International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 15 (2023) pp. 189-200

Received: 19 Jan. 2023; Accepted: 21 May, 2023; Published: 9 June, 2023

An Adaptive Weighted Error Sensitive Fuzzy Clustering Technique for Mammogram Image Segmentation

B.K. Chaudhary[1], S. Agrawal[1], P.K. Mishro[1], L. Dora[2], S. Mahapatra[1] and R. Panda[1]

[1] Dept. of EL&TC Engineering, VSS University of Technology,
Odisha-768018, India
mastbkc@gmail.com, agrawals.72@gmail.com, mailpranaba@gmail.com, mahapatra.shakambhari@gmail.com, r_ppanda@yahoo.co.in

[2] Dept. of EEE Engineering, VSS University of Technology,
Odisha-768018, India
lingraj02uce157ster@gmail.com,

***Abstract*: Mammogram image segmentation plays a crucial role in detecting the lesion region in the breast masses. In this context, the key challenging issue is the false positive detection of pectoral muscles or fatty tissues as the lesion region. Further, the presence of noise and imaging errors degrade the segmentation accuracy. To address these problems, we suggest an Adaptive Weighted Error Sensitive Fuzzy Clustering (AWESFC) technique for delineating the different tissue regions in the mammogram images. The basic idea is to incorporate an error sensitive regulating factor in the objective function of the Fuzzy C-means (FCM) algorithm for enhancing the clustering performance in the noisy environment. The suggested technique is experimented with multiple volumes of mammogram images from standard databases. State-of-the-art methods are compared. Quantitative assessment is done using standard evaluation indices. The results indicate better quality with the proposed method.**

***Keywords*:** Error Sensitivity Modelling, Image Segmentation, FCM Clustering, Mammogram, Image Analysis, Unsupervised Clustering.

I. Introduction

As per the report from the WHO, breast cancer is a highly diagnosed disease. In the year 2020, 2.3 million new patients are diagnosed with breast cancer and 685K deaths occurred in the Global scenario. Till the end of the year, 7.8 million women recovered due to early diagnosis. Expert says, the development of abnormalities, such as: scars, fibroadenomas, lipomas, cysts or overgrowth of ducts, etc. are among the most common symptoms of breast cancer. Detecting these abnormalities in an early diagnosis is an important concern for treatment planning. It is a proficient way in increasing the patient's survival rate [1-2][50]. In image-based diagnostics, many imaging modalities such as X-ray, computerized tomography, magnetic resonance, mammography, etc. are used for detecting the breast cancer. Among them, mammography is a trusted modality for early detection of breast cancer. It is reliable due to its low radiation characteristic. However, a false positive region may appear in the mammogram image. It may be due to the presence of pectoral muscle regions in the mammogram images. Further, fatty tissue regions appear abnormal even though there is no lesion region present in the dense breast masses. A suitable retrospective technique is inevitable which can segment different tissue regions inside the breast masses. It can help in analyzing the detection and classification of tissue regions significantly [51].

In the retrospective techniques, mammogram image segmentation is one of the popular methods used for partitioning the tissue regions as well as the lesion mass. In early days, the classical segmentation techniques, such as: region based, and edge-based segmentation techniques are reported for segmenting the mammogram images. In advancement of time, the learning-based techniques are gaining popularity in solving the problem in hand. In recent years, many unsupervised fuzzy techniques are reported in the literature to isolate the tissue region and lesion or mass in the mammogram images. It may be due to the robust performance of the fuzzy techniques in the noisy environment. In addition to this, the fuzzy methods also perform better for the medical images with intensity inhomogeneity in its tissue regions. Fuzzy clustering is a key approach for mammogram image segmentation. It segments the image into multiple clusters, such that each cluster is having homogeneous tissue regions. The clustering is performed using a similarity distance measure. Although, several clustering techniques are reported in the literature, the techniques which can provide accurate and robust results in real-time applications is still inevitable.

In this paper, an adaptive weighted error sensitive fuzzy clustering (AWESFC) approach is suggested for segmenting the tissue regions as well as the lesion regions in the mammogram images. The input mammogram image is first processed through a pre-processing stage for eliminating the pectoral muscle region from the mammogram images. The image is then enhanced using contrast-limited adaptive

histogram equalization (CLAHE) technique. In this stage, the lower order wavelet coefficients only are enhanced to prevent the structural information from blurring. Then the error sensitive regulating factor is estimated from the enhanced image. Then AWSFCM technique is used for segmentation. The proposed method is examined using multiple mammogram images from different databases. The suggested scheme is compared with some recently reported articles. A set of standard image quality assessment measures are used for comparing the quantitative performance of the suggested technique with the state-of-the-art methods. It seems to be a simple and effective solution for segmenting the tissue regions in the mammogram images.

The rest of the paper is progressed as follows: In Section 2, a brief review of the mammogram image segmentation techniques is explained as the related studies. The suggested technique is explained in Section 3. The experimental results are provided in Section 4. Finally, the work concludes in Section 5.

II. Related Studies

In this section, a brief literature review on mammogram image segmentation techniques is presented. Basically, these techniques are grouped into two classes. They are classical techniques, and learning based techniques. In classical techniques, region based, edge based and thresholding based techniques are in practice. With the advent of time, learning based techniques gained popularity due to performance improvement. They are further grouped into supervised and unsupervised techniques. The supervised techniques need a big training data for obtaining an effective accuracy. On the other hand, the unsupervised techniques are popular due to simplicity and efficacy [5]-[9]. In [10], [11], [12], [13] the authors presented surveys on the mammogram image segmentation techniques. From these surveys, it is observed that fuzzy clustering is a promising approach towards mammogram image segmentation. The advantages and disadvantages of these techniques are presented in Table 1.

Table 1. Summary of the various mammogram image segmentation techniques.

Approaches	Advantages	Disadvantages
Region based	These approaches give good result, while the tissue regions present in the mammogram images have connected regions. These methods are simple and easy to implement.	The seed point must be declared before initiating the segmentation process. The segmentation performance is poor due to the anatomical complexity of the breast masses. They take more execution time for accurate segmentation. Further, the accuracy decreases with noisy mammogram images.
Edge based	These approaches work well, while the tissue regions have prominent edges. These methods are simple and easy to implement.	These approaches are delicate in noisy environment. It decreases the overall contrast of the image. The segmentation performance is poor with weak edges in the mammogram images. Further, the methods are not preferable for the mammogram images with smooth edges.
Thresholding based	These methods are simple and cost effective. They perform well for separating the mammogram images from the background pixels.	The performance of these approaches are poor low contrast mammogram images. It is very difficult to process the mammograms whose histogram is unimodal. Further, it is difficult to find the optimal thresholding values, while the number of homogeneous regions increases.
Supervised	These approaches need a big set of labeled data for the training purpose. Larger the size of the training data, more accurate results can be obtained.	Knowledge on the mammograms is needed before the segmentation process. The unavailability of large number of clinical data for training purpose may degrade the performance.
Un-supervised	These approaches can give better result for a single mammogram image. Size of the data set is not a major concern. These approaches are easy to implement. They can segment the mammograms automatically.	It requires the number of clusters before the segmentation process starts. Labeled data are required.

The FCM is a method of grouping homogeneous unlabeled data points in a feature space. It is a prevailing mechanism in segmenting the medical images, specifically while used in detecting the tissue regions in noisy environment. Considering the mammogram image as a feature space with random variation of intensity levels, the basic structure of

FCM clustering is stated as follows:

Let $Y = \{y_1, ..., y_N\}$ be the mammogram image having N number of pixels. FCM clustering is used to partition the N numbers of pixels into k numbers of clusters. The fitness function is stated, as [4]:

$$J_{fcm} = \sum_{j=1}^{k} \sum_{i=1}^{N} u_{ji}^{m} \left\| y_i - v_j \right\|^2 \quad (1)$$

where $\{u_{ji}\}$, $i = 1, 2, ..., N$, $j = 1, 2, ..., k$ are the membership values, indicating the inclusion of a pixel y_i in the j^{th} cluster. The parameter m is a scalar exponent, defining the degree of fuzziness. From the literature, the value of m is chosen to be 2. ||.|| is a norm metric for computing the Euclidean distance. The membership values are calculated using the cluster centers and the image data. The cluster centers and membership values are updated using the following expressions:

$$v_j = \frac{\sum_{j=1}^{k} u_{ji}^{m} y_i}{\sum_{j=1}^{k} u_{ji}^{m}} \quad (2)$$

$$u_{ji} = \sum_{l=1}^{k} \left(\frac{\left\| y_i - v_j \right\|^2}{\left\| y_i - v_l \right\|^2} \right)^{-1/(m-1)} \quad (3)$$

The algorithm founds to be operative for partitioning the data points in a natural image. However, the clinical images are affected by noise and artifacts during imaging and transmission. The mammogram image segmentation performance using FCM only is limited in presence of noise and artifacts. Kannan et al. [5] suggested a FCM based clustering algorithm by modifying its objective function. They replaced the Euclidian distance measure by hyper tangent function for computing the distance between the cluster centers and the data points. Senthilkumar and Umamaheswari [14] suggested a method by combining the local range modification with the FCM (LR-FCM) technique for enhancing and segmenting the lesion from the mammogram images. However, they selected a region of interest (ROI) and labeled mammogram data for enhancement and segmentation. In [15], Gavrielides et al. suggested a fuzzy rule-based technique for suspicious micro-calcification of clusters in mammogram images. The technique is comprising of three stages, partitioning the ROI, histogram-based segmentation and feature extraction. However, the use of ROI selection and local histogram-based thresholding may not give optimal solution. In [16], [17], the authors reported fuzzy rule-based approaches for microcalcification in mammogram images. They used a series of morphological operation for segmenting the mammogram images into different tissue regions. In [18], the authors suggested support vector machine based fuzzy classifier (FCM-SVM) for separating the malignant microcalcifications from the benign microcalcifications. The method used cluster shape and texture features for classifying the mammogram data. Before segmentation, morphological operations are performed to enhance the mammograms. Hizukuri et al. [19] suggested a computer-aided design for partitioning the lesion regions from the mammogram images. The method used eight different thresholding-based features, while classifying the extracted features using artificial neural network (ANN). In [20], [21] the authors integrated soft clustering with ANN (FCM-ANN) classifier for identifying the lesion regions in the mammogram images. The method used FCM technique for delineating the lesion regions in the mammogram images while ANN is used for classifying the mammogram images with lesion regions. In [22], [23], the authors combined FCM clustering with CNN model for risk prediction due to cancer in breast mass. However, the computation time is significantly high as compared to the state-of-the-art methods.

In [24], the authors suggested a threshold based fuzzy technique for mammogram image segmentation. The method used Gaussian filtering for eliminating the noise contaminant in the image. Further, they used morphological operations for partitioning the mammograms into tissue regions. In [25], the authors suggested a particle swarm optimization (PSO) based fuzzy clustering (FCM-PSO) technique for segmenting the tissue regions in mammograms. They used FCM clustering for initial clustering of the data points while PSO algorithm for optimizing the clusters. In [26], Mughal et al. suggested an intensity variation and color-space based model for segmenting the tumor region from the mammograms. The method used color space histogram for extracting the illumination textural features. Morphological models were used for segmenting the tumor regions from the mammogram images. Letizia et al. [27] suggested a feature based FCM clustering techniques for partitioning the microcalcifications in the mammograms. The method used Laplacian of Gaussian filtering before the clustering for denoising. In [28], [29], [30], the authors fuzzy clustering-based techniques for eliminating the pectoral muscle region while segmenting the mammogram images. These methods used morphological operations for enhancing the image quality. Then a fuzzy based clustering is performed on a selected ROI. The performance of these methods depends on the selected ROI and they need continuous human interaction. Toz and Erdogmos [31] combined the multi-thresholding with the FCM based clustering technique for suspicious regions in the mammogram images. In [32], suggested AI based model for resolving the intrinsic errors while segmenting the tumor region in the breast mass.

In [33], the authors suggested a mean clustering and merging technique for detecting the abnormal regions in mammogram images. They used a 2-dimensional median filtering for eliminating the noise contaminant in the image while used gray level cooccurrence matrix (GLCM) based feature for classifying the malicious tissue regions in the mammogram images. The method used decision tree classifier for classifying the healthy and malicious tissue regions. Saleck et al. [34] proposed a FCM based clustering approach for mammographic image segmentation. The method used GLCM based features for obtaining the optimal threshold values while segmenting the tissue regions using FCM clustering. However, the performance of the technique depends on the ROI selection stage. Fazilov et al. [35] suggested a fuzzy rule based technique for mammogram image segmentation. The method used Otsu thresholding in

its pre-processing stage. A fuzzy rule based technique is used of segmenting the abnormal regions in the mammogram images. However, the method is not solving the issues related to pectoral muscle. In addition, the accuracy of the method is poor based on the qualitative assessment.

In [36], [37], the authors used Hausdorff distance in Bustince intuitionistic fuzzy set (IFCM-1) for locating the poor-membership values in the cluster. Roughness measure is used for estimating the upper and lower bounds in the Rough-set. In [38], [39], suggested clustering approaches based on the Atanassov's intuitionistic fuzzy theory for delineating the lesion regions in the mammogram images. In [38], the author used hesitation degree for computing two membership levels of the interval type-2 fuzzy (IFCM-2) set. The method used restricted equivalence function for computing the thresholding values of the interval type 2 fuzzy membership levels. In [39], the method (IFCM-3) used the Zadeh's min t-conorm for computing the membership matrix. In addition, the method used intuitionistic fuzzy divergence and a fuzzy exponential distance for computing the distance among the cluster centers and data points. The discussed approaches used variants of the intuitionistic fuzzy clustering with different measures for estimating the similarity. However, the performance of these approaches shows declining results while dealing with the noisy mammogram images. In [40], Ittannavar and Havaldar suggested an electromagnetism-like optimization algorithm for distinguishing the cancerous and noncancerous regions in a mammogram image. The method used an electromagnetism-like optimization algorithm for distinguishing the cells while a CLAHE technique is used for contrast enhancement before the segmentation stage. In [41], we reported an error sensitivity fuzzy clustering technique. The methodology addressed the noise sensitivity and spatial correlating factor of the data points in the feature space for improving the segmentation performance.

From the above discussion, it is observed that (1) the detection of false positive due to pectoral muscles is a major problem. It should be eliminated before the segmentation stage, so that the segmentation accuracy can be increased. (2) Noise in the mammogram image is a challenging issue while using FCM based clustering. (3) Accuracy of segmentation is a major challenge while segmenting the medical images. To address these issues, an AWESFC technique is designed by modifying the conventional FCM clustering algorithm.

III. Proposed Methodology

A schematic block diagram of the proposed methodology is shown in Figure 1. It comprises of four sages, (i) pre-processing, (ii) enhancement, (iii) parameter estimation and (iv) segmentation using AWSFCM technique.

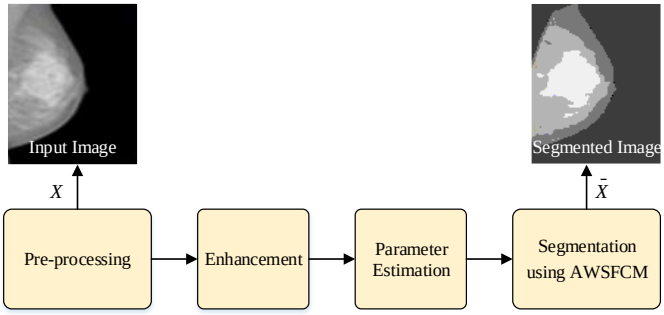


Figure 1. Schematic representation of the proposed adaptive weighted error sensitive fuzzy clustering approach.

A. Pre-processing

Pectoral muscle is a major cause of resulting the false positive region in the mammogram image segmentation. Therefore, it is essential to eliminate pectoral muscle region, if present in the mammogram images. A block diagram-based representation of the pectoral muscle elimination process is shown in Figure 2. It acts as a pre-processing stage before the mammogram image segmentation procedure.

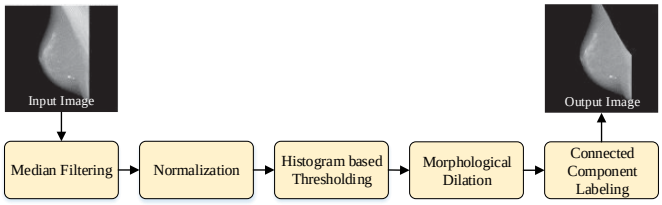
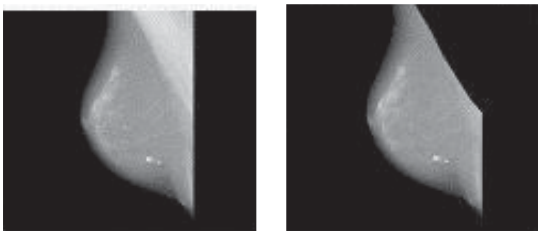


Figure 2. Schematic representation of the pectoral muscle elimination as a pre-processing stage.

A median filter is used for noise removal without blurring the tissue edges. The intensity values are than normalized for avoiding the brightness differences between the left and right mammogram images. The pectoral muscle seems to be a brighter triangular patch near the chest wall and at the lowermost breast mass. Before removing the pectoral muscle, its location must be determined. For this, a simultaneous search for non-zero pixels is conducted from the right top corner and the left top corner. The location where the first non-zero pixel appears is noted. The pectoral muscle is located in the left top corner if it is closest to that corner; otherwise, it is located in the right top corner. A histogram-based thresholding is employed for distinguishing the pectoral muscle from the breast mass. The global optimum in the histogram is selected as the threshold value. The intensity values less than the thresholding values are assigned as black gray levels, and other intensity values are assigned as white gray levels. For linking the gaps in the binary image, a morphological dilation operation is conducted. Connected component labelling is performed on the morphologically dilated image using 8-pixel connectivity. Figure 3 shows the input mammogram image and the resulting pectoral muscle eliminated mammogram image.



(a) (b)

Figure 3. Mammogram images in the pre-processing stage, (a) input image, (b) pectoral muscle eliminated image.

B. Contrast Enhancement

It is an essential stage before the segmentation procedure. Firstly, the subject image is decomposed into four sub-bands using the stationary wavelet transform (SWT) as shown in Figure 4.

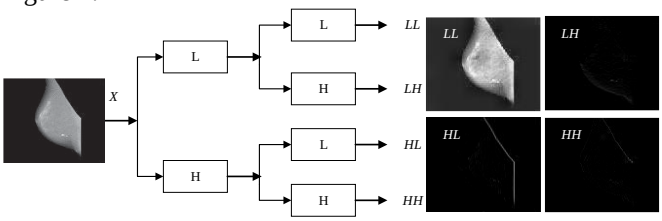


Figure 4. Stationary wavelet transform-based decomposition of the mammogram image.

The lower order wavelet coefficients (*LL*) only are enhanced using CLAHE technique, while the higher order wavelet coefficients (*LH*, *HL*, *HH*) are isolated. The isolation of the higher order wavelet coefficients helps in preventing the structural information from blurring due to enhancement. The enhanced wavelet coefficients are recombined with the higher order wavelet coefficients using inverse wavelet transform as shown in Figure 5.

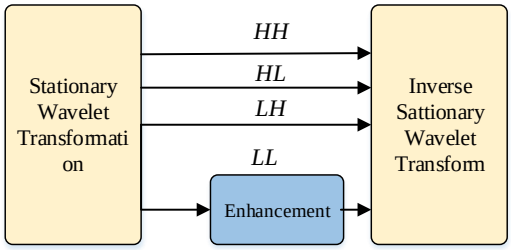


Figure 5. Contrast enhancement in stationary wavelet transform.

C. Error Sensitive Regulating Factor Estimation

Now the error sensitive regulating factor is estimated from the enhanced image. The following paragraph explains the approach of computing the above factors.

Let Y is the subject image with N number of pixels. Considering noise ε in the subject image, it can be represented in (4) as the sum of pristine image X and noise, as:

$$Y = X + \varepsilon \tag{4}$$

The pristine image X is the ideal version of the subject image Y, thus unknown. Here, $\varepsilon = \{e_i : 1,...,N\}$ is the residual error in between the images Y and X. A precise estimation of this error parameter may help in improving the clustering performance of the FCM technique. While considering medical images, the potential noise is usually modeled as Gaussian, Rician or Poisson [42] for simplification. However, the noise is unknown and random in practice. It is hard to define noise in a specific modality. Therefore, we modeled the noise as purely anonymous and a mixture of different noise characteristics, in this work. The pixel intensities of the subject image are expressed in (5) as:

$$y_i = \begin{cases} x_i + e_i & i \in \alpha_1 \\ r_i' & i \in \alpha_2 \end{cases} \tag{5}$$

where, α_1 is the observable region with image data and anonymous noise. α_2 is the region where the image data is missing from $X + \varepsilon$. It is assumed that each data point is drawn from the whole region α using Bernoulli's trial. The error sensitivity in between the pristine and the subject image is quantified using the error sensitive regulating factor ($\Gamma\varepsilon$). Mathematically, it is computed in (6) as:

$$\Gamma\varepsilon = \|\varepsilon_l\|_{l_2}^2 = \sum_{i=1}^{N} |e_{il}|^2 \tag{6}$$

where l is the channel perspective. For instance, the value of l = 3 for an RGB image. Note that $\|\cdot\|_{\ell_2}$ defines the l_2 norm and $|\cdot|$ indicates the absolute value.

D. Segmentation using AWSFCM

Clustering is a method of grouping homogeneous labeled data points in a feature space. It is a prevailing mechanism in segmenting the mammogram images, specifically while used in detecting the lesion region from the whole mass [4]. Considering the mammogram image as a feature space with random variation of intensity levels, the clustering is stated as follows:

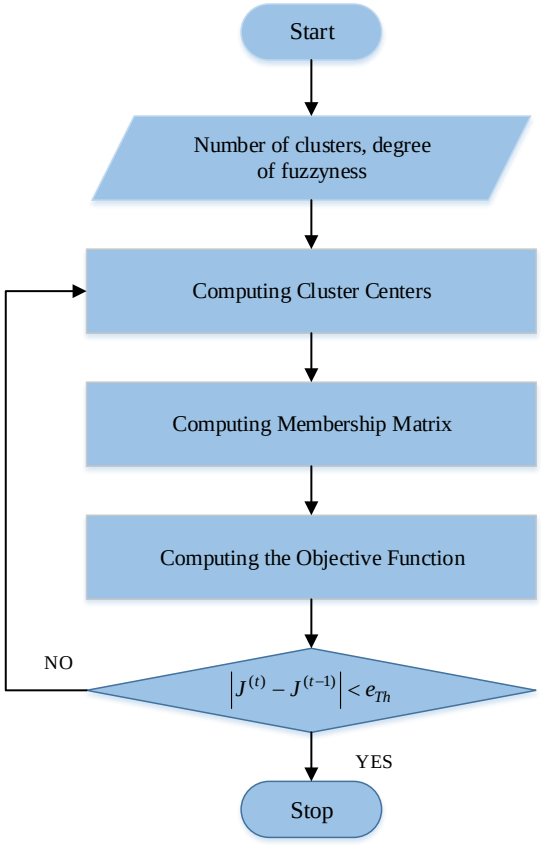


Figure 6. Schematic representation of the segmentation stage using AWSFCM technique.

The error sensitive regulating factor is introduced in

computing the clustering parameters to ensure the minimization of the error sensitivity with the optimization of the clustering algorithm. The accuracy of the segmented regions is significantly affected due to the presence of noise in the medical images. In other words, the segmentation performance of FCM clustering can be significantly improved, if noise can be estimated from the subject image.

To do so, the error sensitive regulating factor is introduced into the objective function of the conventional FCM algorithm. Let N data points in the image $Y = \{y_i\}$ are to be partitioned into k number of clusters. The objective function J_{AWESFC} of the proposed AWESFC approach, including the modification, is defined in (7) as:

$$J_{AWESFC} = \sum_{j=1}^{k}\sum_{i=1}^{N} u_{ji}^{m} \left\| y_i - e_i - v_j \right\|^2 + \beta\, \Gamma(\varepsilon) \tag{7}$$

where, u_{ij} is a membership value in a partition matrix of size $k \times N$ with constraint conditions in (8), such as:

$$\sum_{j=1}^{k} u_{ji} = 1 \quad \forall\, j \text{ and } 0 < \sum_{i=1}^{N} u_{ji} < N \quad \forall\, i \tag{8}$$

Here, β is a controlling parameter that regulates the effect of the error sensitive regulating factor on the objective function. Considering the subject image having red-green-blue channels, the error sensitive regulating factor is defined in terms of the channel perspective in (9) as:

$$\beta\, \Gamma(\varepsilon) = \sum_{l=1}^{L} \beta_l \Gamma(\varepsilon_l) \tag{9}$$

Here, the value of L=3, defining the number of channels in the subject image. Including the spatial information is an important concern while computing the clustering parameters in the AWSFCM technique. In this work, the spatial information is included as a local spatial similarity term (μ_{ij}) that reflects the homogeneity of a data point in its neighborhood. It is calculated in (10) as:

$$\mu_{ij} = \exp\left(\frac{-\left\| y_j - y_c \right\|^2}{\rho \times \sigma_c^2} \right) \tag{10}$$

where y_c is the central pixel value in the neighboring kernel. ρ is the overall scaling parameter. σ_c is the local homogeneity coefficient in the neighboring kernel. It is calculated in (11) as:

$$\sigma_c = \sqrt{\frac{\sum_{j \in Nc} \left\| y_j - y_c \right\|^2}{N_r}} \tag{11}$$

The local spatial similarity term estimates the homogeneity of a data point in its neighborhood. It also helps in including the structural details in computing the clustering parameters. The term is used with the input image data to form a linear weighted parametric image (ξ_j), defined in (12) as:

$$\xi_j = \frac{\sum_{j \in N_i} \mu_{ij} y_j}{\sum_{j \in N_i} \mu_{ij}} \tag{12}$$

Further, an adaptive weight factor is formulated the confines the cluster center to its actual place. The adaptive weight factor is calculated as:

$$w_{ij} = \left(\xi_j - v_i \right)^{-1} \tag{13}$$

Now, the clustering parameters, i.e., the cluster centers and membership matrix are computed in (14) and (15) from the parametric image that includes the local spatial similarity term.

$$v_{jl}^{(t+1)} = \frac{\sum_{j=1}^{N} w_{ij} \left(u_{ij}^{(t+1)} \right)^m y_{il}}{\sum_{j=1}^{N} w_{ij} \left(u_{ij}^{(t+1)} \right)^m} \tag{14}$$

$$u_{ji}^{(t+1)} = \frac{\left(\left\| \xi_j - v_i^{(t)} \right\|^2 \right)^{\frac{-1}{m-1}}}{\sum_{q=1}^{c} \left(\left\| \xi_j - v_q^{(t)} \right\|^2 \right)^{\frac{-1}{m-1}}} \tag{15}$$

The clustering parameters are updated iteratively and the optimal membership matrix and the cluster centers are obtained using an error threshold value (e_{Th}), such as: $\left| J^{(t)} - J^{(t-1)} \right| < e_{Th}$. The proposed AWESFC technique provides optimal clustering parameters with minimum execution time and more accurate segmented tissue regions.

IV. Results and Discussion

In this section, the result analysis of the proposed approach is done. The proposed mammogram image segmentation approach is simulated using MATLAB. The efficacy of the proposed approach is examined in comparison with the state-of-the-art methods, such as: FCM [5], MFCM [6], LR-FCM [14], FCM-ANN [20], FCM-PSO [25], IFCM-1 [36], IFCM-2 [38], IFCM-3 [39] and ESFC [41]. The suggested technique is experimented using DDSM [43], INbreast [44] MIAS datasets [45], and CBIS-DDSM dataset [46]. These datasets consist of 2620, 410, 322 and 306 mammogram images, respectively. The images in each dataset contains normal, benign and malignant tissue masses. The experimental analysis is performed with multiple volumes of mammogram images from each dataset. The validation of the suggested technique is done using the indices: Jaccard [47], Dice coefficients [48], accuracy [49], specificity [49] and sensitivity [49]. The detail information on the validation metrics and datasets are accessible in the corresponding citations.

Figure 7 – Figure 10 show the subjective assessment of the suggested model. Each figure is representing two different class of mammogram images. The benign mass and the malignant mass are presented in row 1 and row 2 of each figure, respectively. Each row of the figures contains the subject image, the segmented image and the lesion region. Inside the breast mass, there may be the Glandular tissue, fatty tissue and high density lesion regions. In addition to this, pectoral muscle may also be observed. The pectoral muscle appears in the mammogram images due to miss alignment of the scanning machine. In Figure 7, the input and the segmented results for the image samples D1_A_1543_1. LEFT_CC (4) and "D1_A_1171_1. LEFT_CC (4)" are shown. In Figure 8, the input and the segmented results for the image samples 20586934 (23) and 20586934 (45)" are shown. The input and the segmented results for the image samples "mdb134 (35) and mdb148 (52)" are shown in Figure 9. Similarly, the input and segmented images for the benign and

malignant mammogram images from CBIS-DDSM dataset is presented in Figure 10. From the figures, it can be clearly observed that the pectoral muscle regions are eliminated, which is leading to the increase in segmentation accuracy.

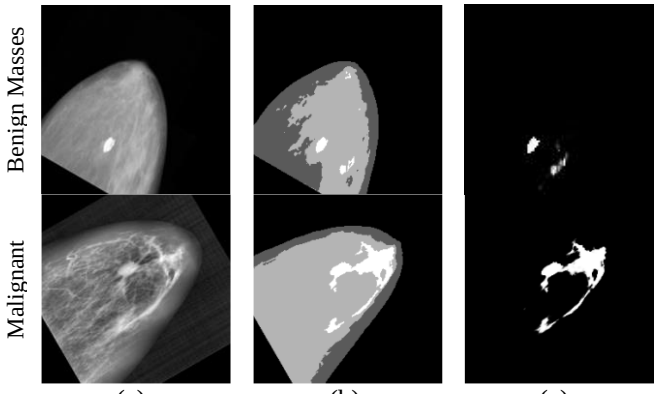


Figure 7. Segmentation results using mammogram images from DDSM dataset. (a) input image, (b) segmented image, (c) lesion region.

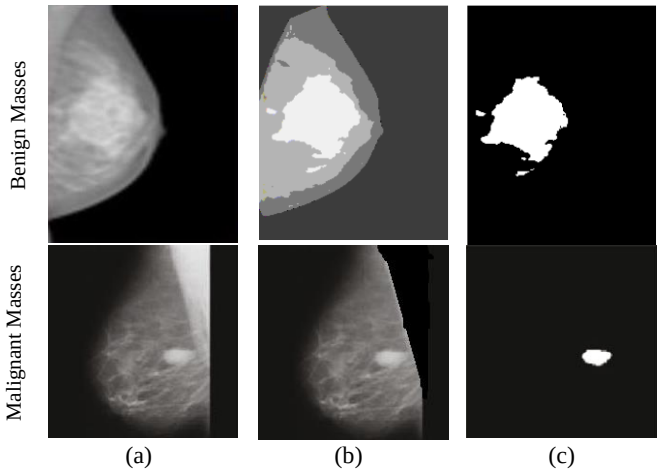
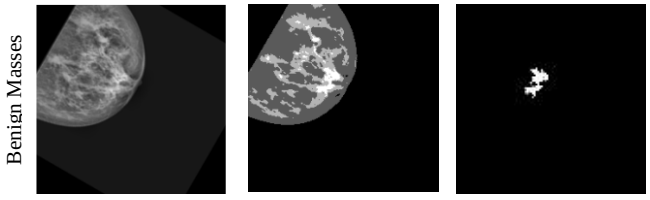


Figure 8. Segmentation results using mammogram images from INbreast dataset. (a) input image, (b) segmented image, (c) lesion region.



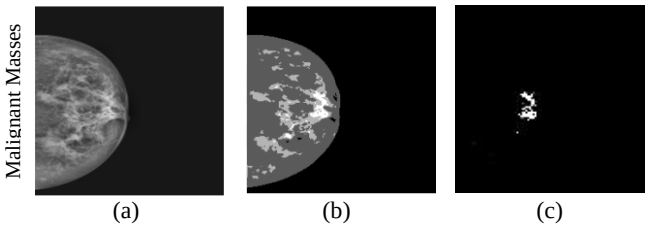


Figure 9. Segmentation results using mammogram images from MIAS dataset. (a) input image, (b) segmented image, (c) lesion region.

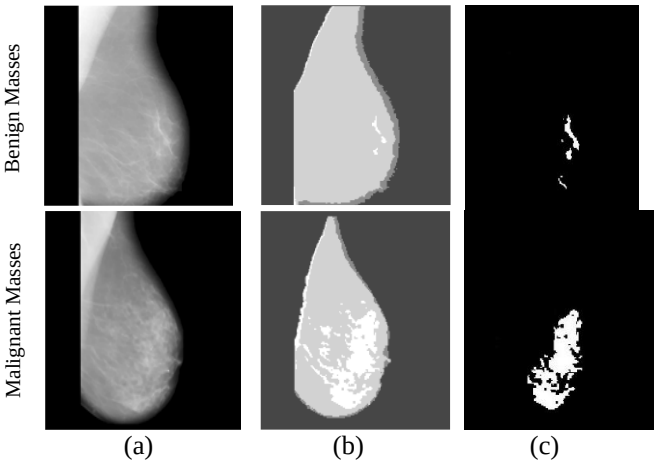


Figure 10. Segmentation results using mammogram images from CBIS-DDSM dataset. (a) input image, (b) segmented image, (c) lesion region.

An objective assessment of the proposed AWESFC technique is shown in Table 2 – Table 5 for DDSM, INbreast, MIAS and CBIS-DDSM datasets, respectively. In each table, one volume of benign and one volume of malignant mammogram images are accessed for comparison purpose. The values in the tables represent average value of the estimated index for one volume of mammogram images. From the tables, it is observed that most of the best-in-class values are obtained using the proposed technique in comparison to the state-of-the-art methods.

Table 2. Performance comparison using Mammogram images from DDSM Dataset.

	Method	Jaccard	Dice	Accuracy	Specificity	Sensitivity
Benign	FCM [5]	0.7710	0.7932	0.8881	0.8812	0.8136
	MFCM [6]	0.7952	0.8151	0.8916	0.8971	0.834
	LR-FCM [13]	0.7642	0.8544	0.9125	0.9092	0.846
	FCM-ANN [19]	0.8214	0.8615	0.9034	0.9144	0.855
	FCM-PSO [21]	0.6745	0.7624	0.8962	0.8895	0.758
	IFCM-1 [30]	0.8648	0.8023	0.8941	0.8542	0.808
	IFCM-2 [32]	0.8313	0.8646	0.8916	0.8651	0.846
	IFCM-3 [33]	0.8464	0.8568	0.9035	0.8648	0.834
	ESFC [35]	0.8568	0.9034	0.9373	0.9289	0.889
	Proposed	**0.8904**	**0.9232**	**0.9424**	**0.9364**	**0.910**
Malignant	FCM	0.8273	0.9152	0.9192	0.8721	0.7701
	MFCM	0.7495	0.8563	0.9105	0.8435	0.6995
	LR-FCM	0.8426	0.8026	0.8825	0.8467	0.8114
	FCM-ANN	0.8496	0.8115	0.9074	0.8726	0.7985
	FCM-PSO	0.7628	0.8739	0.8977	0.8564	0.7167

IFCM-1	0.7547	0.8597	0.8762	0.8722	0.7658
IFCM-2	0.6942	0.8198	0.8535	0.8261	0.6896
IFCM-3	0.8569	0.8149	0.8654	0.8651	0.8013
ESFC	0.8475	0.9354	0.9389	0.9036	0.8025
Proposed	**0.8752**	**0.9428**	**0.9523**	**0.9244**	**0.8346**

Table 3. Performance comparison using Mammogram images from INbreast Dataset.

	Method	Jaccard	Dice	Accuracy	Specificity	Sensitivity
Benign	FCM	0.7842	0.8173	0.8613	0.8264	0.8073
	MFCM	0.8065	0.8245	0.8574	0.8535	0.8095
	LR-FCM	0.7872	0.8594	0.8695	0.8622	0.8124
	FCM-ANN	0.8194	0.8465	0.8846	0.8754	0.8275
	FCM-PSO	0.8465	0.8796	0.8722	0.8466	0.7286
	IFCM-1	0.84221	0.8647	0.8564	0.8614	0.8124
	IFCM-2	0.8164	0.8838	0.8946	0.8797	0.8437
	IFCM-3	0.8347	0.8644	0.8674	0.8924	0.8565
	ESFC	0.8278	0.8585	0.9112	0.8896	0.8496
	Proposed	**0.8546**	**0.8999**	**0.9353**	**0.9244**	**0.9014**
Malignant	FCM	0.8094	0.8501	0.9234	0.8184	0.7815
	MFCM	0.8245	0.9126	0.9345	0.7925	0.7293
	LR-FCM	0.8134	0.8974	0.8786	0.8184	0.7342
	FCM-ANN	0.7847	0.8465	0.8864	0.8314	0.7734
	FCM-PSO	0.7648	0.8272	0.8726	0.7826	0.7052
	IFCM-1	0.8016	0.8514	0.9015	0.8230	0.8012
	IFCM-2	0.7852	0.8466	0.9254	0.8534	0.8454
	IFCM-3	0.8115	0.8945	0.9426	0.8436	0.8235
	ESFC	0.8394	0.9426	0.9344	0.8632	0.7994
	Proposed	**0.8623**	**0.9614**	**0.9725**	**0.9027**	**0.8514**

Table 4. Performance comparison using Mammogram images from MIAS Dataset.

	Method	Jaccard	Dice	Accuracy	Specificity	Sensitivity
Benign	FCM	0.5852	0.6108	0.7794	0.6234	0.4713
	MFCM	0.6865	0.7585	0.86343	0.7495	0.6572
	LR-FCM	0.7434	0.8544	0.8797	0.7865	0.6783
	FCM-ANN	0.6976	0.8213	0.8974	0.7954	0.6976
	FCM-PSO	0.7214	0.8234	0.9155	0.8463	0.7344
	IFCM-1	0.7825	0.8565	0.8944	0.8691	0.7865
	IFCM-2	0.7595	0.8647	0.8672	0.8756	0.8022
	IFCM-3	0.7684	0.8491	0.9044	0.8462	0.7803
	ESFC	0.7242	0.9022	0.9316	0.8955	0.7814
	Proposed	**0.8230**	**0.9115**	**0.9454**	**0.9248**	**0.8242**
Malignant	FCM	0.7594	0.7797	0.7865	0.8542	0.7733
	MFCM	0.7733	0.8385	0.7974	0.8438	0.7034
	LR-FCM	0.8363	0.8715	0.8125	0.8517	0.7642
	FCM-ANN	0.7664	0.8665	0.8797	0.8566	0.7824
	FCM-PSO	0.7855	0.8861	0.8646	0.8677	0.7963
	IFCM-1	0.8013	0.8595	0.8954	0.8948	0.8012
	IFCM-2	0.8344	0.8686	0.9292	0.9127	0.8424
	IFCM-3	0.8525	0.8994	0.9174	0.8944	0.8295
	ESFC	0.8535	0.9047	0.9034	0.8825	0.8271
	Proposed	**0.8992**	**0.9237**	**0.9346**	**0.9372**	**0.8582**

Table 5. Performance comparison using Mammogram images from CBIS-DDSM Dataset.

	Method	Jaccard	Dice	Accuracy	Specificity	Sensitivity
Benign	FCM	0.5569	0.7011	0.7990	0.6643	0.6541
	MFCM	0.6769	0.7862	0.8764	0.7375	0.6836
	LR-FCM	0.7349	0.8796	0.8861	0.7761	0.7038
	FCM-ANN	0.6842	0.8534	0.9026	0.8026	0.7341
	FCM-PSO	0.7346	0.8596	0.9156	0.8615	0.7682
	IFCM-1	0.7462	0.8761	0.8863	0.8764	0.8016
	IFCM-2	0.7649	0.8882	0.8886	0.8919	0.8234
	IFCM-3	0.7739	0.8629	0.9231	0.8561	0.7938
	ESFC	0.7385	0.9341	0.9334	0.9017	0.8139
	Proposed	**0.8383**	**0.9468**	**0.9561**	**0.9346**	**0.8524**
Malignant	FCM	0.7649	0.7924	0.8134	0.8613	0.7863
	MFCM	0.7849	0.8439	0.8234	0.8512	0.7436
	LR-FCM	0.8366	0.8946	0.8469	0.8637	0.7862
	FCM-ANN	0.7764	0.8761	0.8892	0.8674	0.8230
	FCM-PSO	0.7912	0.9034	0.8759	0.8834	0.8213
	IFCM-1	0.8234	0.8761	0.9034	0.9123	0.8346
	IFCM-2	0.8237	0.8813	0.9436	0.9234	0.8653
	IFCM-3	0.8469	0.9134	0.9346	0.9132	0.8462
	ESFC	0.8634	0.9272	0.9234	0.9023	0.8543
	Proposed	**0.9127**	**0.9349**	**0.9436**	**0.9436**	**0.8839**

A graphical assessment of the proposed technique is presented in Figure 11 - Figure 14. The graphs are plotted using the average values of benign and malignant mammogram images from DDSM, INbreast, MIAS and CBIS-DDSM dataset, respectively. The graphs are plotted for presenting a comparison of the proposed technique with the state-of-the-art methods using different evaluation indices. From the graphical assessment, the superiority of the proposed technique can be asserted.

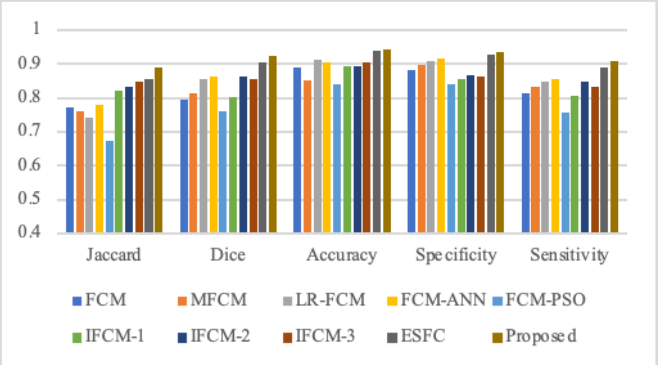


Figure 11. Graphical comparison of different techniques using mammogram images from DDSM dataset.

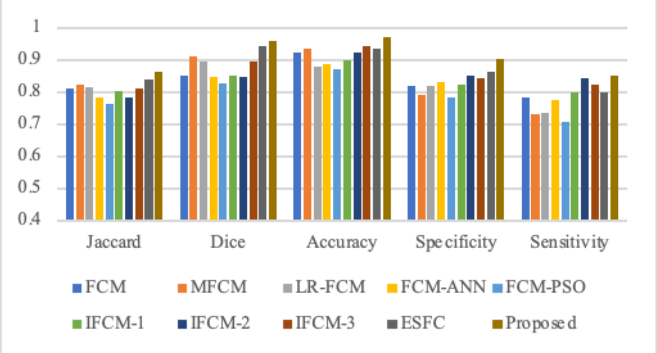


Figure 12. Graphical comparison of different techniques using malignant mammogram images from INbreast dataset.

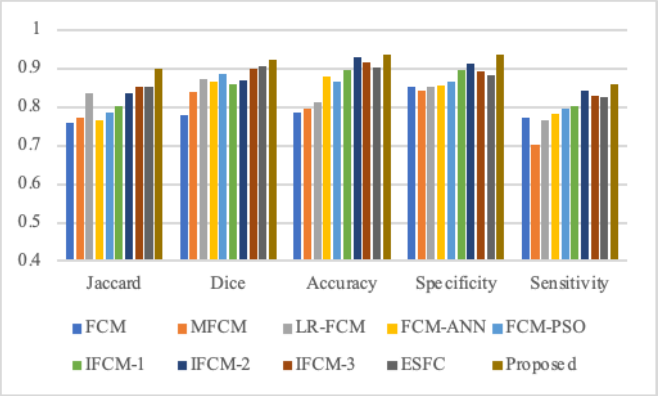


Figure 13. Graphical comparison of different techniques using malignant mammogram images from MIAS dataset.

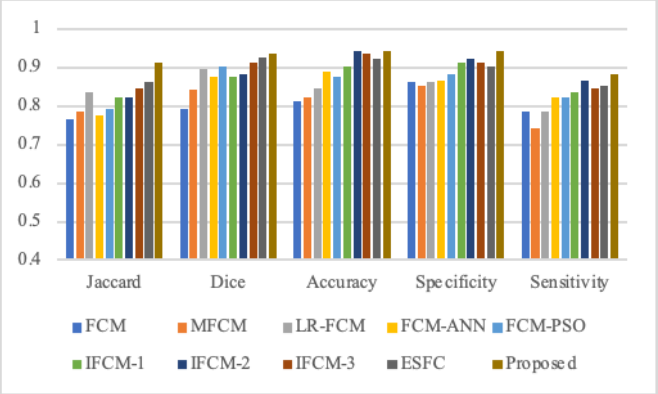


Figure 14. Graphical comparison of different techniques using malignant mammogram images from CBIS-DDSM dataset.

V. Conclusion

In this work, an AWESFC technique is suggested for mammogram image segmentation. The key contributions of the proposed technique are listed as: 1. The method used a pre-processing stage for eliminating the pectoral muscles. Due to the elimination of the pectoral muscles, the false positive rate in the segmented results decreases drastically. 2 The stationary wavelet decomposition of the image isolates the structural details from the enhancement process. Only the lower wavelet coefficient image is enhanced using CLAHE technique. This ensures the enhancement of the mammogram image and preservation of the structural details. The estimation and inclusion of the error sensitive regulating factor ensures the robustness in noisy environment. Further, the one-time estimation of this factor prior to the clustering process reduces the execution time. The clustering process includes the error sensitive regulating factor and the adaptive weighted spatial correlating factor. The inclusion of these factors enhances the segmentation performance as well as reduces the false positive pixels. The suggested technique is experimented with multiple mammogram images from different publicly available datasets. The quantitative assessment shows the superiority of the suggested technique in comparison to the state-of-the-art methods.

References

[1] WHO Report on Breast Cancer: Accessed in May 2022. https://www.who.int/news-room/fact-sheets/detail/breast-cancer#:~:text=In%202020%2C%20there%20were%202.3, the%20world's%20most%20prevalent%20cancer.

[2] A.F. Rositch, K. Unger-Saldana, R.J. DeBoer, A. Ng'ang'a, B.J. Weiner, "The role of dissemination and implementation science in global breast cancer control programs: Frameworks, methods, and examples". *Cancer*. 2020; 126 Suppl 10: 2394-2404 http://www.ncbi.nlm.nih.gov/pubmed/32348574

[3] E. Michael, H. Ma, H. Li, F. Kulwa, and J. Li. "Breast Cancer Segmentation Methods: Current Status and Future Potentials," *BioMed Research International*, 2021, pp. 1-29, 2021.

[4] J.C. Bezdek, L.O. Hall, and L.P. Clark. "Review of MR segmentation technique in pattern recognition", *Medical Physics*, 10(20), pp. 33-48, 1993.

[5] S.R. Kannan, R. Ramathilagam, R. Devi, and A. Sathya. "Robust kernel FCM in segmentation of breast medical images", *Expert Systems with Applications*, 38(4), pp. 4382–4389, 2011.

[6] A. Das, and S.K. Sabut. "Kernelized fuzzy C-means clustering with adaptive thresholding for segmenting liver tumors", *Procedia Computer Science*, 92, pp. 389-395, 2016.

[7] M.S. Yang, and H.S. Tsai. "A Gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction", *Pattern Recognition Letters*, 29(12), pp. 1713-1725, 2008.

[8] S. Chen, and D. Zhang. "Robust image segmentation using FCM with spatial constraint based on new kernel induced distance measure", *IEEE Transaction on Sytems, Man and Cybernetics B*, 34(4), pp. 1906–1917, 2004.

[9] H. Verma, A. Gupta, and D. Kumar. "Modified intuitionistic fuzzy c means algorithm incorporating hesitation degree", *Pattern Recognition Letters*, 122, pp. 45-52, 2019.

[10] S. Padhi, S. Rup, S. Saxena, and F. Mohanty. "Mammogram segmentation methods: A brief review" In *Proceedings of the International Conference on Intelligent Communication and Computational Techniques (ICCT)*, pp. 218-223, 2019.

[11] E. Justaniah, A. Alhothali, and G. Aldabbagh. "Mammogram Segmentation Techniques: A Review", *International Journal of Advanced Computer Science and Applications*, 12(5), pp. 520-529, 2021.

[12] E. Michael, H. Ma, H. Li, F. Kulwa, and J. Li. "Breast cancer segmentation methods: current status and future potentials", *BioMed Research International*, pp. 1-29, 2021.

[13] W. Ansar, and B. Raza. "Breast Cancer Segmentation in Mammogram Using Artificial Intelligence and Image Processing: A Systematic Review", *Current Chinese Science*, 3(1), pp. 3-22, 2023.

[14] B. Senthilkumar, and G. Umamaheswari. "Combination of novel enhancement technique and fuzzy c-means clustering technique in breast cancer detection", *Biomedical Research*, 24(2), pp. 252–256, 2013.

[15] M.A. Gavrielides, J.Y. Lo, R. Vargas-Voracek, and C.E. Floyd. "Segmentation of suspicious clustered microcalcifications in mammograms", *Medical Physics*, 27(1), pp. 13–22, 2000.

[16] R. Erika, E.D.N. Alam, A. Oliver, and R. Zwiggelaar. "Automatic segmentation of microcalcification clusters", in *Medical Image Understanding and Analysis (MIUA)*, 6, pp. 251–261, 2018.

[17] N. Alam, E.R.E. Denton, and R. Zwiggelaar. "Classification of microcalcification clusters in digital mammograms using a stack generalization based classifier", *Journal of Imaging*, 5(76), 2019.

[18] B. Singh, and M. Kaur. "An approach for classification of malignant and benign microcalcification clusters", *Indian Academy of Sciences*, 43 (39), pp. 1–18, 2018.

[19] A. Hizukuri, R. Nakayama, N. Nakako, H. Kawanaka, H. Takase, K. Yamamoto, S. Tsuruoka. "Computerized segmentation method for individual calcifications within clustered microcalcifications while maintaining their shapes on magnification mammograms", *Journal Digit Imaging*, 25(3), pp. 377–386, 2012.

[20] S. Kulkarni, and K.S. Shreedhara. "Stage determination of cancer in mammogram image using soft clustering and ANN", *International Journal of Computer Science and Mobile Computing*, 4 (5), pp. 127–134, 2015.

[21] M. Tardy, and D. Mateus. "Looking for abnormalities in mammograms with self-and weakly supervised reconstruction", *IEEE Transactions on Medical Imaging*, 40(10), pp. 2711-2722, 2021.

[22] W.M. Salama, and M.H. Aly. "Deep learning in mammography images segmentation and classification: Automated CNN approach", *Alexandria Engineering Journal*, 60(5), pp. 4701-4709, 2021.

[23] S. Jha, S. Ahmad, A. Arya, B. Alouffi, A. Alharbi, M. Alharbi, and S. Singh. "Ensemble Learning-Based Hybrid Segmentation of Mammographic Images for Breast Cancer Risk Prediction Using Fuzzy C-Means and CNN Model", *Journal of Healthcare Engineering*, 2023.

[24] A.M. Salih, and M.Y. Kamil. Mammography image segmentation based on fuzzy morphological operations", In *Proceedings of the International Conference on Information and Sciences (AiCIS)*, pp. 40–44, 2018.

[25] N.G. Raju, and P.A.N. Rao. "Particle swarm optimization methods for image segmentation applied in mammography", *International Journal of Engineering Research and Applications*, 3(6), pp. 1572–1579, 2013.

[26] B. Mughal, N. Muhammad, and M. Sharif. "Deviation analysis for texture segmentation of breast lesions in mammographic images", *The European Physical Journal Plus*, 133 (455), pp. 1–15, 2018.

[27] L. Vivona, D. Cascio, F. Fauci, and G. Raso. "Fuzzy technique for microcalcifications clustering in digital mammograms", *BMC Medical Imaging*, 14(1), pp. 1–23 2014.

[28] N. Alam, and M.J. Islam. "Pectoral muscle elimination on mammogram using kmeans clustering approach", *International Journal of Computer Vision and Signal Processing*, 4 (1), pp. 11–21, 2014.

[29] I. Khoulqi, N. Idrissi, and M. Sarfraz. "Segmentation of pectoral muscle in mammogram images using k-means and region growing", *Information Science Letters*, 10(1), pp. 47-57, 2021.

[30] K.A. Gómez, J.D. Echeverry-Correa, Á.Á. Gutiérrez. "Automatic pectoral muscle removal and microcalcification localization in digital mammograms", *Healthcare Informatics Research*, 27(3), pp. 222-230, 2021.

[31] G. Toz, and P. Erdoğmuş. "A novel hybrid image segmentation method for detection of suspicious regions in mammograms based on adaptive multi-thresholding (HCOW), *IEEE Access*, 9, pp. 85377-85391, 2021.

[32] H. Avcı, and J. Karakaya. "A Novel Medical Image Enhancement Algorithm for Breast Cancer Detection on Mammography Images Using Machine Learning" *Diagnostics*, 13(3), p. 348, 2023.

[33] S. Hamissi, and H.F. Merouani. "Novel fully automated computer aided-detection of suspicious regions within mammograms", In *Proceedings of the International Conference on the Innovative Computing Technology (INTECH 2012)*, pp. 153– 157, 2012.

[34] M.M. Saleck, A. El-Moutaouakkil, and M. Mouçouf. "Tumor detection in mammography images using fuzzy c-means and glcm texture features", In *Proceedings of the International Conference on Computer Graphics, Imaging & Visualization*, pp. 121– 125, 2017.

[35] S.K. Fazilov, O.R. Yusupov, K.S. Abdiyeva. "Mammography image segmentation in breast cancer identification using the otsu method", *International Scientific Research Journal*, 3(8), pp. 196-205, 2022.

[36] S.A. Kumar, and B.S. Harish. "A modified intuitionistic fuzzy clustering algorithm for medical image segmentation", *Journal of Intelligent Systems*, 27(4), pp. 593-607, 2017.

[37] Z. Xu, and J. Wu. "Intuitionistic fuzzy C-means clustering algorithms", *Journal of Systems Engineering and Electronics*, 42(4), pp. 580-590, 2010.

[38] T. Chaira. "Intuitionistic fuzzy approach for enhancement of low contrast mammogram images", *International Journal of Imaging Systems and Technology*, 30(4), pp. 1162-1172, 2020.

[39] T. Chaira. "An intuitionistic fuzzy clustering approach for detection of abnormal regions in mammogram images", *Journal of Digital Imaging*, 34(2), pp. 428-439, 2021.

[40] S.S. Ittannavar, and R.H. Havaldar. "Segmentation of breast masses in mammogram image using multilevel multi-objective electromagnetism-like optimization algorithm", *BioMed Research International*, 2022.

[41] B.K. Chaudhary, S. Agrawal, P.K. Mishro, and R. Panda. "An error sensitive fuzzy clustering technique for mammogram image segmentation", in *Intelligent Systems Design and Applications (ISDA)*, 2022.

[42] H. Gudbjartsson, and S. Patz. "The Rician Distribution of Noisy MRI Data", *Magnetic Resonance Med*, 34(6), pp. 910–914, 1995.

[43] S.L. Rebecca, G. Francisco, H. Assaf, R. Daniel. "Curated Breast Imaging Subset of DDSM", *The Cancer Imaging Archive*, 2016. https://doi.org/10.7937/K9/TCIA. 2016.7O02S9CY

[44] I.C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M.J. Cardoso, J.S. Cardoso. "INbreast: toward a full-field digital mammographic database", *Acad Radiol*, 19(2), pp. 236-248, 2012.

[45] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts. "Mammographic Image Analysis Society (MIAS) database v1.21, 2015. https://www.repository.cam.ac.uk/handle/1810/250394.

[46] R.S. Lee, F. Gimenez, A. Hoogi, K.K. Miyake, M. Gorovoy, and D.L. Rubin. “A curated mammography data set for use in computer-aided detection and diagnosis research”, *Scientific Data*, 4, p. 170177, 2017. https://doi.org/10.1038/sdata.2017.177.

[47] P. Jaccard. “The distribution of the flora in the alpine zone”, *New Phytol*, 11 (2), pp. 37–50, 1912.

[48] L.R. Dice. “Measures of the amount of ecologic association between species”, *Ecology*, 26 (3), pp. 297–302, 1945.

[49] K. Chu. “An introduction to sensitivity, specificity, predictive values and likelihood ratios”, *Emergency Medicine*, 11(3), pp. 175-181, 1999.

[50] L Dora, S Agrawal, R Panda, A Abraham, Optimal breast cancer classification using Gauss–Newton representation based algorithm, Expert Systems with Applications, 97: 134-145, 2017.

[51] A. Abraham, Intelligent systems: Architectures and perspectives, Recent advances in intelligent paradigms and applications, Springer, 1-35, 2003.