

Improvement of Segmentation Efficiency in Mammogram Images Using Dual-ROI Method

Venkata Satya Vivek Tammineedi, Koneru Lakshmaiah Education Foundation (Deemed), India*

Raju C., Sri Venkateswara Engineering College, India

Girish Kumar D., Shri Vihnu Engineering College for Women, India

Venkateswarlu Yalla, BVC College of Engineering, India

ABSTRACT

Mammogram segmentation utilizing multi-region of intrigue is a standout amongst the most rising exploration territory in the medical image analysis. The steps engaged with the research are grouped into two kinds: 1) segmentation of mammogram images and 2) extraction of texture features from mammogram images. To overcome these difficulties, a compelling technique is proposed in this paper that comprises of three phases. In the principal arrangement, mammogram images from INbreast database are selected and improved utilizing Laplacian filtering. At that point, the pre-processed mammogram images are utilized for segmentation utilizing modified adaptively regularized kernel-based fuzzy C means (M-ARKFCM). After segmentation, statistical texture FE is connected for recognizing the patterns of cancer and non-cancer regions in mammogram images. Finally, the experimental outcome demonstrated that the proposed approach enhanced the segmentation efficiency by methods of statistical parameters contrasted with the existing operating procedures.

KEYWORDS

Image Segmentation, INbreast Database, Modified Adaptively Regularized Kernel-Based Fuzzy C Means, Texture Features

1. INTRODUCTION

Image segmentation is a basic and essential step in mammography, which is helpful for withdrawal of effectual texture features for mammogram images, Tai et al. 2013. Segmentation provide signs with respect to textures of tumor and non-tumor regions in mammograms, Zhong et al. 2015. Textures are of diverse shapes or noticeable patterns, which is applied for recognizing the textures among malignancy and non-cancer cases for mammogram images in view of the location of inconsistencies, García-Manso et al. 2013. Segmentation based texture is tranquilized of two key advances, such as segmentation of mammogram image and taking out texture features for determined region of mammogram images, Rampun et al. 2017. Region of interest (ROI) is in fact effective support to segmentation that helps to recognize the speculated masses Li, J. B et al. 2012. ROI segmentation is a powerful segmentation, which performs segmentation on different regions of mammogram images Muramatsu et al. 2016. However, ROI can't hold up segmentation for expansive number of mammogram images Beura, S

DOI: 10.4018/IJHISI.305236

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

et al. 2015. Consequently, dual-ROI technique is proposed for inscribing this issue. Dual-ROI based textural features are more eye-catching than the typical ROI based textural features Wang, D et al. 2009.

Additionally, in mammogram images, it is essential to isolate the doubtful cancer region and its ambience Giess, C et al. 2015 Segmentation is one of the key issues for recognizing the masses or relevant objects from the foundation image Kozegar, E. et al. 2017. Usually, the circular based objects are all the more badly characterized masses to defeat this problem a powerful methodology is proposed Berber, T et al. 2013. The proposed approach comprises of three stages: pre-processing, FE and segmentation. At first, the mammogram breast images are gained from the database INbreast. After the procurement of breast images, Laplacian filtering is exploited for denoising or upgrading the quality of breast images. The respective pre-processed mammogram images are used for segmentation by resorting M-ARKFCM. After segmentation, FE is performed utilizing contrast, correlation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, variance, sum average, sum variance, sum entropy, area and difference variance. Finally, the segmented region is contrasted with the stimulated ground truth image to confirm the effectiveness of proposed methodology by the methods of statistical parameters.

This paper is structured as follows. Section 2 gives the overview of a few distinct papers on breast cancer segmentation strategies. In Section 3, statistical texture features are presented with enhanced-ARKFCM segmentation approach. Section 4 demonstrates relative experimental result for proposed breast cancer segmentation artifice using INbreast dataset along with discussions. The conclusion is ended in Section 5.

2. LITERATURE SURVEY

Several approaches are recommended by researchers in breast cancer segmentation framework. In this situation, short assessments of some paramount contributions to the accessible articles are conferred.

Jasmine et al. 2009 presented a paper on a new approach for detecting micro-calcification in digital mammograms employing the combination of wavelet analysis of the image by applying artificial neural networks (ANN) for building the classifiers. The micro-calcification corresponds to high frequency components and the detection of micro-calcification is achieved by extracting the micro-calcification features from the wavelet analysis of the image and we use these results as an input of neural network for classification. The neural network contains one input, two hidden and one output .The system is classified normal from abnormal, mass for micro-calcification and abnormal severity (benign or malignant). The experiments demonstrate that their approach can provide true detection rate approximately 87% and 0 false detection per image which is significant. The evaluation of the system is carried on INbreast dataset.

Mazurowski et al. 2011 specifically, presented a computer-aided detection (CAD) system for mammographic masses that uses a mutual information-based template matching scheme with intelligently selected templates. They presented principles of template matching with mutual information for mammography before. In their paper, they present an implementation of those principles in a complete computer-aided detection system. The proposed system, through an automatic optimization process, chooses the most useful templates (mammographic regions of interest) using a large database of previously collected and annotated mammograms. Through this process, the knowledge about the task of detecting masses in mammograms is incorporated in the system. Then, they evaluate whether there system developed for screen-film mammograms can be successfully applied not only to other mammograms but also to digital breast tomosynthesis (DBT) reconstructed slices without adding any DBT cases for training.

J. Dheeba, et al. 2014 investigated a paper on a new classification approach for detection of breast abnormalities in digital mammograms using Particle Swarm Optimized Wavelet Neural Network (PSOWNN). The proposed abnormality detection algorithm is based on extracting Laws Texture Energy Measures from the mammograms and classifying the suspicious regions by applying

a pattern classifier. The method is applied to real clinical database of 216 mammograms collected from mammogram screening centers. The detection performance of the CAD system is analyzed using Receiver Operating Characteristic (ROC) curve. This curve indicates the trade-offs between sensitivity and specificity that is available from a diagnostic system, and thus describes the inherent discrimination capacity of the proposed system. The result shows that the area under the ROC curve of the proposed algorithm is 0.96853 with a sensitivity 94.167% of and specificity of 92.105%.

M.M. Eltoukhy, et al. 2010 presents an approach for breast cancer diagnosis in digital mammogram using curvelet transform. After decomposing the mammogram images in curvelet basis, a special set of the biggest coefficients is extracted as feature vector. The Euclidean distance is then used to construct a supervised classifier. The experimental results gave a 98.59% classification accuracy rate, which indicate that curvelet transformation is a promising tool for analysis and classification of digital mammograms. After getting the coefficients, a supervised classifier was developed using Euclidean distance. Mammogram breast images from MIAS dataset was availed for examining the projected methodology.

M.M. Eltoukhy, et al. 2012 proffered a method for breast cancer diagnosis in digital mammogram images. Multi-resolution representations, wavelet or curvelet, are used to transform the mammogram images into a long vector of coefficients. A matrix is constructed by putting wavelet or curvelet coefficients of each image in row vector, where the number of rows is the number of images, and the number of columns is the number of coefficients. A feature extraction method is developed based on the statistical t-test method. The method is ranking the features (columns) according to its capability to differentiate the classes. Then, a dynamic threshold is applied to optimize the number of features, which can achieve the maximum classification accuracy rate. The method depends on extracting the features that can maximize the ability to discriminate between different classes. Thus, the dimensionality of data features is reduced and the classification accuracy rate is improved. Support vector machine (SVM) is used to classify between the normal and abnormal tissues and to distinguish between benign and malignant tumors. The proposed method is validated using 5-fold cross validation. The obtained classification accuracy rates demonstrate that the proposed method could contribute to the successful detection of breast cancer.

Elter and Hasslmer 2008 proposed a novel, knowledge-based approach to the computer aided discrimination of mammographic mass lesions that uses computer-extracted attributes of mammographic masses and clinical data as input attributes to a case-based reasoning system. This approach emphasizes a transparent reasoning process which is important for the acceptance of a CADx system in clinical practice. The authors evaluated the performance of the proposed system on a large publicly available mammography database using receiver operating characteristic curve analysis. Our results indicate that the proposed CADx system has the potential to significantly reduce the number of unnecessary breast biopsies in clinical practice.

Bala, B. K., & Audithan, S. 2014 exhibited an efficient classification system for micro-calcification in digital mammogram image. As the early prediction of breast cancer is the key to reduce women mortality The classification of micro-calcification system is presented based on discrete curvelet transform (DCT) and discrete wavelet transforms (DWT). The energy features are extracted from the mammogram images by using aforementioned transformations at various level of decomposition and k nearest neighbour (KNN) classifier is used for classification task. Experimental results show that the DCT based classification system provides satisfactory result over DWT.

Li Y et al. 2016 presented a paper on a mass classification method in mammograms is proposed based on two-concentric masks and discriminating texton. First, the two-concentric masks are employed, dividing each mass region into the centre region and the periphery region. Then integrating linear discriminant analysis (LDA) with traditional texton, the discriminating texton is proposed. The shortage of not considering the class information in traditional texton is improved. Finally, features are extracted with discriminating texton for both the centre region and the periphery region. Thus, the problem of disregarding the spatial layout information is alleviated. The proposed method is tested on 130 mass regions from Digital Database for Screening Mammography (DDSM) database. The

classification accuracy rate reaches 86.92% and the area under the receiver operating characteristics (ROC) curve is 0.91, which is higher than traditional texton and some other texture-based methods. Discriminating texton improved the shortage of not considering class information. Employing the concentric masks alleviated the disregard of the spatial information. The periphery region is more effective to classify masses than the centre region.

Sampaio et al. 2011 presented a computational methodology for detection of masses in mammogram images which can be described in following steps: (1) removing noise and objects outside the boundary and highlighting the internal structures of the breast, (2) regions containing mass are segmented using cellular neural network, (3) Thereafter the shape of these regions are analyzed through shape descriptors, (4) classification of candidate region is classified as masses or non-masses through Support Vector Machine.

Dalmiya et al. 2012 introduced a segmentation method for mammograms using wavelet and k-means clustering. Authors defined their method in following steps: (1) Discrete wavelet transform is used to extract high level details from MRI images, (2) the outputted image is then added to original input image to get sharpened image, (3) k-means clustering is performed on sharpened image to locate the tumor region. Final tumor region is extracted by performing thresholding on clustered image.

Mubarak et al. 2012 differentiated the consistency of the image and represented by quad trees. In quad tree, each node has four descendants and the root represents the entire image. The contribution of this method is to split the image in rapport of our requisition, because the splitting level depends on criteria. The major restriction is, it may create blocky segments. To avoid blocky segment, fragment into better level. But it acquires computation time.

To overcome the above drawbacks, modified-ARKFCM is actualized with statistical texture features for improving the performance of mammogram breast image segmentation.

3. PROPOSED METHODOLOGY

The proposed strategy for segmenting the cancer and non-cancer region from mammogram breast image is separated into three noteworthy steps: image pre-processing, FE and segmentation.

3.1 Image Acquisition and Pre-processing

In the antecedent phase of the mammogram breast cancer segmentation, the mammogram images are taken into account from the established benchmark dataset (INbreast dataset). The original INbreast dataset digitized at 50 micron pixel edge, at that point the database is reduced to 200 micron pixel edge with 1024×1024 pixels. Subsequently by obtaining the mammogram images, an important impression (pre-processing) is carried out using Laplacian filter.

In the proposed framework, Laplacian channel is used as the preprocessing procedure, on the grounds that contrasted with other filters it is easy to cut down the noise. This channel comprises of smoothing operator, which is used to transform the noise image into a noiseless image to achieve the objects from original mammogram images and furthermore Laplacian filter assumes an outstanding role in edge detection of mammogram images.

Assuming, a couple of smooth images in a linear scale space, which is represented as $G_{\sigma_1}, G_{\sigma_2}, \dots, G_{\sigma_n}$. It is formed by convolving the original image G_0 with a Gaussian kernel $K_{\tilde{\lambda}}$. Here, the goal of this examination is to choose the appropriate index (σ) from the scale-space at each pair (x, y) and the paradigm for such choice relies upon the Minimal Description Length (MDL) principle. It furnishes an effective descriptive model with less complexity or minimal. Laplacian filtering is illuminated in (1).

$$G_0(x, y) = G_{\sigma}(x, y) + \epsilon_{\sigma}(x, y) \quad (1)$$

where, ϵ_σ is represented as a residual of smooth image.

However, the optimal model incorporates a predetermined number of bits, which helps to increase the associate information to the model. While, the maximal smoothness with minimum residual is exemplified in (2).

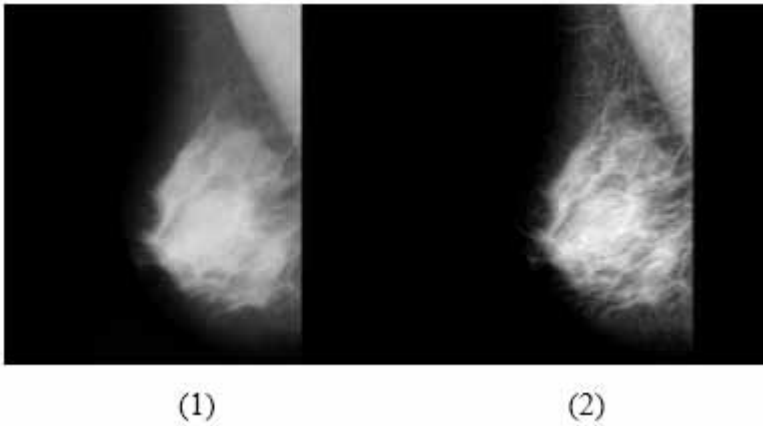
$$dlG_0(x, y) = dlG_\sigma(x, y) + dl\epsilon_\sigma(x, y) \quad (2)$$

Based on MDL principle, select the minimum value dlG_σ at (x, y) . This $mindlG_\sigma$ helps to determine the optimal smoothness (σ^*) in (x, y) . The Laplacian filter is achieved by employing the local variance σ at (x, y) . Each point (x, y) is smoothed as $\sigma^*(x, y)$ and it is given in (3).

$$G_0(x, y) = G_\sigma(x, y) * e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3)$$

where, x is represented as the horizontal axis, y is denoted as the vertical axis, and σ is mentioned as the standard deviation. The sample pre-processed mammogram image is spoken to it in the Fig 1. The respective noiseless images are utilized for segmentation by employing modified-ARKFCM.

Figure 1. 1) Original image 2) Pre-processed image



3.2. Segmentation Utilizing Modified-ARKFCM

The pre-processed mammogram image is used for segmentation; a compelling procedure modified-ARKFCM is attempted for segmenting the tumor and non-tumor regions of mammogram breast images. Let I be an image that consists of a set of x_i grey scale images at pixel i ($i = 1, 2, \dots, N$) and $X = \{x_1, x_2, x_3 \dots x_N\} \subset R^k$ respectively in the k -dimensional area and the cluster centers $v = \{v_1, v_2, v_3 \dots v_c\}$, in which c is said to be a positive integer ($2 < c \ll N$) and u_{ij} is the membership value for each i pixel in j -th cluster ($j = 1, 2, \dots, c$). The clusters framed in the image space are

joined by assigning a different participation membership value to all pixels in the FCM algorithm. At that point, the objective function or general equation of FCM is written in (4).

$$J_{FCM} = \sum_{i=1}^N \sum_{j=1}^c u_{ij}^m x_i - v_j^2 \quad (4)$$

In which, m is an exponent of regularization to the degree of fuzziness, $m > 1$ and $x_i - v_j^2$ be the gray scale Euclidean distance between i and v_j , which is stated in (5).

$$\sum_{j=1}^c u_{ij} = 1, u_{ij} \in [0, 1], 0 \leq \sum_{i=1}^N u_{ij} \leq N \quad (5)$$

Using membership function from the alternate optimization the cluster centers are updated iteratively using (6) and (7).

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(x_i - v_j^2 / x_i - v_k^2 \right)^{1/(m-1)}} \quad (6)$$

$$v_j = \frac{\sum_{i=1}^N u_{ij}^m x_i}{\sum_{i=1}^N u_{ij}^m} \quad (7)$$

The presence of noise is decreased by adding the spatial information of neighboring pixels that is denoted in (8).

$$J_{FCM-S} = \sum_{i=1}^N \sum_{j=1}^c u_{ij}^m x_i - v_j^2 + \frac{\alpha}{N_R} \sum_{i=1}^N \sum_{j=1}^c u_{ij}^m \left(\sum_{r \in N_i} x_r - v_j^2 \right) \quad (8)$$

where, α is denote as spatial information, N_i and N_r is defined as the set of pixel and cardinality of the pixels employed in this system. To avoid the neighborhood function, the term $\frac{1}{N_R} \sum_{r \in N_i} x_r - v_j^2$ is replaced with $\bar{x}_i - v_j^2$. Where, $\bar{x}$ is a gray scale filtered image and a kernel function is used to replace Euclidean distance. The updated equation is represented in the (9)

$$J_{FCM-S(1,2)} = \sum_{i=1}^N \sum_{j=1}^c u_{ij}^m x_i - v_j^2 + \alpha \sum_{i=1}^N \sum_{j=1}^c u_{ij}^m \left(\bar{x}_i - v_j^2 \right) \quad (9)$$

In addition, a Gaussian kernel-based FCM is proposed, which calculates the parameter η_j at every step of the iterations to replace α for every cluster. The kernel functions are used to calculate the parameter value, which is represented in (10).

$$\eta_j = \frac{\min_{j' \neq j} \left(1 - K(v_j', v_j) \right)}{\max_K \left(1 - K(v_K, \bar{x}) \right)} \quad (10)$$

Here, K is the kernel function. The general identification of K requires a large number of patterns and also many cluster centers are required to find the optimal value for η_j . To overcome this problem, the combination of spatial context and grayscale information are made using a fuzzy factor. The fuzzy factor G_{ij} is included in the objective function of the FCM, is stated in (11).

$$J_{FLICM} = \sum_{i=1}^N \sum_{j=1}^c [u_{ij}^m x_i - v_j^2 + G_{ij}] \quad (11)$$

Then, the altered fuzzy factor G'_{ij} is derived using (12).

$$G'_{ij} = \sum_{K \in N_i, i \neq K} w_{ik} \left(1 - u_{ij} \right)^m \left(1 - K(x_i, v_j) \right) \quad (12)$$

This altered fuzzy controls the local neighbor relationship and replace the distance with a kernel function, where w_{ik} denotes the fuzzy factor i and $1 - K(x_i, v_j)$ denotes kernel metric function. In the proposed methodology (M-ARKFCM), the kernel function k is replaced by correlation function c , so the equation (10) is updated as (13).

$$\eta_j = \frac{\min_{j' \neq j} \left(1 - c(v_j', v_j) \right)}{\max_K \left(1 - c(v_K, \bar{x}) \right)} \quad (13)$$

After segmentation phase, a dual-ROI is applied on the segmented output, which is briefly described below.

3.3. Dual-Region of Interest

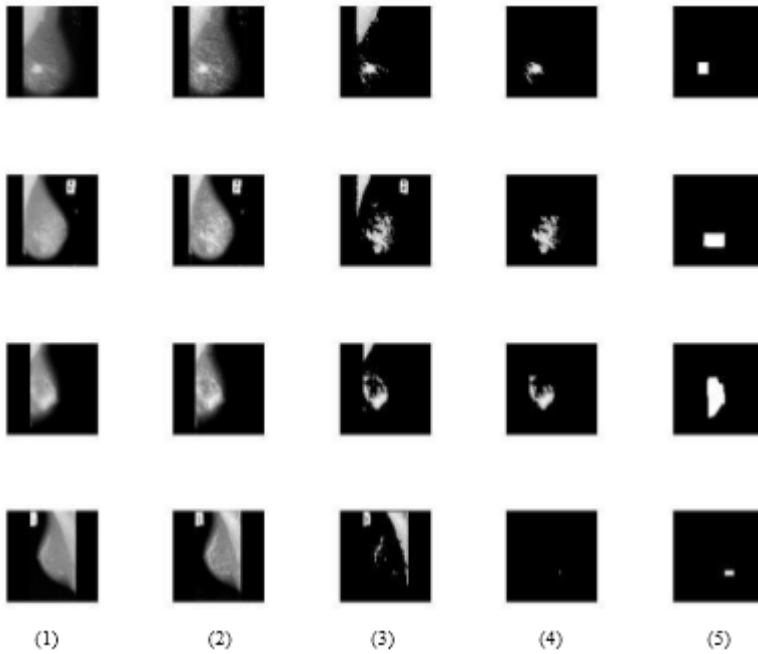
The primitive performance is common for both ROI and dual-ROI, however dual-ROI is an repetitious interpretation. Along these lines, it is adaptable and faster for the selection of ROI for extensive image datasets. Textural properties of segmented object assume an indispensable role for disclosure of breast cancer or collocation of abnormalities of mammogram images.

The dual-ROI groundwork is advanced for choosing the limit of ROI and furthermore to demonstrate the interior intensity dissemination of huge gathering of mammogram images. In this situation, the image intensity probability outline figured for a given set of mammogram images. At that point, the mean probability is computed for an arrangement of mammogram images and it is applied on image magnitude probability map PI for evaluating binary image BPI , whose pixels are elevated than the threshold value I . Binary mask of dual-ROI is represented as BIr and then work out the shape image of BIr . The dual-ROI intensity information term is spell out in (14).

$$D - ROI(i) = \frac{1}{n} \iint (\varphi_M(x) - \varphi_r(N)).dx \quad (14)$$

where, n designate to the entire amount of mammogram images and x accredits the overall outline or replica of mammogram images and shape of non-selected region of mammogram image is represented as i . The general flow illustration of prospective methodology is represented in the Fig 2.

Figure 2. 1) Original image, 2) Pre-processed image, 3) segmented output image, 4) Dual-ROI image, 5) Ground truth image



3.4. Statistical Feature Extraction

In the wake of applying dual-ROI in fragmented output then the FE is performed on the portioned mammogram images. FE is characterized as the plan of aligning the image from image space to the feature space. Here, the FE is enforced on the texture modes such as, contrast, correlation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, variance, sum average, sum variance, sum entropy, area and difference variance for extracting the texture features of mammogram images. These features ascertain the feature information about the mammogram images.

4. RESULTS AND DISCUSSIONS

In the scheme of investigational simulation, MATLAB (version 2017a) was engaged on PC with 3.2 GHz with i5 processor. Keeping in mind the end goal to surmise the effectiveness of anticipated algorithm, the execution of planned strategy was contrasted with FCM and ARKFCM on the presumed database INbreast. A set of mammograms were selected from the publicly available INbreast database. Every image with 1536x1024 pixels, sampled at 200 micrometer pixel size resolution. The execution of the proposed methodology was looked at as far as of Area under the Curve (AUC), dice-coefficient, and Jaccard coefficient.

4.1. Execution/Conduct Measures

The correlation among the input and output variables of a framework comprehend by exploiting the appropriate performance metrics like AUC, dice-coefficient, and Jaccard coefficient. In segmentation validation, the dice coefficient is communicated as in terms of TP, TN, FP and FN counts, which is obtained by matching the segmented result to the ground truth image. These values are used to ascertain the dice coefficient as shown in (15).

$$Dice\ coefficient = \frac{2TP}{(2TP + FP + FN)} \quad (15)$$

where, the dice coefficient value “0” shows, no similarity between the outcomes and the value “1” demonstrates the similarity between the output and ground truth image.

In Jaccard coefficient, the TP esteems are identified by the overlaps between the manually segmented ground truth cancer labels and the machine generated cancer labels. The general formula to calculate Jaccard co-efficient is denoted in (16).

$$Jaccard\ coefficient = \frac{TP}{FP + FN + TP} \quad (16)$$

where, TP is portrayed as true positive, FP is imported as false negative, TN is expressed as true negative and FN is stated as false negative.

Area under the curve (AUC) is a simple measurement metric used to measure the accuracy by reducing ROC curve result into a scalar value 14 . The value of this method is normalized between the range of 0 and 1. Higher value of AUC indicates a better performance of the segmentation. It is calculated using the formulae:

$$AUC = \int_x^y f(a) da \quad (17)$$

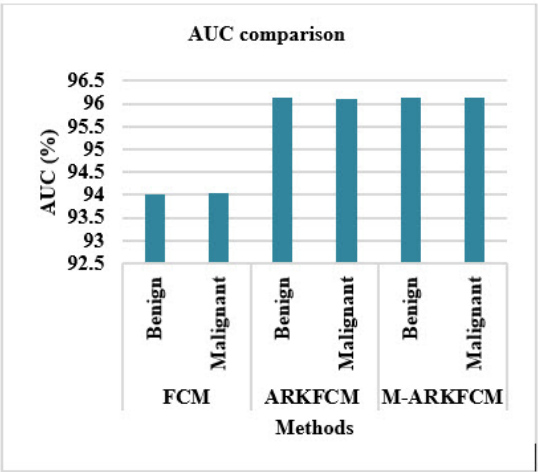
where ‘ x ’ and ‘ y ’ are the minimum and maximum axis points in the curve with ‘ $f(a)$ ’ a function partly above and below the curve. In simple words, AUC is the difference between the area above ROC curve and area below ROC curve.

4.2. Experimental Test Result on INbreast Database

In this experimental investigation, INbreast dataset is assessed for contrasting the performance assessment of accessible strategies and the anticipated conspire. In table 1, the performance assessment of proposed and existing techniques are contrasted with the both benign and malignant (cancer stages).

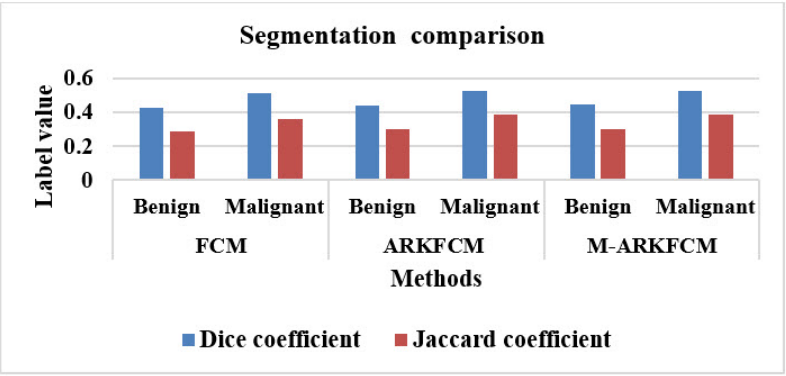
The AUC of proposed procedure for benign is 0.96142 ± 0.0031 and the existing methodologies offer 0.94 ± 0.0083 and 0.9613 ± 0.0067 . Similarly, the AUC of proposed methodology for malignant is 0.9613 ± 0.0029 and the existing methodologies offer 0.944 ± 0.0079 and 0.9611 ± 0.0052 . The graphical portrayal of AUC comparison is denoted in the Fig 3.

Figure 3. AUC comparison of existing and proposed methodology



Similarly, the Dice coefficient and Jaccard coefficient of proposed strategy for benign is 0.443 and 0.2994. The current techniques delivers 0.422 and 0.441 of Dice coefficient and 0.282 and 0.299 of Jaccard coefficient. Similarly, the Dice coefficient and Jaccard coefficient of proposed methodology for malignant is 0.5243 and 0.3874. The existing methodologies delivers 0.509 and 0.5232 of Dice coefficient and 0.357 and 0.3863 of Jaccard coefficient. The graphical representation of Dice coefficient and Jaccard coefficient comparison is insinuated in the Fig 4.

Figure 4. Dice and Jaccard coefficient similarity of accessible and projected line of attack



The Table 1 confirmed that the anticipated tactic performs better than the obtainable methods on the INbreast mammogram images; because modified-ARKFCM encodes the both local and shape features of wavelet transformed mammogram medical images for improving the segmentation efficiency of breast cancer.

Table 1. Performance similarity of accessible and projected line of attack

Clustering approach	Cancer stage	AUC	Dice coefficient	Jaccard coefficient
FCM	Benign	0.94 ± 0.0083	0.422	0.282
	Malignant	0.944 ± 0.0079	0.509	0.357
ARKFCM	Benign	0.9613 ± 0.0067	0.441	0.299
	Malignant	0.9611 ± 0.0052	0.5232	0.3863
M-ARKFCM	Benign	0.96142 ± 0.0031	0.443	0.2994
	Malignant	0.9613 ± 0.0029	0.5243	0.3874

4.3. Relative Examination

The Table.2 presents the relative investigation of on hand works and planned work presentation. C. Varela, et al. 2007 recommended a impressive procedure named as adaptive threshold segmentation, which depends on contour-related, gray level texture and morphological highlights. This existing work accomplishes 0.7877 ± 0.003 of AUC value. Similarly, P.Agrawal, et al. 2014 proposed another new segmentation methodology using saliency based segmentation. After segmentation, a compelling classifier SVM was implemented for classifying the cancer and non-cancer region of segmentation output. This analysis was carried out on a overtly accessible database (i.e. INbreast database) to certify its outcome in terms of AUC and achieved 0.8917 ± 0.001 .Whereas, the proposed work accomplishes 0.96136 ± 0.0030 of AUC that is advanced than the offered works.

Table 2. Performance comparison of accessible and projected work

Clustering approach	AUC
Adaptive threshold segmentation	0.7877 ± 0.003
Saliency based segmentation	0.8917 ± 0.001
M-ARKFCM	0.96136 ± 0.0030

4.4 Pre-predicting the Classification Result Using Feature Esteems

In this section, the mammogram breast cancer classification result is pre-predicted utilizing statistical feature values, Table 3 indicates the feature vector estimations of statistical highlights. Totally, fifteen statistical features are considered for texture FE such as contrast, correlation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, variance, sum

Table 3. Feature vector values of statistical features

Images	Benign 1	Benign 4	Benign 8	Benign 13	Benign 17	Malignant 1	Malignant 6	Malignant 9	Malignant 14	Malignant 20
Area	3841	4349	269	1467	44	1153	3651	2706	28	3744
Sum entropy	12.94	19.06	4.84	8.788	4.121	7.776	16.27	13.08	4.771	16.61
Sum variance	2.683	2.932	2.057	2.314	2.009	2.243	2.782	2.58	2.053	2.799
Sum average	3.657	5.151	1.227	2.379	1.011	2.061	4.48	3.571	1.21	4.556
Difference variance	0.334	0.267	0.033	0.131	0.007	0.102	0.255	0.203	0.031	0.255
Maximum probability	0.988	0.995	0.998	0.994	0.999	0.996	0.99	0.993	0.998	0.991
Variance	0.94	0.93	1	0.97	1	0.98	0.94	0.95	1	0.94
Entropy	0.88	0.87	0.99	0.95	1	0.96	0.88	0.91	0.99	0.88
Homogeneity	0.346	0.27	0.033	0.135	0.007	0.105	0.261	0.208	0.032	0.261
Energy	0.04	0.03	0.01	0.04	0	0.02	0.07	0.05	0.01	0.06
Dissimilarity	77.985	143.79	9.854	51.908	1.467	42.092	120	93.254	9.021	123.2
Cluster Shade	895.442	1882.4	134.385	698.771	19.658	5572.29	1576.35	1240.83	122.582	1618.83
Cluster Prominence	0.954	0.965	0.843	0.861	0.727	0.905	0.91	0.91	0.82	0.921
Correlation	0.172	0.225	0.063	0.282	0.018	0.157	0.465	0.349	0.067	0.415
Contrast	3.61	5.08	1.23	2.27	1.03	2.02	4.29	3.44	1.21	4.39

average, sum variance, sum entropy, area and difference variance. Among these statistical features, area is the most critical segmentation feature for classification result. By setting the cumulative score between the ranges of 1000-3700, we can able to achieve 80% of classification accuracy. Moreover, by considering the feature methods like correlation, cluster shade, dissimilarity, entropy and area, we can able to achieve >70% of classification accuracy by setting the cumulative score < 800.

5. CONCLUSION

In this paper, a new texture based mammogram segmentation methodology is proposed, which depends on modified-ARKFCM with dual-ROI. Modified-ARKFCM is the most effective methodology in mammogram breast cancer segmentation. In this experimental inquire, Modified-ARKFCM is used for segmenting the cancer and non-cancer regions in mammogram images. Subsequently, the statistical texture features are utilized for recognizing the patterns of cancer and non-cancer regions in mammogram images. The proposed methodology adequately combines the advantages of statistical texture features and modified-ARKFCM. The trial investigation is confirmed on overtly obtainable databases (INbreast dataset), which demonstrates a predominance of the recommended framework. The segmentation rate on the mammogram image is more efficient in modified-ARKFCM than the FCM, ARKFCM on hand methodologies. In the forthcoming work, feature vectors obtained from the statistical texture features are employed in an appropriate binary classification methodology for classifying the cancer regions in mammogram images.

FUNDING AGENCY

The publisher has waived the Open Access Processing fee for this article.

REFERENCES

- Agrawal, P., Vatsa, M., & Singh, R. (2014). Saliency based mass detection from screening mammograms. *Signal Processing*, 99, 29–47. doi:10.1016/j.sigpro.2013.12.010
- Bala, B. K., & Audithan, S. (2014, July). Wavelet and curvelet analysis for the classification of macro calcification using mammogram images. *International Conference on Current Trends in Engineering and Technology-ICCTET 2014*, 517–521. doi:10.1109/ICCTET.2014.6966351
- Berber, T., Alpkocak, A., Balci, P., & Dicle, O. (2013). Breast mass contour segmentation algorithm in digital mammograms. *Computer Methods and Programs in Biomedicine*, 110(2), 150–159. doi:10.1016/j.cmpb.2012.11.003 PMID:23273502
- Beura, S., Majhi, B., & Dash, R. (2015). Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer. *Neurocomputing*, 154, 1–14. doi:10.1016/j.neucom.2014.12.032
- Dalmiya, S., Dasgupta, A., & Datta, S. K. (2012). Application of wavelet based k-means algorithm in mammogram segmentation. *International Journal of Computers and Applications*, 52(15).
- Dheeba, J., Singh, N. A., & Selvi, S. T. (2014). Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach. *Journal of Biomedical Informatics*, 49, 45–52. doi:10.1016/j.jbi.2014.01.010 PMID:24509074
- Elter, M., & Haßlmeyer, E. (2008, March). A knowledge-based approach to CADx of mammographic masses. In *Medical Imaging 2008: Computer-Aided Diagnosis* (Vol. 6915, p. 69150L). International Society for Optics and Photonics. doi:10.1117/12.770135
- Eltoukhy, M. M., Faye, I., & Samir, B. B. (2010). Breast cancer diagnosis in digital mammogram using multiscale curvelet transform. *Computerized Medical Imaging and Graphics*, 34(4), 269–276. doi:10.1016/j.compmedimag.2009.11.002 PMID:20004076
- Eltoukhy, M. M., Faye, I., & Samir, B. B. (2012). A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation. *Computers in Biology and Medicine*, 42(1), 123–128. doi:10.1016/j.combiomed.2011.10.016 PMID:22115076
- García-Manso, A., García-Orellana, C. J., González-Velasco, H., Gallardo-Caballero, R., & Macías, M. M. (2013). Consistent performance measurement of a system to detect masses in mammograms based on blind feature extraction. *Biomedical Engineering Online*, 12(1), 2. doi:10.1186/1475-925X-12-2 PMID:23305491
- Giess, C. S., Poole, P. S., Chikarmane, S. A., Sipko, D. A., & Birdwell, R. L. (2015). Screening breast MRI in patients previously treated for breast cancer: Diagnostic yield for cancer and abnormal interpretation rate. *Academic Radiology*, 22(11), 1331–1337. doi:10.1016/j.acra.2015.05.009 PMID:26142951
- INbreast Dataset. (n.d.). http://medicalresearch.inescporto.pt/breastresearch/index.php/Get_INbreast_Database
- Jasmine, J. L., Govardhan, A., & Baskaran, S. (2009, June). Microcalcification detection in digital mammograms based on wavelet analysis and neural networks. In *2009 International Conference on Control, Automation, Communication and Energy Conservation* (pp. 1–6). IEEE.
- Kozegar, E., Soryani, M., Behnam, H., Salamati, M., & Tan, T. (2017). Breast cancer detection in automated 3D breast ultrasound using iso-contours and cascaded RUSBoosts. *Ultrasonics*, 79, 68–80. doi:10.1016/j.ultras.2017.04.008 PMID:28448836
- Li, J. B., Wang, Y. H., & Tang, L. L. (2012). Mammogram-based discriminant fusion analysis for breast cancer diagnosis. *Clinical Imaging*, 36(6), 710–716. doi:10.1016/j.clinimag.2012.01.041 PMID:23153999
- Li, Y., Chen, H., Wei, X., Peng, Y., & Cheng, L. (2016). Mass classification in mammograms based on two-concentric masks and discriminating texton. *Pattern Recognition*, 60, 648–656. doi:10.1016/j.patcog.2016.06.021
- Mazurowski, M. A., Lo, J. Y., Harrawood, B. P., & Tourassi, G. D. (2011). Mutual information-based template matching scheme for detection of breast masses: From mammography to digital breast tomosynthesis. *Journal of Biomedical Informatics*, 44(5), 815–823. doi:10.1016/j.jbi.2011.04.008 PMID:21554985
- Mubarak, D. M. N., Sathik, M. M., Beevi, S. Z., & Revathy, K. (2012). A hybrid region growing algorithm for medical image segmentation. *International Journal of Computer Science and Information Technologies*, 4(3), 61–70. doi:10.5121/ijcsit.2012.4306

Muramatsu, C., Hara, T., Endo, T., & Fujita, H. (2016). Breast mass classification on mammograms using radial local ternary patterns. *Computers in Biology and Medicine*, 72, 43–53. doi:10.1016/j.combiomed.2016.03.007 PMID:27015322

Rampun, A., Morrow, P. J., Scotney, B. W., & Winder, J. (2017). Fully automated breast boundary and pectoral muscle segmentation in mammograms. *Artificial Intelligence in Medicine*, 79, 28–41. doi:10.1016/j.artmed.2017.06.001 PMID:28606722

Sampaio, W. B., Diniz, E. M., Silva, A. C., De Paiva, A. C., & Gattass, M. (2011). Detection of masses in mammogram images using CNN, geostatistic functions and SVM. *Computers in Biology and Medicine*, 41(8), 653–664. doi:10.1016/j.combiomed.2011.05.017 PMID:21703605

Tai, S. C., Chen, Z. S., & Tsai, W. T. (2013). An automatic mass detection system in mammograms based on complex texture features. *IEEE Journal of Biomedical and Health Informatics*, 18(2), 618–627. PMID:24608061

Varela, C., Tahoces, P. G., Méndez, A. J., Souto, M., & Vidal, J. J. (2007). Computerized detection of breast masses in digitized mammograms. *Computers in Biology and Medicine*, 37(2), 214–226. doi:10.1016/j.combiomed.2005.12.006 PMID:16620805

Wang, D., Shi, L., & Heng, P. A. (2009). Automatic detection of breast cancers in mammograms using structured support vector machines. *Neurocomputing*, 72(13-15), 3296–3302. doi:10.1016/j.neucom.2009.02.015

Zhong, X., Li, J., Ertl, S. M., Hassemer, C., & Fiedler, L. (2015). A system-theoretic approach to modeling and analysis of mammography testing process. *IEEE Transactions on Systems, Man, and Cybernetics. Systems*, 46(1), 126–138. doi:10.1109/TSMC.2015.2429643

T. Venkata Satya Vivek received his Bachelor's Degree in Computer Science & Engineering from Jawaharlal Nehru Technological University affiliated college way back in 2012. Thereafter he did his Master of Technology (M.Tech) from KL University, Vijayawada under the specialization of Computer Networks and Security. He completed his PhD thesis from Dr. M.G.R Educational & Research Institute (Deemed to be University), Chennai. Presently he is working as Assistant Professor in the Department of Computer Science & Engineering from Koneru Lakshmaiah Education Foundation (Deemed to be University), Vaddeswaram, Guntur, Andhra Pradesh, India. To his credit he is having more than 15 research papers from Scopus indexing, UGC care journals and One (1) Indian Patent. His research areas include Applications of Cryptography, Data Hiding in Images and Machine Learning.

Raju C. received his B.Tech and M.Tech degrees in Electronics and Communication Engineering from Jawaharlal Nehru Technological University Anantapur (JNTUA), Andhra Pradesh, India in 2009 and 2012 respectively. He received Doctoral degree from Sri Venkateswara University, Tirupati. Currently, he is working as Associate Professor in Department of Electronics and Communication Engineering, SV Engineering College, Tirupati. His research focuses on areas of Radar signal processing, microwave remote sensing and image processing.

D. Girishkumar is working as Assistant Professor in the Department of ECE at Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh. He has 6.5 years teaching as well as research experience. He received his M.Tech degree from V. R. Siddhartha engineering college in Communication and signal Processing, B.Tech degree in ECE from St'Anns College of Engineering and Technology, Chirala in 2009. His focused areas are optimization algorithms and Digital image processing.

Yalla Venkateswarlu obtained Triple PG and PhD in Computer Science Engineering. He is a currently working as a Professor in Department of Computer Science & Engineering of BVC College of Engineering, Palacharla, Rajahmundry. He is a versatile senior educationist, administrator, confident leader, and self-motivated expertise with excellence in subject knowledge more than 25 years of solid professional teaching, research experience in reputed engineering colleges. He is a Dynamic instructor, motivational speaker and mentor focused on providing students with rigorous and challenging education to make students to National builders. Dr Yalla shaping the student's career towards innovative ideas to encourage the entrepreneur for solving real world problems. His Ideology imports the students to learn advanced cutting edge emerging technologies in engineering field. His Research Areas include Image Processing, Machine Learning, and Big Data Analytics.