International Journal of Engineering & Technology, 8 (1) (2019) 44-52

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 19689
Research paper

Mammogram image segmentation for improving the diagnosis of dense breast issues

Marwah Thamer Ali [1] *, Faiez Musa Lahmood [2], Raad Farhood Chisab [1]

[1] Department of Electrical Techniques, Technical Institute/ Kut, Middle Technical University Baghdad, Iraq
[2] Department of Automobiles, Technical Institute/ Kut, Middle Technical University Baghdad, Iraq
**Corresponding author E-mail: thamer.marwa@yahoo.com*

Abstract

Dense breast tissues can raise the hazard of breast cancer for women because breast of that type has more dense tissues than normal do. It also reduces the effective Mammogram when the produced diagnosed images present lapping intensities for multiple different tissues. However, image segmentation algorithms present the property of separation a whole image into sub sections, which is called regions, and reduce the effects of additive noise caused by artifacts. This paper presents the performance evaluation of Fuzzy segmentation using C-mean algorithms (FCM). Clustering classifier pursued by feature enhancement was the effective way to establish the required results. The last are analyzed to emphasize the accurate segment for diagnosis. The implementation consists of three image samples from real patients executed by the main algorithm with and without adding special information.
The final results confirmed the influence of the investigated work for dense breast diagnosis. The performance evaluation is measured by the statistical significance of mean square error (MSE) for both cases of algorithms. Some specialists have supported the results and promoted it since it offered better visible dense images. In addition, the accuracy of segmentation can be achieved by increasing the number of clustering. Correspondingly, the segmentation besides clustering of dense breast has aided in making a good diagnosis.

Keywords: *Dense Breast; Mammogram; Image Segmentation; Fuzzy C-Mean; Competitive Learning Clustering.*

1. Introduction

Recently, much attention has been drawn to various imaging modalities and becoming more available for acquiring complementary information for variety anatomy. Mammogram is one of these modalities; it uses a low dose of x-ray to scan the inside breast tissues. Furthermore, this advantageous tool is considered to be the most efficacious device for the early find of breast abnormalities in women [1].
The problem in dense breasts lies in the danger of having less fatty composition and non-fattier one compared to the normal breasts. On the other hand, dense breasts have more gland tissues and fibroids. It can be inherited; so the early diagnosis is concerned by women through measuring the tissues' density and thickness. Breast tissues are classified into groups: Mostly fatty, scattered low density, consistent density, and extremely dense [2]. Researches have demonstrated the probability of intensive breasts can develop cancer cells inside and make it harder for diagnosis [3 - 5].
Image segmentation has a significant role in separating an image into sub-parts of different clusters. In general, image segmentation is classified into three methods: region based, edge based, and pixel based [6]. There are two main kinds of classification in all these categories which are supervised and unsupervised. Firstly, the data of classes are defined in advance whereas no need to define it in the unsupervised classification. In addition to that, there are only two types of clustering in pixel based technique which are hard and fuzzy clustering. The first type shows data items is belong to only one cluster while fuzzy clustering belongs its data to the degree of membership to each cluster [7], [22], and [8].

Copyright © 2019 Marwah Thamer Ali et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

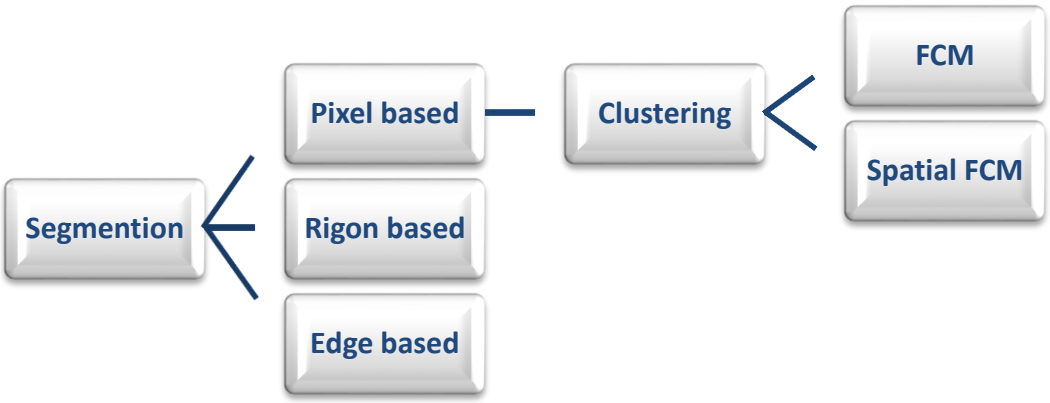


Fig. 1: Classification Categories of Digital Image Segmentation.

Figure 1 shows the three main classification categories of digital image segmentation. The main end of any medical image segmentation is to differentiate anatomical characteristics and separate the component of interest such as tumors, fibroids and glands from image background. The contrast weakness and poor resolution makes the segmentation of computerized medical image is a challenging problem. Furthermore, some of these challenges come from the presence of noise of instrumental limitation and artifacts caused by the patient movement [9], [10], [23], and [24]. The process of identifying multiple groups of related image is called clustering, which can be classified to achieve super pixel resolution. Results quality and efficiency should be established to achieve accurate diagnosis and minimum storage size respectively [11], and [12]. Moreover, in this study, Mammogram image analysis tool is introduced to enable the radiologists in analyzing breast regions and detecting dense masses. As a result, specialists can effectively segment and enhance the required breast image in order to clarify any suspicious dense region.
The concept of this research is illustrated by five sections, the introduction is provided in section 1. Section 2 deals with the mathematical expressions of the presented study. Section 3 describes the procedure of the investigated algorithm and shows the implemented flow chart. Section 4 presents the experimental results. Finally, the conclusion is presented in section 5 with an insight to future work.

2. Methods and theory

In this paper, Mammogram image analysis tool is introduced to enable the radiologists in detecting dense regions and analyzing dense masses. It could effectively segment and enhance the required image in order to clarify any suspicious dense regions. The algorithm behind these works is Fuzzy C-mean (FCM) clustering technique. FCM is a formalized clustering technique that has been used for feature analysis and classification. It plays a significant role in the field of pattern identification and modality recognition. Many fuzzy clustering matters are introduced recently and most of them depend on measuring the distance among pixels and the cluster centers based on distance criteria. The presented algorithm is accomplished by minimizing a function which is called a cost function iteratively to compute fuzzy weights. It concentrates on calculating the centers that minimize the dissimilarity function [13 - 15].
The method of FCM in [16], and [17] is adapted and used to find the optimum cluster centers. To clarify that, a membership function U_{ij} of pixel X_i in the ith cluster V_i is randomly initialized according to the following equation:

$$\sum_{i=1}^{C} u_{ij}=1;\ 0\leq u_{ij} \leq 1\ ;\ 0< \sum_{j=1}^{N} u_{ij} < N \tag{1}$$

Where
N: describe the total pixels in the image.
C: describe the total clusters which are used in algorithm.
The objective function J that is usually used in FCM is given in eq.2.

$$J = \sum_{j=1}^{N}\sum_{i=1}^{C} u_{ij}^{m}\| x_{j-}v\|^2 \tag{2}$$

Where
m: the weighting factor.
To achieve minimum of objective function, cluster centers and membership function are updated iteratively. Euclidean distance between X_j and V_i refers to $\| X_j - V_i \|$. See eq.3 & 4.

$$u_{ij} = \sum_{k=1}^{C}\left(\frac{\| x_j-v\|}{\|x_j-v_k\|}\right)^{\frac{-2}{(m-1)}} \tag{3}$$

Where
k: The weighting factor.
In order to solve the problem of neighboring pixels, spatial information has been added to the standard algorithm which would develop a modified method. The spatial function is defined as h_{ij} which illustrates the likelihood that pixel X_j belong to V_i cluster inside each group. The output value would be large if the pixel neighborhoods belong to the same cluster.

$$h_{ij}=\sum_{k\in NB(x_j)} u_{ij} \tag{4}$$

Where
NB (x_j): a spatial domain squared window settled on one pixel.

The spatial information is introduced to the membership function as u'_{ij}

$$u'_{ij}=\frac{u_{ij}^{p}h_{ij}^{q}}{\sum_{k=1}^{c}u_{ij}^{p}h_{ij}^{q}} \quad (5)$$

Where p, and q are variables to control both the spatial function and membership function. However, the mentioned adjusting parameters play an important role in implementation.
The validity function of fuzzy clustering method is classified based on partition coefficient, partition entropy, and samples structure.
Two more coefficients are used to classify the validity functions which are separating entropy V_{pe} and separating coefficients V_{pe} , these are defined respectively as:

$$V_{pc}=\frac{\sum_{j}^{N}\sum_{i}^{c}u_{ij}^{2}}{N} \quad (6)$$

And

$$V_{pe}=\frac{-\sum_{j}^{N}\sum_{i}^{c}[u_{ij}*\log u_{ij}]}{N} \quad (7)$$

The importance of these functions is to estimate the effectiveness of partitioning which means for better performance, clustering can be established with less fuzziness. The good fuzzy clustering is achieved when $V_{pc} \leq 1$ and $V_{pe} \leq 2$. However, fuzzy clustering based on geometric samples structure computes the samples within one partition and separate samples between different clusters [18, 19, and 20].
In order to illustrate a good segmentation, image enhancement is considered in the first stage using adaptive histogram filtering. Prior investigations have tried several methods for filtering; therefore this part of section will covers those three main types.

1) Adaptive filter: it is also called adaptive histogram equalization (AHE). it works on the image contrast which makes the distribution of gray levels is adaptively distributed along the axis. As a result, the hidden features become more visible especially in homogeneous regions.
2) Median filter: It is a non-linear, spatial, and smoothing filter. It works by changing a pixel by its neighborhoods median.
3) Gaussian smoothing: It is low pass Gaussian filter based on Fourier transform criteria. It is used the image by attenuating the high frequencies and retains the low ones. [21].

3. The simulated FCM method

The investigated algorithm is implemented using Mammogram images with different M x N pixels. The acquired experimental results of the applied method have been considered for three image samples. This section deals with those samples which suffer from dense breast issues in different stages primary, secondary, and extremely dense. Figure 2 presents the flow chart of the implemented FCM algorithm. The discussed FCM segmentation algorithm has been simulated to the following procedure.
Step 1: Reading the input image as png or tiff format.
Step 2: Defining Mammogram image as an uncolored image otherwise using RGB-Gray Scale image conversion function.
Step 3: Pixel by pixel multiplication using input filter function $H(u,v)$ and the input image$I(u,v)$:

$$F(u,v) = I(u,v) * H(u,v)$$

The output resultants F (u, v) from multiplying the input image with the filter function in Fourier domain.
Where I (u, v) is the input image in the Fourier domain, H (u, v) is the filter function, and F (u, v) is the output resultant image.
Step 4: Initialization the input parameters: weighting factor, number of clusters, iterations number, and absolute error.
Step 5: Constructing a vector $x_1, x_2 \ldots x_n$ form the input image.
Step 6: Initializing randomly the membership function of algorithm u_{ij} shown in eq. 3 according to the summation of all values must be equal to one.
Step 7: Minimizing the objective function by updating C and u_{ij} iteratively.
Step 8: Adding spatial information to the membership function by u_{ij} and adjusting p and q parameters to control the function.
Step 9: updating C and choosing $\max_{ij} \{ |V^{(t+1)} - V^{(t)}| \} < \varepsilon$ then end, otherwise substitute t+1 and go to the next step.
By the end of implementation, the results were determined, and the best algorithm for testing image has been chosen.

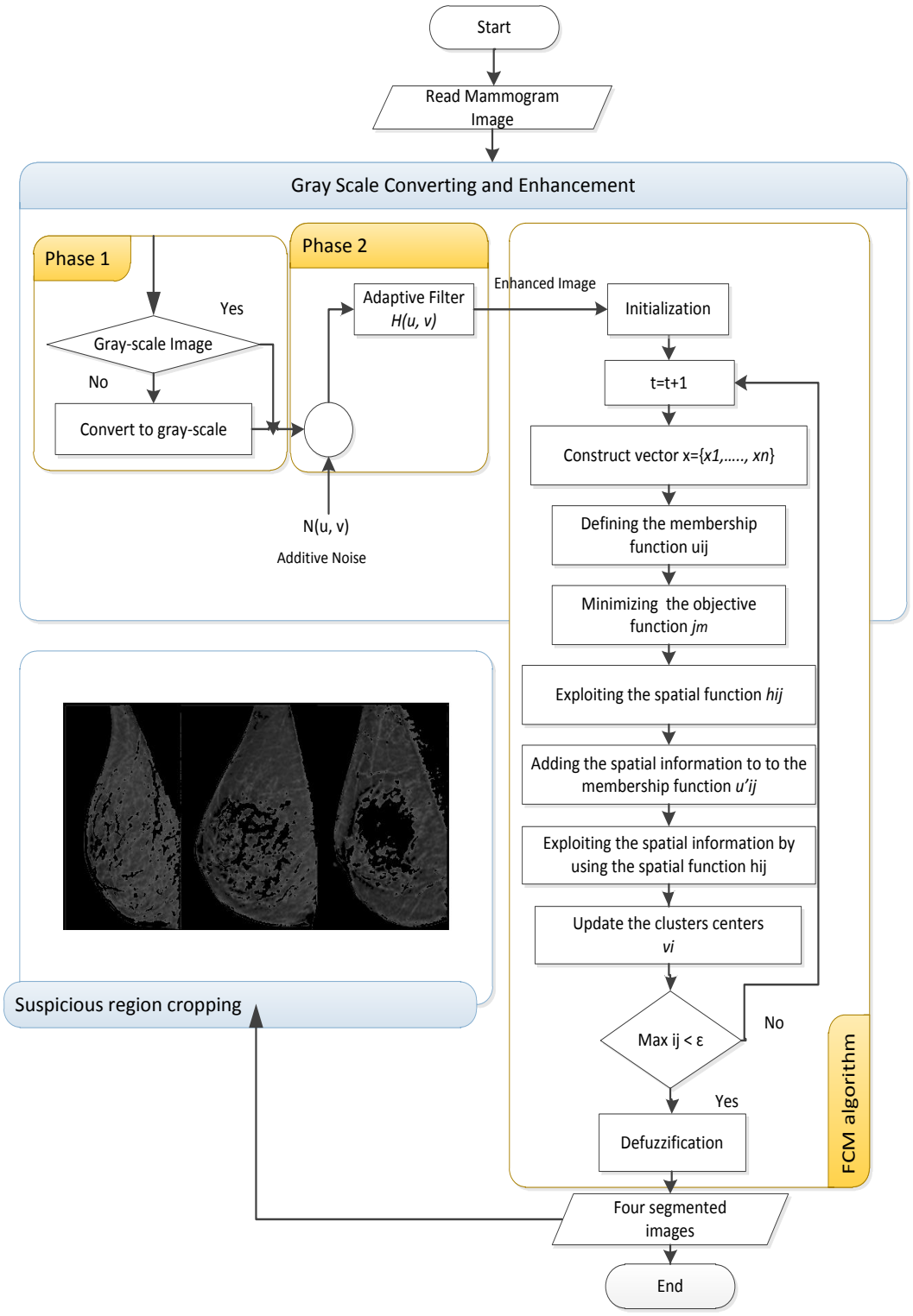


Fig. 2: The Flow Chart of FCM Algorithm.

4. The experimental results

The results of dense breast segmentation are presented in this part of the study which was obtained by applying FCM clustering with and without adding spatial information. Three input Mammogram image samples have been used in to explain the method of segmentation for improving medical image diagnosis. These samples include low, dense, and extremely dense respectively as shown in fig. 4. The mentioned cases have been selected for their significance effects of hiding abnormal complication that happens inside breast tissues.

The applied segmentation algorithm is not just for separating pixels for number of clusters, but also for enhancing the noisy pixel caused by artifacts or other resources. Furthermore, the section of enhancement was also tested by adding Gaussian noise (GN) to these samples which represent the common type of noise that may affect the medical images. Fig.5 shows the mean square error for the input images which have been corrupted by GN. Wiener filter was used to reconstruct the original image and MSE was calculated in order to show the effectiveness of filtering. Filtering part is considered a very significant step regarding the corruption which happens through imaging process.

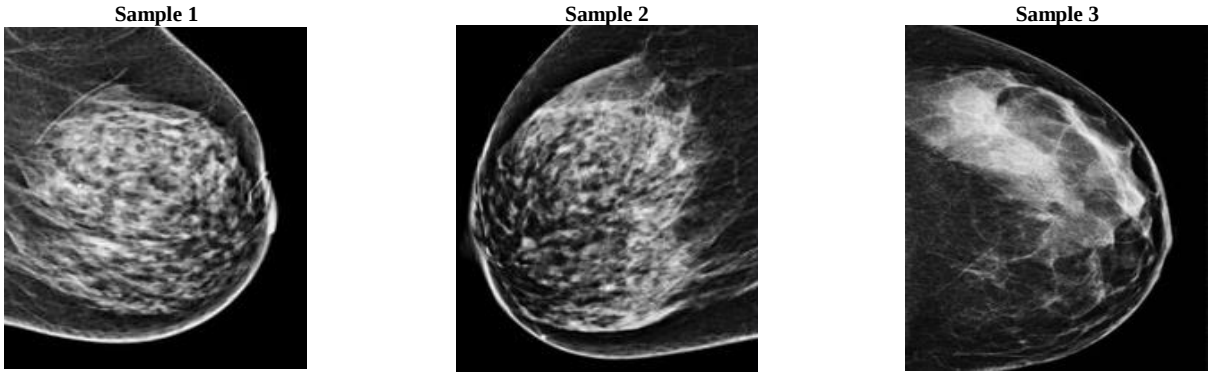


Fig. 3: Mammogram Images Samples.

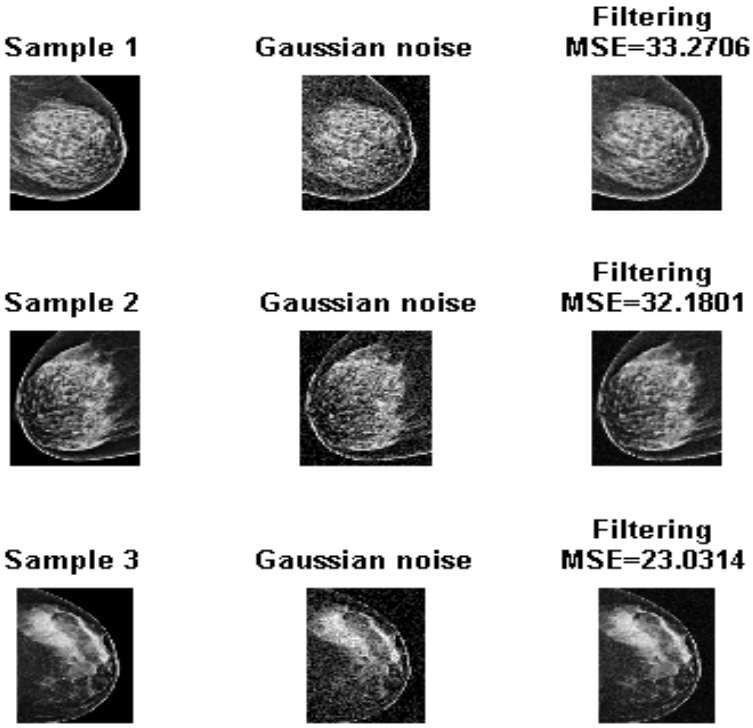


Fig. 4: MSE of FCM Algorithm for Reducing Noise Effects.

There are many filtering approaches have been used in the previous studies, but the adaptive method, which is wiener filter, showed it in comparison with other filters such as median filter. Fig.5 presents the mentioned comparison for random sample.

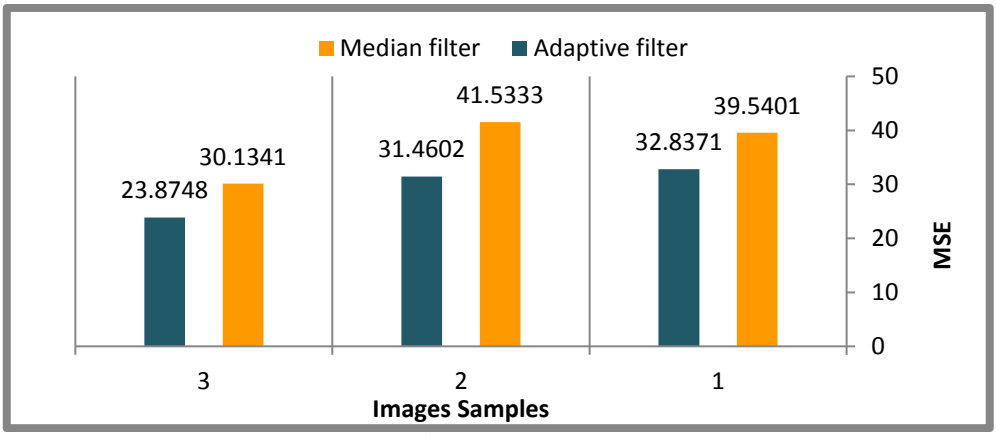


Fig. 5: MSE of Adaptive and Median Filtering.

For initialization, four different gray levels are required to estimate the clusters centers which are 50, 100, 150, and 200 respectively. In this case, these levels are chosen to be updated for partitioning the original image into four segments. Figure 8 shows the clusters centers optimization after number of iterations.

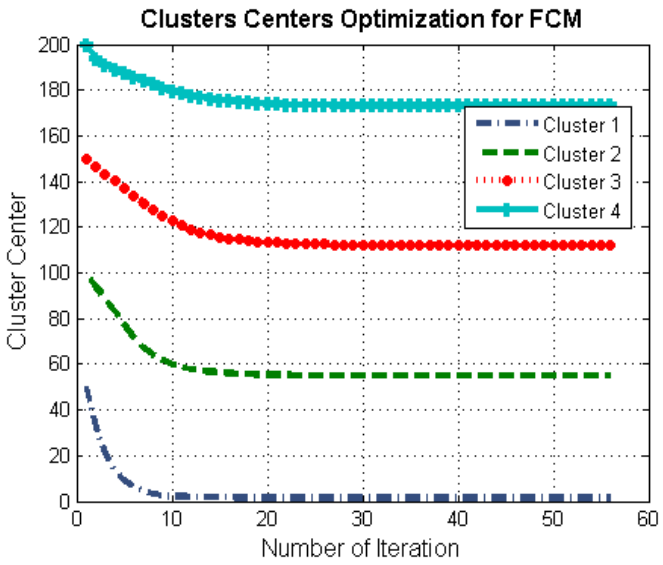


Fig. 6: The Optimized Clusters Centers after 56 Iterations.

Adding spatial information to the algorithm which is called spatial FCM has increased the quality of output segmented image. For the same initialized centers, Fig.8 shows the optimized clusters centers for spatial FCM.

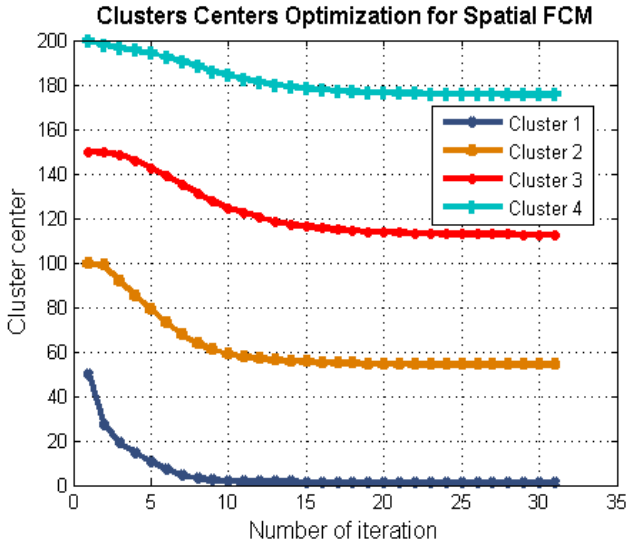


Fig. 7: The Optimized Clusters Centers after 31 Iterations.

The optimized clusters centers of images samples with and without spatial information are shown in Table 1.

Optimized clusters centers for FCM				
Samples No.	C1	C2	C3	C4
Sample1	1.557	54.936	112.190	173.188
Sample2	5.816	40.908	102.587	164.171
Sample3	3.666	41.785	107.511	185.122
Optimized clusters centers for spatial FCM				
Sample1	1.185	54.074	112.570	175.581
Sample2	4.965	40.592	103.373	166.513
Sample3	3.531	42.171	109.115	186.081

The time of optimizing the clusters centers for segmentation algorithm is shown in fig.9

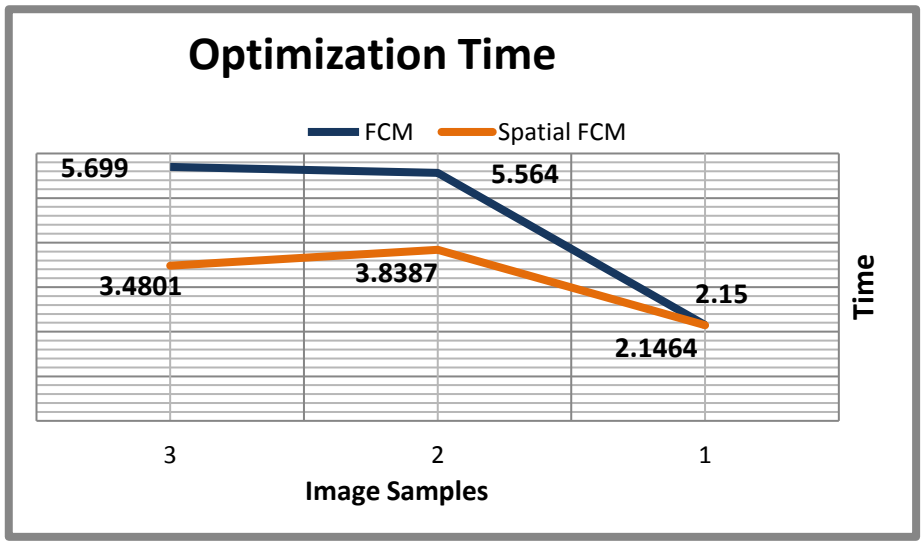


Fig. 8: Clusters Centers Optimization Time.

The output images for these optimized clusters are achieved at the end of section. MSE of segmented image with and without spatial information are shown in Fig.9.

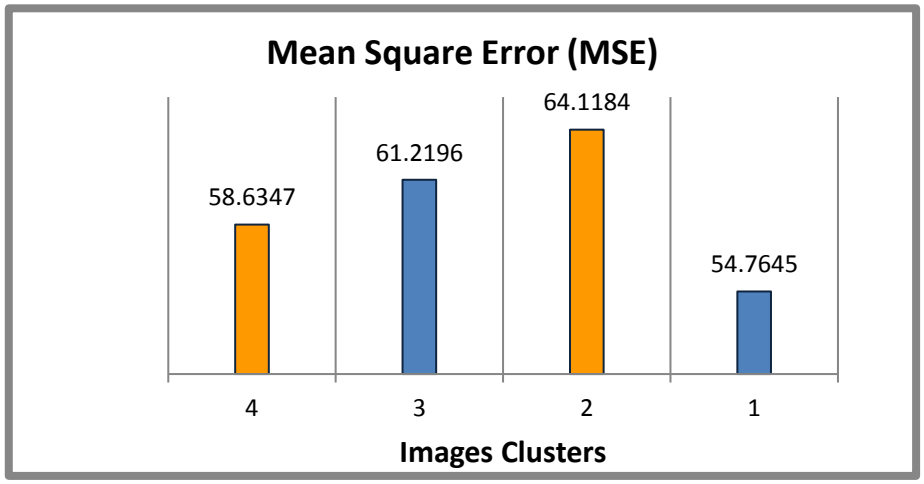
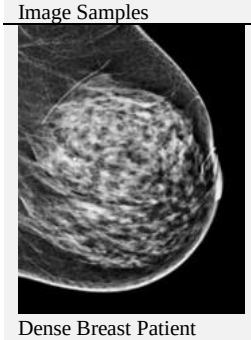
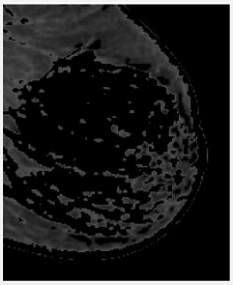
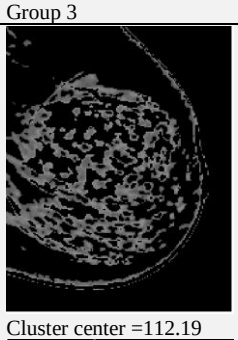
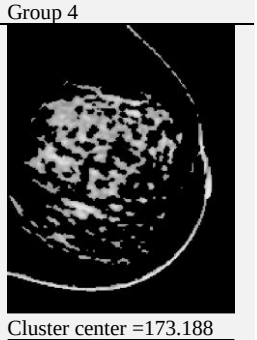
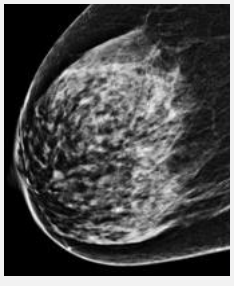
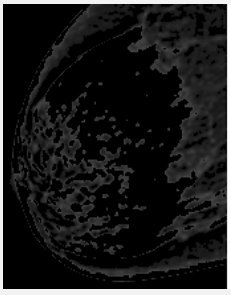
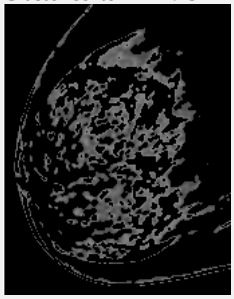


Fig. 9: MSE of Output Images.

Figure 10 and Fig. 11 show the output segmented images of the three dense breast image samples.

Table 1: FCM without Spatial Information

Image Samples	Group 1	Group 2	Group 3	Group 4
				
Dense Breast Patient	Cluster center =1.55	Cluster center =54.93	Cluster center =112.19	Cluster center =173.188
				
High Dense Breast Patient	Cluster center =5.816	Cluster center =40.908	Cluster center =102.587	Cluster center =164.171

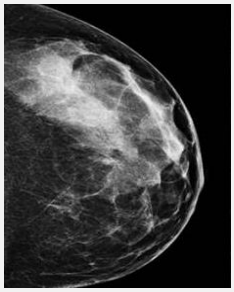
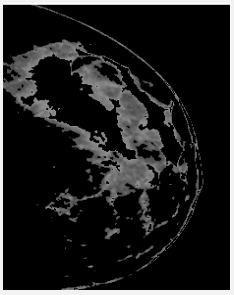
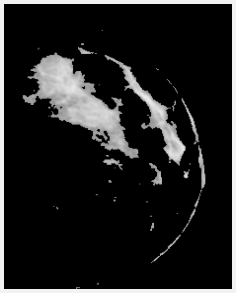
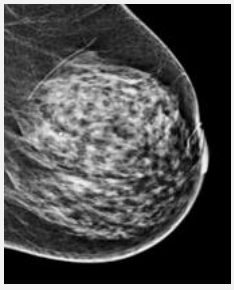
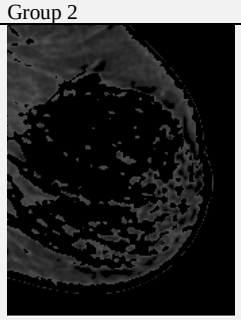
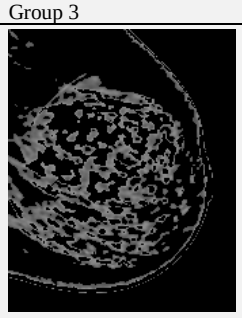
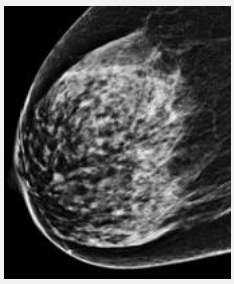
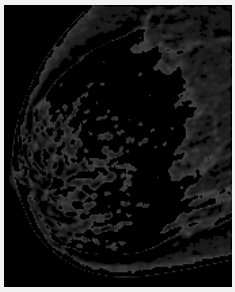
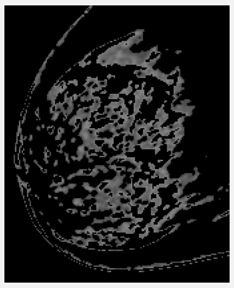
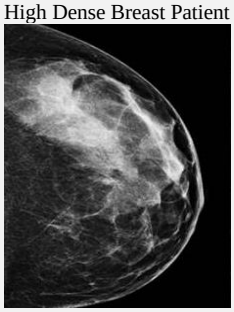
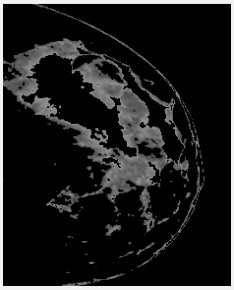
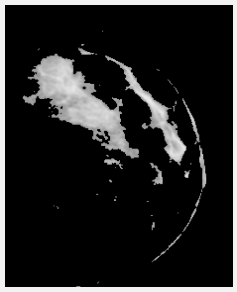
				
Extremely Dense Breast Patient	Cluster center =3.666	Cluster center =41.785	Cluster center =107.511	Cluster center =185.122

Table 2: FCM with Spatial Information

Image Samples	Group 1	Group 2	Group 3	Group 4
				
Dense Breast Patient	Cluster center =1.185	Cluster center =54.074	Cluster center =109.115	Cluster center =175.581
				
High Dense Breast Patient	Cluster center =4.965	Cluster center =40.592	Cluster center =103.373	Cluster center =166.513
				
Extremely Dense Breast Patient	Cluster center =3.531	Cluster center =42.171	Cluster center =109.115	Cluster center =186.081

5. Conclusion

An algorithm for segmenting a dense breast image by using FCM has been proposed in this study. In addition to that, performance evaluation has been conducted with and without adding spatial information. This study indicates that the submitted study is an dynamic method for partitioning breast images into different clusters. Results showed that both FCM+ and spatial FCM were capable of producing the same regions of interest in all selected clusters. However, spatial FCM provided better quality images to FCM. Furthermore, more precision can be achieved by increasing the clusters' number. Therefore, both methods are considered a very effective tool to detect the dense breast areas according to the optimized clusters centers.

References

[1] American College of Radiology (ACR) Breast Imaging Reporting and Data System (BI-RADS) 3rd ed. Reston, VA: American College of Radiology; 2003.

[2] D.W. Kufe, R.E. Pollock, R.R. Weichselbaum, et al., Holland-Frei Cancer Medicine, 6th edition, BC Decker Inc., Hamilton, 2003.

[3] Vachon CM, van Gils CH, Sellers TA, Ghosh K, Pruthi S, Brandt KR, Pankratz VS: Mammographic density, breast cancer risk and risk prediction. Breast Cancer Res. 2007. https://doi.org/10.1186/bcr1829.

[4] McCormack VA, dos Santos Silva I: Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev. 2006. https://doi.org/10.1158/1055-9965.EPI-06-0034.

[5] Boyd NF, Martin LJ, Bronskill M, Yaffe MJ, Duric N, Minkin S. Breast Tissue Composition and Susceptibility to Breast Cancer. JNCI Journal of the National Cancer Institute. 2010.

[6] M. Beltawi, N. Al-Najdawi, S. Tedmori, Mammogram enhancement and segmentation methods: classification, analysis, and evaluation, in: Proceedings of the 13th International Arab Conference on Information Technology, 2012, pp. 477–485.
[7] Paresh Chandra et.al. "MRI image Segmentation using level set method and implement an medical diagnosis system", Computer science & Engineering: An international journal (CSEIJ).Vol.1, No.5, 2011.
[8] Tsai, C.S., Chang, C.C., "An Improvement to image segment based on Human Visual system for object based coding", Fundamental information, Vol.58,No.2,2004.
[9] Chuang, K., Tzeng, H., Wu, J., and Chen, T., 2006, *Fuzzy C-Means Clustering with Spatial Information for Image Segmentation*," Elsevier Ltd. Computerized Medical Imaging and Graphics, vol. 30, pp. 9-15. https://doi.org/10.1016/j.compmedimag.2005.10.001.
[10] M.B.K. Shahedi, R. Amirfattahi, F.T. Azar, S. Sadri, Accurate breast region detection in digital mammograms using a local adaptive thresholding method, in International Workshop on Image Analysis for Multimedia Interactive Services,2007, pp. 26–30. https://doi.org/10.1109/WIAMIS.2007.15.
[11] J. Scharcanski, C.R. Jung, Denoising and enhancing digital mammographic images for visual screening, Comput. Med. Imaging Graph. 30 (2006) 243–254. https://doi.org/10.1016/j.compmedimag.2006.05.002.
[12] S. Sing, R. Al-Mansoori, Identification of regions of interest in digital mammograms, J. Intell. Syst. 10 (2000) 183–217. https://doi.org/10.1515/JISYS.2000.10.2.183.
[13] M. Bhattacharya, A. Das, Fuzzy logic based segmentation of micro calcification in breast using digital mammograms considering multi-resolution, in: Proceedings of the International Machine Visual and Image Processing Conference, 2007, pp. 98–105.
[14] H.D. Cheng, Y.M. Lui, R.I. Freimanis, A novel approach to micro-calcification detection using fuzzy logic technique, IEEE Trans. Med. Imaging 17 (1998) 442–450. https://doi.org/10.1109/42.712133.
[15] H.D. Cheng, H. Xu, A novel fuzzy logic approach to mammogram contrast enhancement, Int. J. Inf. Sci. Appl. 2002 (2002) 167–184. https://doi.org/10.1016/S0020-0255(02)00293-1.
[16] Uri Kroszynski and Jianjun Zhou, Fuzzy Clustering Principles, Methods and Examples, IKS, December 1998.
[17] K. C.Lin, M.Yang, H.Liu, J.Lirng and P.Wang, "*Generalized Kohonen's competitive learning algorithms for ophthalmological MR image segmentation*" Magnetic Resonance Imaging , vol. 21 pp. 863–870,2003. https://doi.org/10.1016/S0730-725X(03)00185-1.
[18] K. Shih, H. Tzeng, S. Chen, J. Wu and T.Chen, "*Fuzzy c-means clustering with spatial information for image segmentation*," *Computerized Medical Imaging and Graphics*, vol. 30, 2006.
[19] M. Eltoukhy, I. Faye, B. Samir, A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation, Comput. Biol. Med. 42 (1) (2012) 123–128. https://doi.org/10.1016/j.compbiomed.2011.10.016.
[20] I. Maitra, S. Nag, S. Kumar, Technique for preprocessing of digital mammogram, Comput. Methods Programs Biomed. 107 (2) (2012) 175–188. https://doi.org/10.1016/j.cmpb.2011.05.007.
[21] R. Gonzalez, R. Woods, Digital Image Processing, 3rd edition, Addison-Wesley Longman Publishing, 2008.
[22] Khusheef, Ahmed Shany. "VISION-BASED OBJECT DISTANCE MEASUREMENT USING MONO AND STEREO VISION." *Iraqi journal of mechanical and material engineering*17.4 (2017): 734-744.
[23] Mnati, Mohannad Jabbar, Raad Farhood Chisab, and Alex Van den Bossche. "A smart distance power electronic measurement using smartphone applications." *Power Electronics and Applications (EPE'17 ECCE Europe), 2017 19th European Conference on*. IEEE, 2017. https://doi.org/10.23919/EPE17ECCEEurope.2017.8099394.
[24] Chisab, Raad Farhood, and C. K. Shukla. "Comparative study in performance for subcarrier mapping in uplink 4G-LTE under different channel cases." *International Journal of Advanced Computer Science and Applications (IJACSA)* (2014).