, "Iterative fuzzy segmentation for an accurate delimitation of the breast region," *Computer methods and programs in biomedicine,* vol. 132, pp. 137-147, 2016.

[8] C. J. Baines, D. V. Mcfarlane, and A. B. Miller, "The Role of the Reference Radiologist; Estimates of Inter-Observer Agreement and Potential Delay in Cancer Detection in the National Breast Screening Study," *Investigative Radiology,* vol. 25, pp. 971-976, 1990.

[9] K. El Fahssi, A. Elmoufidi, A. Abenaou, S. Jai-Andaloussi, and A. Sekkaki, "Mass segmentation in mammograms based on the minimisation of energy and active contour model," in *Medical Measurements and Applications (MeMeA), 2014 IEEE International Symposium on*, 2014, pp. 1-5.

[10] W. I. Suleiman, S. Lewis, D. Georgian-Smith, M. G. Evanoff, and M. F. McEntee, "Number of mammography cases read per year is a strong predictor of sensitivity," *Journal of Medical Imaging,* vol. 1, p. 015503, 2014.

[11] M. Y. Kamil, "Morphological gradient in brain magnetic resonance imaging based on intuitionistic fuzzy approach," in *Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA), Al-Sadeq International Conference on*, 2016, pp. 1-3.

[12] M.Y. Kamil and A. M. Salih, "Mammography Images Segmentation via Fuzzy C-mean and K-mean," International Journal of Intelligent Engineering and Systems, Vol.12, No.1, p. 22-29, 2019.

[13] B. O. Alijla, A. T. Khader, L. C. Peng, M. A. Al-Betar, and W. L. Pei, "Fuzzy Rough Set Approach for Selecting the Most Significant Texture Features in Mammogram Images," in *Information and Communication Technology (PICICT), 2013 Palestinian International Conference on*, 2013, pp. 51-56.

[14] J. Kaur and M. Mahajan, "Hybrid of Fuzzy Logic and Random Walker Method for Medical Image Segmentation," *International Journal of Image, Graphics and Signal Processing,* vol. 7, p. 23, 2015.

[15] V. Ananthi, P. Balasubramaniam, and P. Raveendran, "A thresholding method based on interval-valued intuitionistic fuzzy sets: an application to image segmentation," *Pattern Analysis and Applications,* pp. 1-13, 2017.

[16] S. Arora and Q. Kaur, "Modified edge detection technique using fuzzy inference system," *Computer Application,* vol. 44, pp. 9-12, 2012.

[17] A. D. Borkar and M. Atulkar, "Fuzzy Inference System for Image Processing," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET),* vol. 2, pp. pp: 1007-1010, 2013.

[18] A. D. Borkar and M. Atulkar, "Detection of edges using fuzzy inference system," *International Journal of Innovative Research in Computer and*

Communication Engineering, vol. 1, 2013.

[19] C. L. Chowdhary and D. Acharjya, "Segmentation of Mammograms Using a Novel Intuitionistic Possibilistic Fuzzy C-Mean Clustering Algorithm," in *Nature Inspired Computing*, ed: Springer, 2018, pp. 75-82.

[20] P. Kaur, A. Soni, and A. Gosain, "RETRACTED: A robust kernelized intuitionistic fuzzy c-means clustering algorithm in segmentation of noisy medical images," ed: Elsevier, 2013.

[21] A. M. Salih and M.Y. Kamil, "Mammography Image Segmentation Based on Fuzzy Morphological Operations," 1st Annual International Conference on Information and Sciences (AiCIS'18), 2018.

[22] L. A. Zadeh, "Fuzzy sets," *Inform. control*, vol. 8, pp. 338-365, 1965.

[23] J. M. Mendel, "A comparison of three approaches for estimating (synthesizing) an interval type-2 fuzzy set model of a linguistic term for computing with words," *Granular Computing*, vol. 1, pp. 59-69, 2016.

[24] D. Yu and S. Shi, "Researching the development of Atanassov intuitionistic fuzzy set: Using a citation network analysis," *Applied Soft Computing*, vol. 32, pp. 189-198, 2015.