Mammogram image modalities utilizing novelty of fuzzy segmentation for image enhancement of breast cancer identification

Jeevitha V.* and Laurence Aroquiaraj I.
Department of Computer Science, Periyar University, Salem, Tamil Nadu, INDIA
*jeevithapu@periyaruniversity.ac.in

Abstract
Breast Cancer is one of the most dangerous and deadly diseases. It is related to early identification of a mammogram X-ray tool diagnostic, the breast cancerous, non-cancerous and normal tissue identification with radiologist findings. It uses a Mammogram Image Analysis of Society (MIAS) database utilizing benchmark dataset to identify breast cancer with enhancing images. The main methodology of image segmentation utilizing the heart of the methods is K-Means (KMs) based on clustering and classification. There are also findings about image enhancement for multi-models such as the performance of KMs, KM++, GMM, FKM, FCM and FRR. These methods are evaluated for image enhancement.

Image segmentation of machine learning approaches is one of the methods: K-Means based image segments to various methods enhancing statistical measurements of PSNR, SNR, MSE, IoU, DSC, JI. These metrics are image quality metrics. The classification and prediction-based result findings are precision, accuracy, specificity, sensitivity and F-measures. Finally, performance computing with python uses better results for image quality metrics and image segmentation of breast cancer identification.

Keywords: K-Means (KMs), K-Means++ (KM++), Fuzzy K-Means (FKM), Fuzzy C-Means (FCM), Fuzzy Relative Reduct (FRR), Breast Cancer.

Introduction

Mammography segment allows for differentiation among harmless, cancerous and healthy tissue, it is essential for the early detection and diagnosis of breast tumors[1]. Evaluating mammography images can be difficult due to variations in appearance, noise and overlaying anatomical features. These are conventional methods like recognizing edges or threshold[3]. When there is uncertainty about the constraints such as a little variance in distinct tissue forms, they are inadequate for obtaining precise results. Because of its ease of application and efficacy in dividing images into distinct parts based on pixel similarity[4], K-Means clustering emerged as a popular image segmentation technique.

K-Means has limitations, especially when handling complex tissue boundaries and intensity variations, which are typical in medical imaging procedures like mammograms[5]. Fuzzy C-Means (FCM) is a strong alternative to K-Means allowing each pixel to be part of several clusters with different degrees of membership[5]. This is very important in mammogram images, where tumor boundaries are often unclear. FCM addresses the limitations of K-Means by handling uncertain or ambiguous boundaries making it effective in distinguishing between benign and malignant areas[5]. However, it faces limits like as noise sensitivity as well as computing complexity, which limit its use in real-time medical applications[7].

Fuzzy relative reduction with fuzzy set theory is proposed as an approach for improving segmentation accuracy[10]. The most important features like texture, shape as edge information are retained after unnecessary or duplicate features are eliminated reducing the complexity of the dataset[11]. This improves classification performance and computational efficiency for segmentation models[10]. A hybrid strategy combining K-Means clustering as well as fuzzy relative reduction is presented to improve mammography image segmentation. This method uses K-means for initial segmentation and FCM to refine these clusters in regions with unclear boundaries.

By reducing the feature space fuzzy relative reduction ensures that only pertinent characteristics are taken to heart for additional analysis[13]. This improves precision and computing efficiency in the segmentation process which is critical for distinguishing between malignant as well as benign tumors from normal tissue[11]. In terms of accuracy of detecting cancerous breast, this hybrid also reduces false positives and negatives making it valuable as a tool for the diagnosis of radiologists with challenges remaining in aspects including computational complexity real-time processing and image variability.

Real-time approach for mammographic image segmentation uses edge generation based on k-min and fuzzy c-means (FCM) algorithmic clustering to accurately diagnose breast cancer[9]. This method eliminates mammography image noise and expensive computing expenses. The fuzzy C-method allows soft classification improving decomposition results and enhancing the decomposition of complex structures such as tumors with blurred boundaries[9]. Edge detection improves the accuracy of imaging the boundaries between different tissue types in mammography images.

This hybrid methodology is appropriate towards real-time applications in breast cancer diagnostics since it outperforms

https://doi.org/10.25303/2011rjbt2490257

existing methods in terms of separation accuracy processing speed and noise resilience[9]. The techniques are used in clinical situations where prompt and precise mammography analysis are required highlighted by the authors[9]. The novel feature selection approach for the segmentation of mammogram images is by proposing a Fuzzy Rough Set (FRRS) method[15]. It is significant in the medical image analysis process since it provides features that are considered crucial for the enhancement of the performance of segmentation algorithms, where the traditional segmentation algorithms tend to fail due to noisy or irrelevant data affecting the tumor detection accuracy[15].

This combination ensures that only the most relevant features are selected which improve the accuracy of tumor segmentation. The study shows that the method of feature selection[15] based on diffuse relational approximations improves segmentation performance by removing redundant data and focusing on the most discriminative features. Tested on mammography data, the method showed significant improvements in anomaly detection and reduction of false positives[15]. The method improves the efficiency and accuracy of the automatic analysis of mammograms and offers a promising technique for the early detection of breast cancer[15].

A hybrid feature selection approach was proposed for mammogram image segmentation combining fuzzy relative reduction with particle swarm optimization (PSO)[12]. This is important for the detection of breast cancer because it helps in identifying potential abnormalities such as tumors[12]. High dimensionality and irrelevant features might degrade the performance of segmentation algorithms[12]. The authors use fuzzy relative reduction in combination with PSO to improve feature selection, thereby highlighting relevant characteristics while rejecting irrelevant ones[12]. The hybrid methodology considerably increases mammography segmentation precision and productivity resulting in better tumor identification and fewer false positives when compared to previous methods[12].

This research helps to advance the development of automated mammography analysis systems[12]. There are some traditional methods struggling with high dimensional feature space, mostly containing redundant data or even irrelevant data[14]. To achieve a correct diagnosis of breast cancer, feature selection can be the most important method and fuzzy relative reduct is a method applied based on fuzzy set theory to select the most appropriate features from mammogram images[14]. It minimizes the feature set and increases computational performance[14]. Authors said collate newest classifiers like support vector machines and decision trees with the fuzzy relative reduct method in order to improve segmentation[14].

The multi model has both the advantages[14] providing a robust approach to the detection of tumors and abnormalities in mammogram images. The results indicate a substantial improvement in segmentation accuracy while minimizing false positives and ensuring higher reliability of the system in detecting breast cancer using the combination of fuzzy relative reduct with hybrid classifiers[14]. These studies are more effective automated mammogram analysis tools[14]. The suggested a modified K-means clustering approach for mammogram classification with the objective of accurate tumor detection[6]. Conventional K-means clustering has problems with fuzzy borders and susceptibility to initialization particularly in mammograms where tumour regions may be as intense as the tissues that surround them.

To overcome these limitations, we introduced preprocessing steps such as edge detection and histogram equalization to improve segmentation accuracy. Edge detection makes the boundary edges clearer which facilitates the differentiation of areas containing healthy tissue from areas with tumours. By adjusting the image contrary histogram equalization creates a hue dispersion that is better suited for clustering. The improved K-means clustering method greatly enhances segmented reliability particularly in finding tumor spots on mammograms. The clustering algorithm enhancement shows that the fundamental drawbacks of the conventional K-means approach are solved giving more dependable findings with automated breast cancer screening.

This study contributes to attempts to improve[6] the exactness and dependability of automated computers to assist breast cancer diagnosis and emphasizes the necessity of precondition improvements in the tasks of segmentation[6]. Given the intricacy and noise in mammography, it is a crucial step in the correct identification of breast cancer. Fuzzy relative reduct is proposed by the authors which is a method based on fuzzy set theory for managing uncertainty in mammogram images[8]. This approach determines the most important features by their contribution to the segmentation task and retains only the important attributes.

This decreases computational complexity and improves the general efficacy of this separation method. The authors use choosing features with a mixed approach to segmentation that combines the benefits of fuzzy clustering and standard segment approaches. The resulting model can deal with both clear and unclear tumour boundaries it is noise-resistant[8].

The results reveal that fuzzy relative reduct based feature selection greatly increases segmentation accuracy resulting in better abnormality identification and fewer false positives[8]. This work makes an important contribution to autonomous breast cancer detection systems and proposes a potential strategy for more reliable mammography processing[8].

Medical image analysis in modalities, low radiation of x-ray images using the MIAS database such as benign, malignant and normal is seen in the image segmentation (Table 1). Normalization of preprocessing methods with min-max for data processed for tasks in image segmentation based on selection methods is K-Means (KMs). Noise removal of filtering techniques is like utilizing bilateral filters by enhancing images. Image segmentation of proposed methods is: K-Means (KMs), K-Means++ (KMs++), Gaussian Mixture Model (GMM), Fuzzy K-Means (FKM), Fuzzy C-Means (FCM) and Fuzzy Relative Reduct (FRR).

These methods are evaluated by the feature extraction for reduced unwanted features and image segmentation of finding the best models for image enhancement of breast cancer identification. Performance metrics such as MSE, PSNR, SNR, IoU, JI and DSC image quality are evaluated for the best signal frequency of noise removal. Enhanced images and classification metrics would be accuracy, precision, f-measures, sensitivity and specificity for image segmentation with better performance to the best results found from the mammogram images. Fig. 1 provides proposed methods of step-by-step processing in the section of this study.

Material and Methods

A study comparing preprocessing processes utilizing 322 mammograms collected by the Mammogram Image Analysis Society (MIAS). UK studies used a digital dataset with a dimension over 1024 x 1024 pixels and a size of 1 megabyte. The images were taken from 161 patients based on level of severity: benign, malignant and normal[2]. The images were divided into three categories: 61 benign, 54 malignant with 207 normal mammography.

Pre-Processing: K-Means (KMs) is a popular algorithm for image segmentation and is the classification of image and clustering based methods. To preprocess the normalization means pixel values correspond within a standard range usually between 0 and 1 or 0 and 255 as depending on the requirements for further processing. This is an essential phase because it eliminates any distinctions in pixel intensity that can result in better image enhancement. The following formula is used for normalization:

$$Normalized\ Value = \frac{X - X_{min}}{X_{min} - X_{max}} \quad (1)$$

while X is the beginning frequency and X_{min} and X_{max} are the feature's minimum and maximum values respectively.

Bilateral Filtering Method: Bilateral filtering reduced noise removal of smoothing filters. Gaussian functions based on two methods are mainly appropriate, such as spatial proximity and intensity similarity. Gaussian is spatial means more to distant pixels and intensity, similarity contributes less to others, but both methods are filtered by gaussian functions.

Filtering the bilateral by image I and a pixel at position (i, j), filtered value I' (i, j) are:

$$I'(i,j) = \frac{1}{W_p} \sum_{\kappa,\iota} I(\kappa,\iota).exp\left(-\frac{(i-k)^2+(j-l)^2)}{2\sigma^2 s} - \frac{(I(i-k)^2+I(j-l))^2)}{2\sigma^2 r}\right) \quad (2)$$

where (κ, ι) are the neighbouring pixel Coordinates, σ_s is spatial standard deviation, σ_r is intensity range and W_p is Normalization with weights sum to 1.

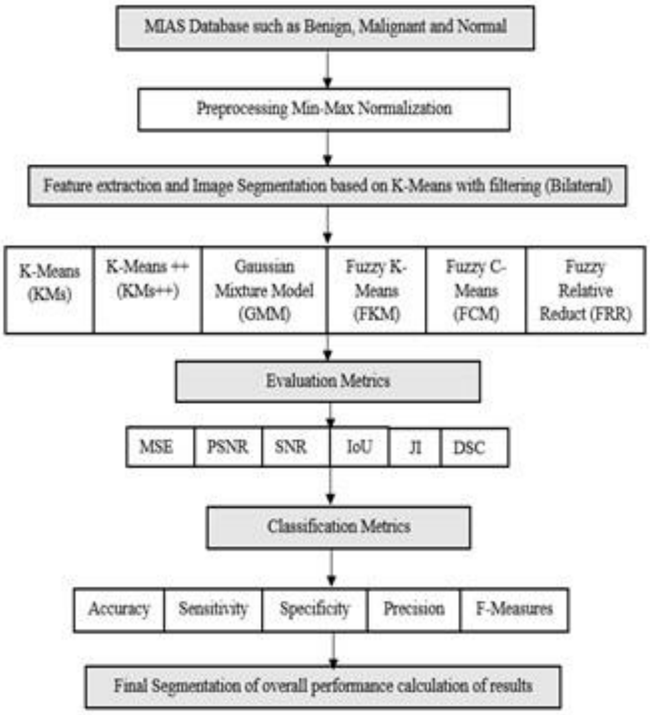

Figure 1: Proposed K-Means (KMs) Image Segmentation based on Fuzzy Relative Reduct with Mammogram Images of Breast Cancer Identification

Image Segmentation with Fuzzy Set: A fuzzy set A is a subset of the universe of discourse X that admits partial memberships. The fuzzy set A is defined as an ordered pair A = {x, μ(x)} where, x ∊ X and $0 \le \mu_A(x) \le 1$. The membership function $\mu_A(x)$ describes the degree to which the object x belongs to the set A. $\mu_A(x) = 0$ represents no membership and $\mu_A(x) = 1$ represent full membership. There are several types of membership functions that characterize A. In this research, we used the Generalized Bell Membership function (GBMF) defined as follows:

$$\mu_A(x) = \frac{1}{1+\left|\frac{x-c}{a}\right|^2} \quad (3)$$

where c is the centre of the membership function, as is the width of the set at the cross-over point and b is the slope of the curve.

K-Means (KMs)
Algorithm 1: K-Means (KMs) Image Segmentation
Input: Parameters are X, K, max_iter, ϵ.
Mammogram Images as Benign, Malignant and Normal data are represented by pixel values. K is number of clusters, max_iteration is 100, ϵ is threshold.

Output: Segmented images processed get k-means images segmentation.

Step 1: Initialize to randomly select *K* pixels from the image as the initial cluster centres μ_k.

Step 2: Provide criteria to nearest centroid over each pixel x_i, give it to the closest centroid μ_k according to the distance calculated by Euclid:

$$distance(xi, \mu k) = \left|\left|x_i - \mu_k\right|\right| \quad (4)$$

Step 3: Upgrade centroids after assigning every pixel and determine new centroids by calculating the mean for every point assigned to each cluster whereas N_k is the number of points in cluster K:

$$\mu_k = \frac{1}{N_k}\sum_{i=1}^{N_k} x_i \quad (5)$$

Step 4: To objective function (minimized) whereas, r_{ik} = 1, pixel i assigned to cluster K, otherwise 0, μ_k is centroid of cluster K:

$$J = \sum_{i=1}^{N}\sum_{k=1}^{k} r_{ik} \cdot ||x_i - \mu_k||^2 \quad (6)$$

Step 5: ϵ is stopping criteria for convergence threshold.

K-Means ++ (KMs++)
Algorithm 2: K-Means++ (KMs++) Image Segmentation

Input: Parameters as X, K, max_iter, ϵ.

Mammogram Images as Benign, Malignant and Normal data are represented by pixel values. K is number of clusters, max_iteration is 100, ϵ is threshold.

Output: Segmented Images processed to get K-Means++ Images Segmentation.

Step 1: Initialize to randomly select *K* pixels from the image as the initial cluster centres μ_k.

Step 2: Estimate Distances over each pixel x_i, determine the squared distance D_i from the closest centre:

$$D_i = \min_{\mu_k \in C} ||xi - \mu_k||^2 \quad (7)$$

Step 3: Choose the next centroid to chance of selecting the next centre is proportional to Di:

$$P(xi) = \frac{D_i}{\sum_{j=1}^{N} D_j} \quad (8)$$

Step 4: Repeat until K centroids are chosen, to assign each pixel to the nearest centroid:

$$r_{ik} = \begin{cases} 1 & if\ pixel\ i\ is\ assigned\ to\ centroid\ k \\ 0 & otherwise \end{cases} \quad (9)$$

Step 5: Recalculate centroids as the mean of assigned pixels:

$$\mu_k = \frac{1}{N_k}\sum_{i=1}^{N_k} r_{ik} \cdot x_i \quad (10)$$

Step 6: To objective function (minimized), Whereas, r_{ik} = 1, if pixel i assigned to cluster K, otherwise 0, μ_k is centroid of cluster K:

$$J = \sum_{i=1}^{N}\sum_{k=1}^{k} r_{ik} \cdot \left|\left|x_i - \mu_k\right|\right|^2 \quad (11)$$

Step 7: The convergence of verify to halt if the centroids do not change substantially (change < ϵ) or max_iter has been reached.

Gaussian Mixture Model (GMM)
Algorithm 3: GMM Image Segmentation

Input: Parameters as Images, k, m, max_iter, ϵ.

Mammogram Images as Benign, Malignant and Normal data are represented by pixel values. k is number of clusters, m is fuzzification parameter, max_iteration is 100, ϵ is threshold for stopping criteria.

Output: Segmented Images processed to get GMM Image Segmentation.

Step 1: Randomly initialize the parameters for K Gaussian distributions: means μ_k, covariances Σ_k and mixing coefficients π_k (weights of each Gaussian).

Step 2: For each pixel x_i in the image, compute the responsibility r_{ik} that each Gaussian component k takes for generating pixel x_i using the current parameters:

$$r_{ik} = \frac{\pi_k N(xi|\mu k,\Sigma k)}{\sum_{j=1}^{k} \pi_j N(xi|\mu j,\Sigma kj)} \quad (12)$$

where π_i is the mixing coefficient, μ_i is the mean and Σ_i is the covariance matrix of the k-th Gaussian component.

Step 3: Update the parameters of the Gaussians (μ_k, Σ_k, π_k) using the responsibilities calculated in the steps:

Mean: $$\mu_k = \frac{\sum_{i=1}^{N_k} r_{ik} . x_i}{\sum_{i=1}^{N} r_{ik}} \quad (13)$$

Covariance: $$\Sigma_k = \frac{\sum_{i=1}^{N} r_{ik}(xi-\mu_k)(xi-\mu_k)^T}{\sum_{i=1}^{N} r_{ik}} \quad (14)$$

Mixing Coefficient: $$\pi_k = \frac{1}{N} \sum_{i=1}^{N} r_{ik} \quad (15)$$

Step 4: Repeat steps 2 and 3 until the change in the log-likelihood or parameters is smaller than ϵ or until max_iter is reached.

Step 5: After convergence, assign each pixel to the cluster with the highest responsibility for that pixel, which corresponds to the Gaussian with the highest r_{ik} for each pixel.

Step 6: Objective function in GMM is the log-likelihood of the data, which is maximized during the Expectation-Maximization (EM) process:

$$L = \sum_{i=1}^{N} \log \left(\sum_{k=1}^{N} \pi_k N(xi|\mu_k,\Sigma_k) \right) \quad (16)$$

where N is the number of pixels, π_k is the mixing coefficient, μ_k is the mean and Σ_k is the covariance matrix for each cluster.

Fuzzy K-Means (FKM)
Algorithm 4: Fuzzy K-Means Image Segmentation

Input: Parameters as Images, k, m, max_iter, ϵ.

Mammogram Images as Benign, Malignant and Normal data are represented by pixel values. k is number of clusters, m is fuzzification parameter, max_iteration is 100, ϵ is threshold for stopping criteria.

Output: Segmented Images processed to get Fuzzy-K-Means Images Segmentation.

Step 1: Initialize to randomly select K pixels from the image as the initial cluster centres.

Step 2: Randomly begin a fuzzy membership matrix U, in which each $u_i k$ is a degree of membership of pixel i to cluster K as well as normalize each row for U so that the total across all clusters exceeds1.

Step 3: To each cluster k, as modify the centre μ_k with a weighted average in pixel values:

$$\mu_K = \frac{\sum_{i=1}^{N} (u_{ik})^m . x_i}{\sum_{i=1}^{N} (u_{ik})^m} \quad (17)$$

Step 4: To membership matrix (U) for each pixel i and each cluster K, update $u_i k$ based on the distance to the cluster centres:

$$\mu_{iK} = \frac{1}{\sum_{j=1}^{k} \left(\frac{||x_i-\mu_k||}{||x_i-\mu_k||} \right)^{2/(m-1)}} \quad (18)$$

Step 5: It changes in U and u_k is smaller than ∈, stop; otherwise, repeat from step 3.

Step 6: FKM Algorithm minimizes to objective function measures of the weighted sum of squared distances between image pixels and their corresponding cluster centres:

$$J(U,\mu) = \sum_{i=1}^{N} \sum_{k=1}^{k} (u_{ik})^m . ||xi-\mu||^2 \quad (19)$$

where u_{ik} is the membership degree of data point x_i to cluster k, μ_k is the centre of cluster k, x_i is the image pixels, m is the fuzzification parameter (typically 2).

Step 7: To set each pixel to the cluster with the highest membership degree.

Fuzzy C-Means (FCM)
Algorithm 5: Fuzzy C-Means Image Segmentation

Input: Parameters as Images, c, m, max_iter, ϵ.

Mammogram Images as Benign, Malignant and Normal data are represented by pixel values. c is number of clusters, m is fuzzification parameter, max_iteration is 100, ϵ is convergence threshold.

Output: Segmented Images processed to get Fuzzy-C-Means Image Segmentation

Step 1: Initialize to randomly select K pixels from the image as the initial cluster centres.

Step 2: Randomly begin a fuzzy membership matrix U, in which each $u_i k$ is a degree of membership of pixel i to cluster K as well as normalize each row for U so that the total across all clusters exceeds1.

$$\sum_{k=1}^{C} u_{ik} = 1 \quad (20)$$

Step 3: To modify the centre μ_K with a weighted average in pixel values:

$$\mu_K = \frac{\sum_{i=1}^{N} (u_{ik})^m . x_i}{\sum_{i=1}^{N} (u_{ik})^m} \quad (21)$$

Step 4: To membership matrix (U) for each pixel i and each cluster K, update $u_i k$ based on the distance to the cluster centers:

$$\mu_{i\kappa} = \frac{1}{\sum_{j=1}^{c}\left(\frac{||x_i-\mu_k||}{||x_i-\mu_k||}\right)^{2/(m-1)}} \quad (22)$$

Step 5: If the change in $\boldsymbol{U}$ and μ_κ is smaller than ∈, or after *max_iter* iterations.

Step 6: FCM Algorithm minimizes objective function, measures of the weighted sum of squared distances between image pixels and their corresponding cluster centres:

$$J(U,\mu) = \sum_{i=1}^{N}\sum_{k=1}^{C}(u_{ik})^2 \,.\, ||xi-\mu||^2 \quad (23)$$

where u_{ik} is membership degree of data point x_i to cluster k, μ_k centre of cluster k, x_i is the image pixels, m is the fuzzification parameter (typically 2), C is the number of clusters, N is the total number of pixels.

Step 7: To each pixel to the cluster with the highest membership degree.

Fuzzy Relative Reduct (FRR)
Algorithm 6: Fuzzy Relative Reduct Image Segmentation

Input: Parameters as Images, k, m, max_iter, ϵ.

Mammogram Images are Benign, Malignant and Normal data represented by pixel values. k is number of clusters, m is fuzzification parameter, max_iteration is 100, ϵ is convergence threshold.

Output: Segmented Images processed to get Fuzzy Relative Reduct Images Segmentation.

Step 1: Fuzzy Relative Reduction (FRR) is used in FRR for Feature selection to eliminate redundant attributes from the image while maintaining classification capacity.

Step 2: Initialize Cluster Centers (μ_k) to choose at random k initial centers for clustering from the image pixel values (or reduced features).

Step 3: Randomly begin a fuzzy membership matrix U, in which each $u_i k$ is a degree of membership of pixel i to cluster K as well as normalize each row for U so that the total across all clusters exceeds 1:

$$\sum_{k\,=\,1}^{C} u_{ik} = 1 \quad (24)$$

Step 4: Initialize Membership Matrix (U) Set up U, guaranteeing each row sums to 1, Initiate to μ_k with a weighted average in pixel values:

$$\mu_\kappa = \frac{\sum_{i=1}^{N}(u_{ik})^m \,.\, x_i}{\sum_{i=1}^{N}(u_{ik})^m} \quad (25)$$

Step 5: Update $u_i k$ based on the distance to the cluster centers:

$$\mu_{i\kappa} = \frac{1}{\sum_{j=1}^{k}\left(\frac{||x_i-\mu_k||}{||x_i-\mu_j||}\right)^{2/(m-1)}} \quad (26)$$

Step 6: If the change in U and μ_κ is smaller than ∈, or after *max_iter* iterations.

Step 7: FRR is minimizes objective function, measures of the weighted sum of squared distances between image pixels and their corresponding cluster centres:

$$J(U,\mu) = \sum_{i=1}^{N}\sum_{k=1}^{k}(u_{ik})^m \,.\, ||xi-\mu_k||^2 \quad (27)$$

where u_{ik} is membership degree of data point x_i to cluster k, μ_k centre of cluster k, x_i is the image pixels, m is the fuzzification parameter (typically 2), k is the number of clusters, N is the total number of pixels.

Step 8: Assign each pixel to the cluster with the highest membership degree.

Results and Discussion

The comparison of various approaches to clustering for mammogram image analysis of fig. 2 is seen as image segmentation into K-Means (KMs), K-Means++ (KM++), Gaussian Mixture Model (GMM), Fuzzy K-Means (FKM), Fuzzy C-Means (FCM) and Fuzzy Relative Reduct (FRR) demonstrating that Fuzzy Relative Reduct (FRR) is better than the others in terms of accuracy and integrity.

The K-Means (KMs) serves as a method of clustering that splits data into K clusters; K-Means++ improves on this through choosing the initial centroids with greater efficiency. GMM represents a probabilistic model that assumes data comes from an assortment of distributions that are Gaussian.

Fuzzy K-Means (FKM) as well as Fuzzy C-Means (FCM) enable measurements to be assigned for multiple clusters, with FCM reducing fuzzy membership objectives.

Fuzzy Rough (FRR) is a technique that blends fuzzy with rough set models to improve clustering accuracy. When IoU, JI and DSC quantify overlap in segmentation tasks, evaluation measures such as PSNR, SNR and MSE evaluate quality and inaccuracy. The model's ability to accurately detect positive and negative occurrences is assessed by precision, accuracy, specificity, sensitivity along with F-measure.

Evaluation Metrics of Mammogram Image Segmentation to K-Means model: Peak Signal-to-Noise Ratio (PSNR) evaluates the quality of the image signals that are based on improved quality because it is the most effective of improving the image.

$$PSNR = 10\, log_{10}\left(\frac{\mathrm{Max}^2}{MSE}\right) \quad (28)$$

Max is Maximum pixel value and MSE is Mean Square Error indicating original and processed images. Signal-to-Noise Ratio (SNR) is greater, the SNR shows an improved indication,

$$SNR = Signal\ Power / Noise\ Power \quad (29)$$

Mean Squared Error (MSE) is a method to calculate the typical level variance between the expected and actual values. It frequently finds its way to task involving regression analysis or restoration.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (30)$$

Therefore, y_i is the real value and $\hat{y}_i$ is the predicted value. Intersection over Union (IoU) determines the similarity among two distinct sets (for instance predicted and actual regions in segmentation). Better overlap is indicated by a greater IoU.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (31)$$

where A and B represent the predicted and true sets respectively.

The Jaccard Index (JI) evaluates the degree of similarity among two sets of data frequently utilized in grouping or segmentation evaluation.

$$JI = \frac{|A \cap B|}{|A \cup B|} \quad (32)$$

Dice Similarity Coefficient (DSC) was an indicator of resemblance among two sets of dice with principles below 1 suggesting more effectively intersect.

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \quad (33)$$

Precision measures the number of those optimistic forecasts have proven proper.

$$Precision = TP/(TP + FP) \quad (34)$$

Accuracy determines the number of right forecasts (which include those true positives and true negatives).

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (35)$$

The models capacity to detect instances of negativity is measured by its specificity.

$$Specificity = TN/(TN + FP) \quad (36)$$

The representations sensitivity also known as recall gauges how well it can detect positive cases.

$$Sensitivity = TP/(TP + FN) \quad (37)$$

The F-Measure (as well as F1 score) integrates both recall and accuracy to one rating in order to evaluate the equilibrium among them.

$$F1 = 2 \cdot \frac{(\text{Precision} \cdot \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (38)$$

Despite previous approaches that depend largely on proximity measures, FRR uses fuzzy logic to limit the effect of extraneous information, allowing for more exact categorization of benign, malignant, as well as normal categories. In light of this, FRR, which stands is very efficient at managing overlapping data and ambiguity for medical image.

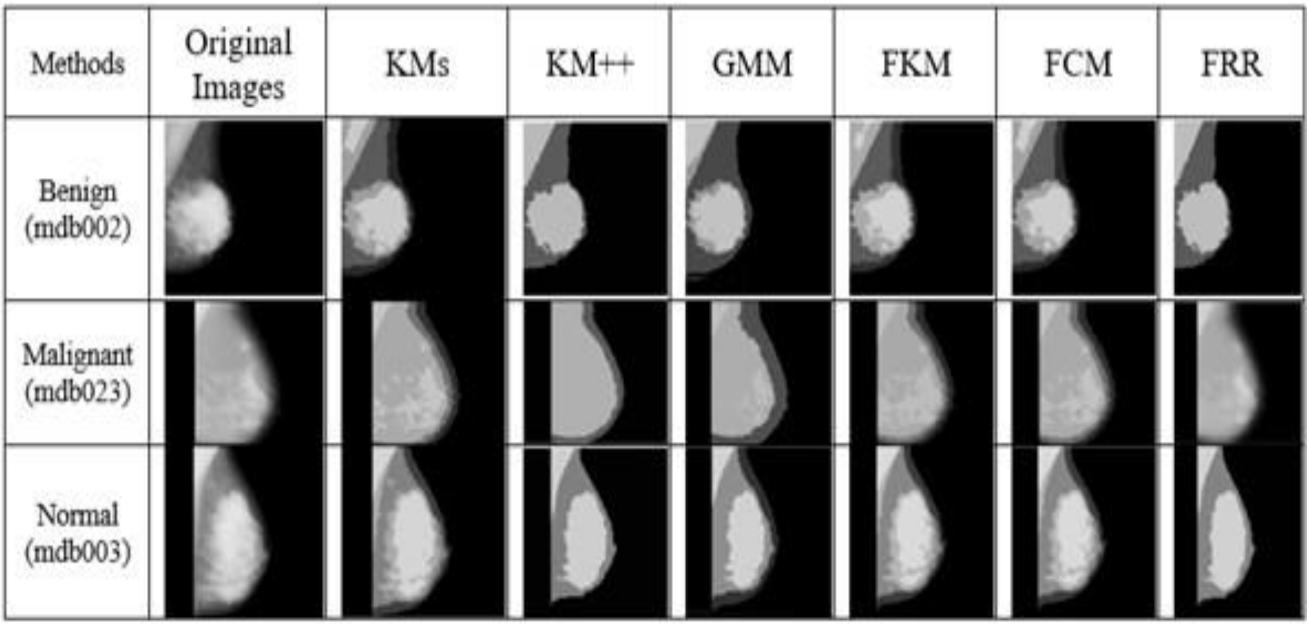

Figure 2: Comparison of K-Means based Image Segmentation Methods with Mammogram Images

Table 1
Comparison of Overall Metrics in Image Segmentation (IS) of Proposed Method

(KMs-IS) Methods	KMs	KM++	GMM	FKM	FCM	Proposed FRR
PSNR	33.53	33.14	33.13	33.73	33.59	37.90
SNR	9.64	9.45	9.44	9.74	9.67	11.83
MSE	29.42	32.17	32.21	27.97	29.00	10.72
IoU	87	62	87	81	90	81
JI	87	62	87	81	90	81
DSC	92	74	91	87	94	88
Precision	96	92	93	96	96	97
Accuracy	98	97.34	97.66	98	98	99.10
Specificity	99	97	98	98	99	99
Sensitivity	96	98	97	95	95	99
F1score	96	94	95	95	95	98

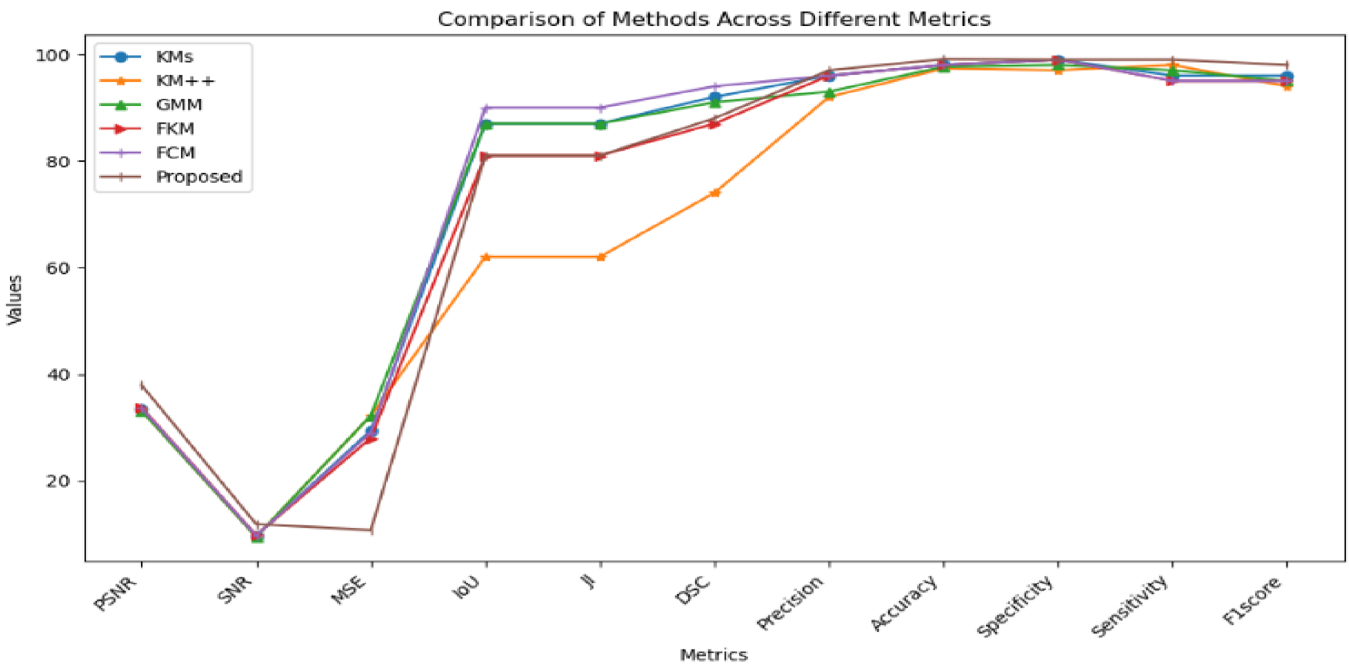

Figure 3: Comparison of KMs Image Segmentation to Overall Metrics Evaluations

When the clustering algorithms of KMs, KM++, GMM, FKM, FCM and FRR are evaluated using several distinct measures, FRR consistently performs better than the others. PSNR (37.90), SNR (11.83), MSE (10.72), Precision (0.97), Accuracy (99.10), Specificity (0.99), Sensitivity (0.99) as well as F-Measure (0.98) represent the important metrics in which it gets the best results. Based on the findings in table 1, FRR performs most well with regard to accurately recognizing positive as well as negative cases generating a minimal amount of noise and provides the most realistic simulations.

FCM, the opposite together outperforms in segmentation related measures like IoU (0.90), JI (0.90) and DSC (0.94), implying especially effective at clustering tasks requiring spatial precision. The remaining methods KMs, KM++, GMM and FKM show a few merits, particularly in particular fields like precision as well as IoU. Thus FRR appears as a more dependable and successful strategy in this analysis.

Conclusion

Medical image analysis based on performace methods k-means and Fuzzy Relative Reduct (FRR) for mammogram image segmentation revealed that FRR outperforms k-means in essential performance metrics. The overall accuracy of the FRR is 99.10%, which is very close to the 98% of K-means which shows that the FRR is better than k-means in correct classification of pixels in an image.

The FRR and k-means PSNRs were 37.90 dB and 33.53 dB respectively. This demonstrates the way an FRR reduces noise and produces a more legible segmentation result. Similar to this, the FRR's SNR is 11.83 dB whereas K-means is 9.64 dB, indicating that the FRR lowers noise levels and produces more accurate results for image enhancement. In terms of Intersection over Union (IoU), the Jaccard Index with Dice Coefficient FRR, surpassed the others; achieving 81% for IoU and Jaccard, as well as 88% for Dice, compared to K-means which obtained 87% for both IoU and Jaccard as well as 92% for Dice.

In medical imaging, improved overlap with FRR is especially important since precise segmentation of certain traits is necessary for accurate diagnosis and therapy forecasting. Therefore, for all aspects of K-means based methods, Fuzzy Relative Reduct has surpassed the previous method, thus it was reliable precise and preferable as an approach for segmenting mammography images efficiently managing the complicated boundaries of classified regions, overlaps and even noise.

Acknowledgement

The authors express their sincere gratitude to Periyar University, Salem, Tamil Nadu, India, for providing University Research Followship (URF) - PU/AD-3/URF Order/14214/23F82123/2023, funds for research in the Department of Computer Science.

References

1. Jagya P. and Dubey R., Comparison of Two Segmentation Methods for Mammographic Image, *International Journal of Computer Applications*, **126(1)**, 31-43 **(2015)**

2. Jeevitha V. and Aroquiaraj I.L., Mammogram Images Using Noise Removal of Filtering Techniques, *International Research Journal on Advanced Engineering Hub (IRJAEH)*, **2(2)**, 118–123 **(2024)**

3. Jeevitha V. and Aroquiaraj I., Novel Medical Image Analysis for Breast Cancer Identification, *International Journal of Education, Modern Management, Applied Science & Social Science,* **6(3(II))**, 1-6 **(2024)**

4. Jeevitha V. and Aroquiaraj I., Mammogram Image Classification for Filtering Methods with 3-Dimensional Fitness Model, *Mukt Shabd Journal*, **14(1)**, 1226-1240 **(2025)**

5. Jeevitha V. and Aroquiaraj I., K-Means of Canny Edge Detection Methods Utilizing Mammogram Image Enhancement, *Indian Journal of Technical Education*, **48(2)**, 16-23 **(2025)**

6. Krasnov D. et al, Fuzzy C-Means Clustering: A Review of Applications in Breast Cancer Detection, *Entropy*, **25(7)**, 1-14 **(2023)**

7. Langarizadeh M., Mahmud R., Ramli A.R., Napis S., Beikzadeh and Rahman W.E.Z.W.A., Improvement of digital mammogram images using histogram equalization, histogram stretching and median filter, *Journal of Medical Engineering & Technology*, **35(2)**, 103–108 **(2011)**

8. Punitha M. and Perumal K., Hybrid segmentation and feature extraction approach to detect tumour based on fuzzy rough-in mammogram images, *Procedia Computer Science*, **165**, 478–484 **(2019)**

9. Ramani R., Valarmathy S. and Vanitha N.S., Breast Cancer Detection in Mammograms based on Clustering Techniques- A Survey, *International Journal of Computer Applications*, **62(11)**, 17-21 **(2013)**

10. Rashed B.M. and Popescu N., Critical Analysis of the Current Medical Image-Based Processing Techniques for Automatic Disease Evaluation: Systematic Literature review, *Sensors*, **22(18)**, 7065 **(2022)**

11. Reddy V.N. and Rao P.S., Comparative analysis of breast cancer detection using K-means and FCM and EM segmentation techniques, *Ingénierie Des Systèmes D Information*, **23(6)**, 173–187 **(2018)**

12. Sharma R. and Kamra A., Enhancing diagnosis of breast cancer through mammographic image segmentation using Fuzzy C-Means, *International Journal of Sustainable Building Technology and Urban Development*, **14(4)**, 488-499 **(2023)**

13. Singh N. and Mohapatra A.G., Breast Cancer Mass Detection in Mammograms using K-Means and Fuzzy C-Means Clustering, *International Journal of Computer Applications*, **22(2)**, 15–21 **(2011)**

14. Srinivas A., Prasad V. and Kumari B.L., Computer-aided diagnosis for early cancer detection using Adaptive Kernel Based Fuzzy Cuckoo Search Optimization Clustering from mammogram images, *Computers & Electrical Engineering*, **103**, 108343 **(2022)**

15. Vivona L., Cascio D., Fauci F. and Raso G., Fuzzy technique for microcalcifications clustering in digital mammograms, *BMC Medical Imaging*, **14(23)**, 1-18 **(2014)**.

(Received 06th May 2025, accepted 09th July 2025)