Computer Systems Science & Engineering
DOI:10.32604/csse.2022.018837
Article

Tech Science Press

Hybrid Active Contour Mammographic Mass Segmentation and Classification

K. Yuvaraj* **and U. S. Ragupathy**

Department of Electronics and Instrumentation Engineering, Kongu Engineering College Perundurai, Erode, 638060, India
*Corresponding Author: K. Yuvaraj. Email: kyuvarakeie2020@gmail.com, yuvarajkumarasamy@gmail.com
Received: 23 March 2021; Accepted: 28 May 2021

Abstract: This research implements a novel segmentation of mammographic mass. Three methods are proposed, namely, segmentation of mass based on iterative active contour, automatic region growing, and fully automatic mask selection-based active contour techniques. In the first method, iterative threshold is performed for manual cropped preprocessed image, and active contour is applied thereafter. To overcome manual cropping in the second method, an automatic seed selection followed by region growing is performed. Given that the result is only a few images owing to over segmentation, the third method uses a fully automatic active contour. Results of the segmentation techniques are compared with the manual markup by experts, specifically by taking the difference in their mean values. Accordingly, the difference in the mean value of the third method is 1.0853, which indicates the closeness of the segmentation. Moreover, the proposed method is compared with the existing fuzzy C means and level set methods. The automatic mass segmentation based on active contour technique results in segmentation with high accuracy. By using adaptive neuro fuzzy inference system, classification is done and results in a sensitivity of 94.73%, accuracy of 93.93%, and Mathew's correlation coefficient (MCC) of 0.876.

Keywords: Feature optimization; hybrid active contour; segmentation; mass classification; mass feature extraction; medical image analysis

1 Introduction

The X-ray of breast is called a mammogram, which is used to locate cancer at an early stage. Mass is nothing but a lump that has its own shape and marginal properties. Benign is an initial stage of cancer and marginated smoothly, whereas malignancy becomes more violent with time. Mass can be categorized into three types, namely, well-defined, speculated, and ill-defined [1,2]. Numerous studies have been conducted to identify a mass in its early stages. The literature has indicated that segmentation, which is highly efficient, accurate, and fully automatic, is needed for spotting a mass at its beginning phase [3,4].

This paper is organized as follows. Section 1 describes a mammographic mass and its three types. Section 2 discusses mammographic mass segmentation based on iterative active contour. Section 3 illustrates the automatic seed selection technique. Section 4 discusses an automatic mask selection-based active contour. Section 5 deals with classification. Lastly, Section 6 draws the conclusion and future scope.

This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

2 Iterative Active Contour-Based Mass Segmentation

Mammographic image with mass is considered an input and preprocessing techniques are used, which include cropping and equalization using histogram. Preprocessed images are subjected to iterative thresholding to extract a rough boundary. Active contour is applied to the iterative threshold image, leading to final segmentation. The main advantage of contour-based segmentation is that it preserves mass region information, edge, and shape. All processes are carried out and the results are obtained using MATLAB. Fig. 1 shows the architecture diagram of the iterative active contour segmentation.

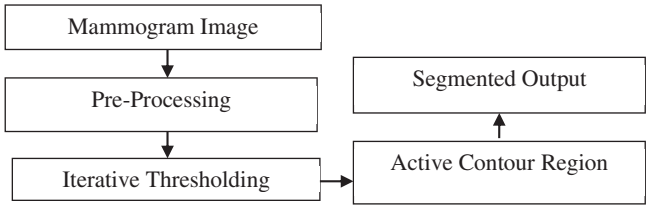


Figure 1: Architecture diagram of the iterative active contour technique

Images are obtained from the MIAS database and collected from the Clarity Imaging Center, Coimbatore. The mass region is cropped and the preprocessed gray scale picture is converted into a binary image using the iterative threshold technique [5,6]. The threshold value θ is calculated iteratively based on the foreground and background strengths. The threshold value of the image shown in Fig. 2 is calculated as 128. The original, preprocessed, and threshold images are shown in Fig. 2.

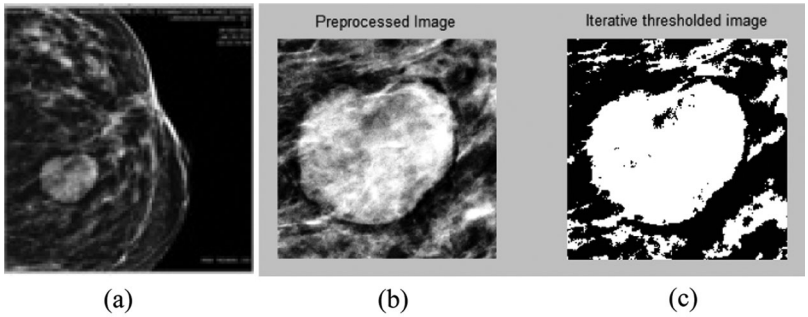


Figure 2: (a) Actual, (b) preprocessed, and (c) iterative threshold images

Threshold image is taken as input. The mask image size is taken as the same size of the iterative threshold image. Within this image, a square mask of size 111 × 111 is defined. On the basis of energy minimization, the defined mask will fit to the boundary. Gradient descent minimization is used to minimize snake energy [7]. The force acting on the contour is as follows:

$$G_1(i) + G_2(i) = \int_{\text{inside(i)}} |u0 - i1|^2 \, dx + \int_{\text{outside(i)}} |u0 - i2|^2 \, dx. \tag{1}$$

The total energy of the contour is as follows:

$$En^*_{snake} = \int_0^1 \left(En_{internal}(S(v)) + En_{image}(S(v)) + En_{con}(S(v))\right) ds, \tag{2}$$

where $S(v)$ gives the derivative of S with respect to v. The internal energy of the snake is given as follows:

$$\nabla En_{int}(s) = \left[\left(\alpha(v)\ \nabla\left|\left|\frac{dv}{ds}(s)\right|\right|^2 + \beta(s)\nabla\left|\left|\frac{d^2v}{ds^2}(v)\right|\right|^2\right)/2\right]. \tag{3}$$

The external energy acting on the contour is as follows:

$$En_{external} = En_{image} + En_{con}, \tag{4}$$

where En_{image} gives the force acting on the contour and En_{con} depicts the external constraint force as follows:

$$En_{image} = w_{line}En_{line} + w_{edge}En_{edge} + w_{term}En_{term}. \tag{5}$$

$En_{line} = I(x, y)$, which depends on the sign of w_{line}, and the width of the line varies. All edge information will be found by the edge energy function as follows:

$$E_{edge} = -|\nabla I(x, y)|^2, \tag{6}$$

$$En_{term} = \frac{i_{yy}i_x^2 - 2i_{xy}i_xi_y + i_{xx}i_y^2}{\left(i_x^2 + i_y^2\right)^{\frac{3}{2}}}. \tag{7}$$

Eq. (8) represents the next iteration of the contour:

$$x_{t+1} = x_t + \gamma\nabla f(x_t). \tag{8}$$

The iteration repeats until the snake converges its minimal boundary [8]. Fig. 3 illustrates various stages involved in the proposed method of mass segmentation [9,10]. The output obtained (450^{th} iteration) for a sample image is given in the following section.

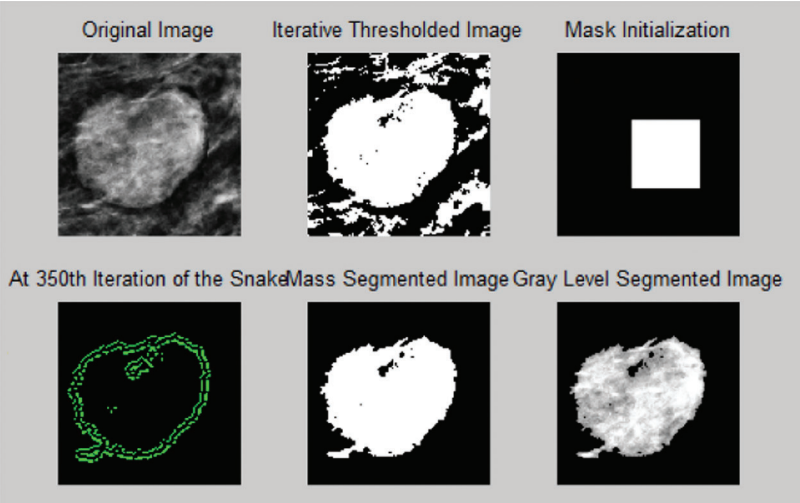


Figure 3: Various stages of output in the proposed iterative active contour-based segmentation

The mass affected area in the mammographic image is cropped manually and allowed for further processing. This drawback is avoided by adopting a new segmentation scheme, in which the entire image is considered without cropping.

3 Automatic Seed Selection Based on Region Growing Technique

Fig. 4 shows the flow diagram of the automatic selection of seed based on the region growing [11,12] mass segmentation technique [13,14]. As a pre-study, a mask of size 20 × 20 is initialized for various reference images. For regions within the mask, statistical features are extracted, namely, contrast, dissimilarity, sum average, sum variance, and auto correlation, as listed in Tab. 1. For the image under testing, if the extracted features lie in the range of the predefined statistical features of mass, then the seed is fixed. The input image is subjected to thresholding. Thereafter, the final gray level segmented picture is obtained. The mask is moved in case of a mismatch in the extracted features.

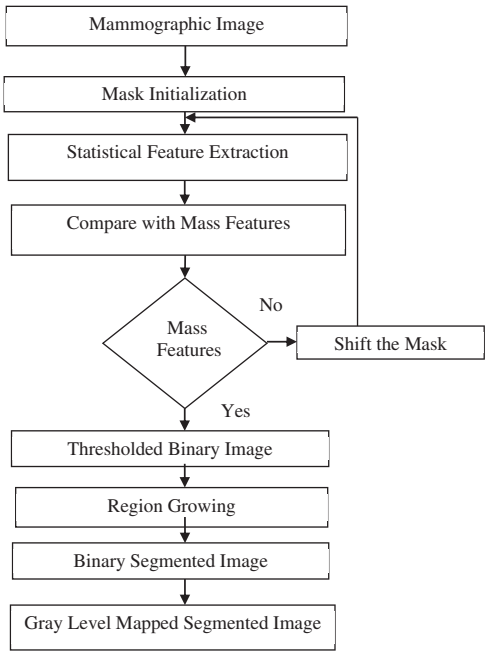


Figure 4: Flow diagram of the automated seed selection-based region growing mass segmentation technique

Table 1: Statistical features extracted for image SE003

Features	Values
Mean	192
Dissimilarity	0.2762
Sum average	7.7810
Sum variance	40.2691
Auto correlation	24.2833

Fig. 5 shows the various stages of mass segmentation using automatic seed selection-based region growing technique.

Fig. 6 shows over segmentation. A single pixel can act as a bridge for leakage. The active contour technique is utilized to solve this problem.

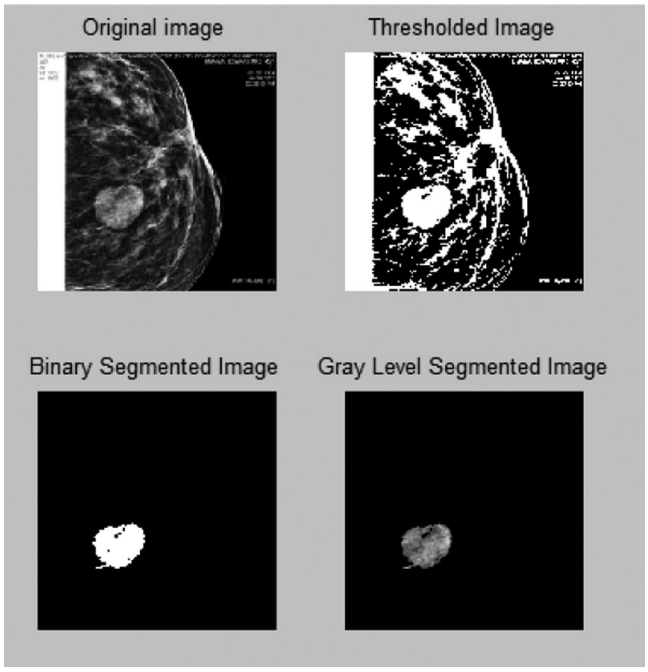


Figure 5: Various stage outputs of the proposed automatic region growing-based segmentation

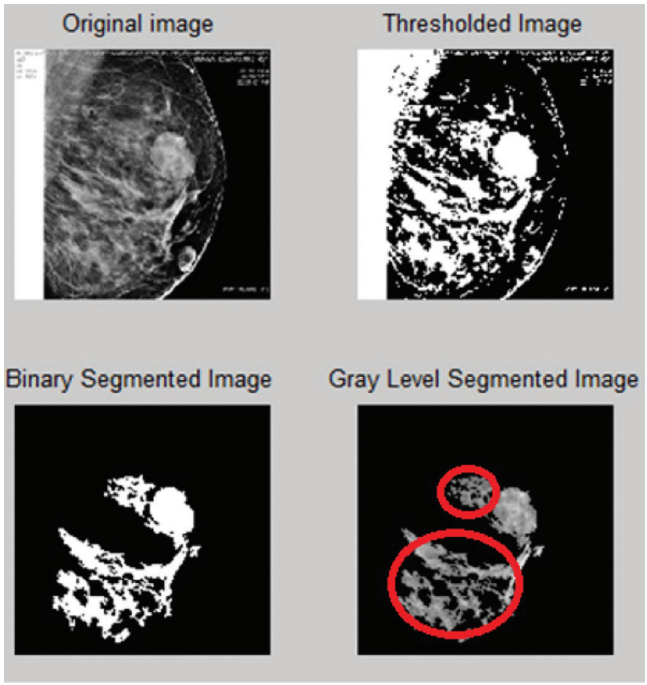


Figure 6: Over segmentation for the proposed automatic region growing-based segmentation

4 Automatic Mass Selection-Based Active Contour Technique

This method completely aims at automatic mask selection based on the active contour technique. The mask for the active contour technique is chosen automatically by reading the mass features. Fig. 7 illustrates the flow diagram of the fully automatic active contour-based mass segmentation.

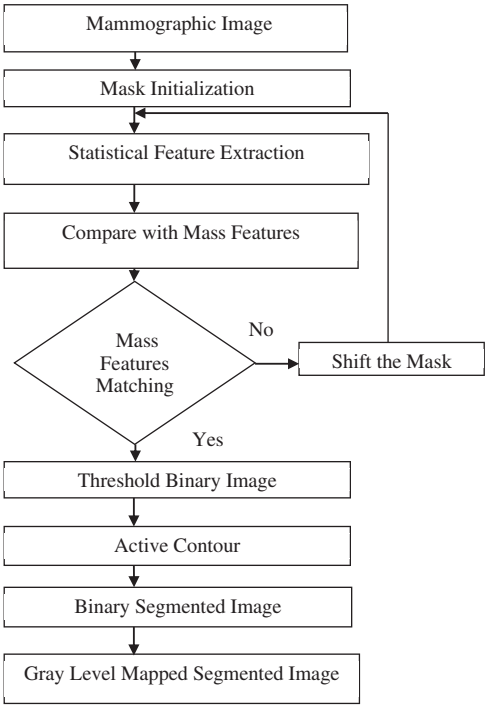


Figure 7: Flow diagram of the proposed fully automatic mass segmentation based on active contour

A mask of size 20 × 20 is initialized. Five statistical features are extracted for the regions within the mask (same as in method 2). If the extracted features lie in the range of the predefined features of mass, then the mask is fixed. The input image is subjected to thresholding and active contour technique is implemented thereafter [15]. Lastly, gray level segmented image is obtained. Thereafter, the mask is moved in case of mismatch in the extracted features [16]. Fig. 8 depicts the output of the preceding method.

Fig. 9 shows the output of the three proposed methods for various mammographic images. Evidently, over segmentation for image 2 is prevented by using the automatic active contour technique.

The expected output is marked manually by the physician, and the mean values of the three proposed methods are calculated in Tab. 2. Note that the difference in the mean values of the proposed method 3 is 1.085. This value is closer to zero, thereby indicating the closeness of the segmentation.

The proposed method is compared with the existing methods, such as the fuzzy C means (FCM) clustering and level set methods (see Fig. 10).

The statistical means of the preceding methods are compared using Fiji (see Tab. 3).

Tab. 3 shows that the proposed algorithm holds good because its difference of mean value is low, which shows the closeness of segmentation.

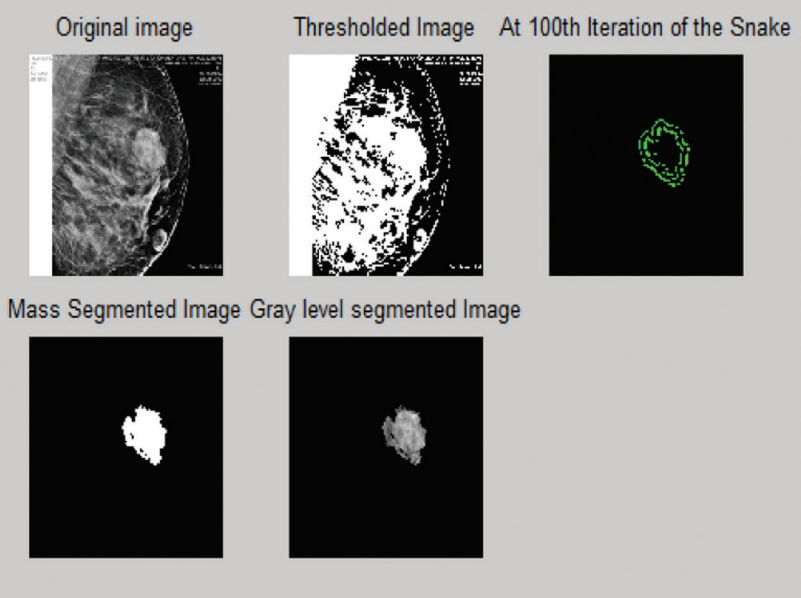


Figure 8: Various stages of outputs in the proposed automatic active contour-based segmentation

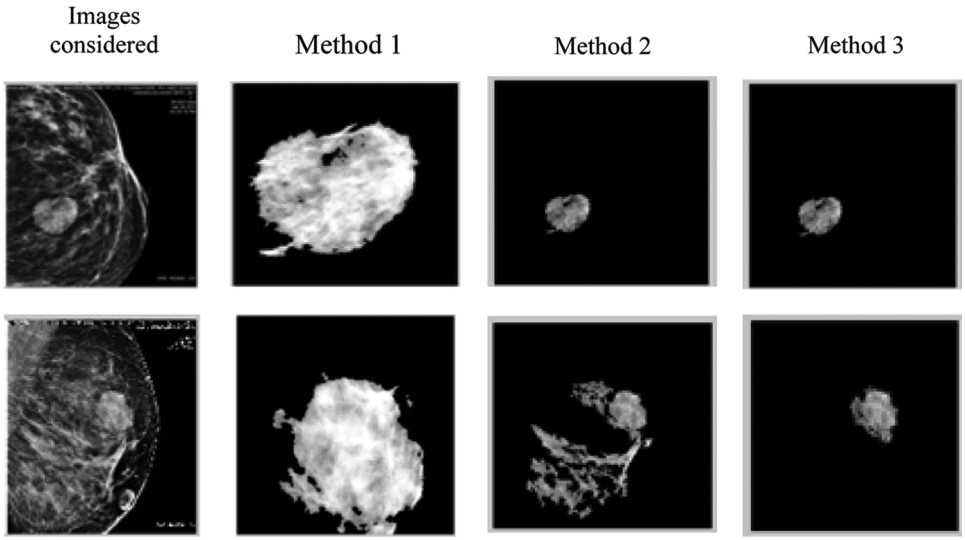


Figure 9: Output of the three proposed methods for various mammographic images

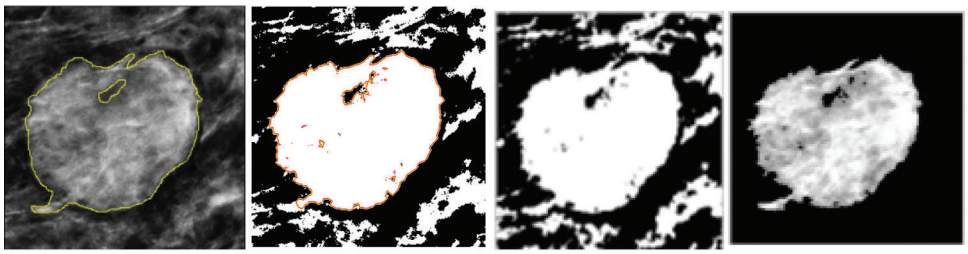
5 Classification Using ANFIS

From the segmented result of the automatic active contour technique, a 20 × 20 mask is defined and the following statistical features are extracted from its formulated gray level co-occurrence matrix (GLCM). Initially, 16 statistical features are extracted, as shown in Tab. 4. Based on variance, five features are selected and listed in Tab. 5.

ANFIS is used as a classifier [17,18]. Various membership functions and their testing errors are shown in Fig. 11. The triangular membership function with minimum error of 0.463 is considered. Three triangular membership functions are used for each input. A total of 729 rules are framed. Tab. 6 lists the number of images considered for classification.

Table 2: Output comparison of the three proposed methods

Mean Values Image Samples	Mean values of the Manual Segmented Image (M)	Mean Values of the Iterative Active Contour-based Segmented Image (x)	Difference in the Mean values (X-M)	Mean values of the Automatic Region-based Active Contour Segmented Image (Y)	Differences in the Mean values (Y-M)	Mean Values of the Automatic Active Contour-based Segmented Image (z)	Difference in the Mean values (Z-M)
se003	189.399	191.317	1.918	190.251	0.852	189.755	0.356
se001	167.825	169.5	1.675	171.221	3.396	169.14	1.315
mdb184	72.255	73.298	1.043	73.654	1.399	72.998	0.733
mdb025	83.411	84.728	1.317	83.655	0.244	83.342	0.069
mdb015	59.067	62.015	2.948	61.431	2.364	61.587	2.52
sc1	47.519	49.261	1.742	50.362	2.843	49.110	1.519
Average Values of the Difference in Mean Values			1.773		1.849		**1.0853**



(a) (b) (c) (d)

Figure 10: Comparison of the segmented outputs: Proposed method with the existing algorithms: (a) manually marked, (b) level set, (c) FCM, and (d) proposed methods

Table 3: Performance comparison of the proposed automated segmentation against the existing methods

Techniques	Processes	Area	Mean
Level set approach	Manual segmentation	179104	189.396
	Automated segmentation	179104	191.47
	Difference in the mean values		2.074
FCM clustering	Manual segmentation	179104	189.395
	Automated segmentation	179104	192.078
	Difference in the mean values		2.683
Proposed method	Manual segmentation	179104	189.399
	Automated segmentation	179104	191.317
	Difference in the mean values		**1.918**

Table 4: Features extracted from GLCM

Image Categories Features Extracted from GLCM	Normal Mdb123 (p)	Benign Mdb69 (q)	Malignant Mdb267 (r)	Mean S = (p + q + r)/3	Variance ((S − p) + (S − q) + (S − r))/3
Contrast	0.3806	0.4527	0.0736	0.3023	0.17
Correlation	−1.3024	−1.2568	−1.261	−1.2734	0.0193
Dissimilarity	0.0567	0.0704	1.0457	0.3909	0.4376
Energy	0.5348	0.4865	0.4843	0.5018	0.01230
Entropy	−0.423	−0.3952	−0.3932	−0.4038	0.0120
Homogeneity	0.9925	0.9893	0.9865	0.9894	0.0052
Maximum Probability	0.6433	0.4975	0.5032	0.5486	0.0638
Sum average	9.0695	6.9192	7.1434	7.7107	1.0849
Sum variance	120.664	90.7035	93.4085	101.5919	12.7139
Sum entropy	0.3297	0.4142	0.4216	0.3885	0.040
Difference variance	0.3808	0.4524	0.6038	0.479	0.0783
Difference entropy	0.0616	0.0867	0.0993	0.082533	0.0132
Information measure of correlation	0.9846	0.9864	0.986	0.985667	0.0015
Inverse difference normalized	0.996	0.9952	0.9934	0.994867	0.0005
Inverse difference moment normalized	0.9967	0.9954	0.9947	0.9956	0.0003
Auto correlation	41.6198	31.8956	32.8252	35.44687	4.1143

Table 5: Features selected for classification

Image Categories Features	Normal mdb047	Benign mdb025	Malignant mdb264
Contrast	0.3660	0.3122	0.7849
Dissimilarity	0.0549	0.0449	0.1205
Sum average	7.4158	7.0525	7.1308
Sum variance	97.8039	93.1585	92.7945
Auto correlation	34.1820	32.5815	32.6760

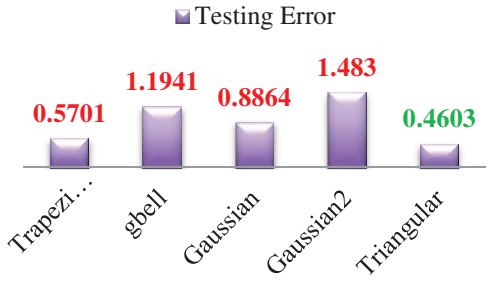


Figure 11: Various membership functions and their testing errors

Table 6: Image data sets

Image categories	No. of training images	No. of testing images	Total images
Normal	18	14	32
Benign	20	11	31
Malignant	11	8	19
Total images considered			82

The classified ANFIS plot is shown in Fig. 12. As presented in Fig. 12, a confusion matrix is formed and shown in Tab. 7.

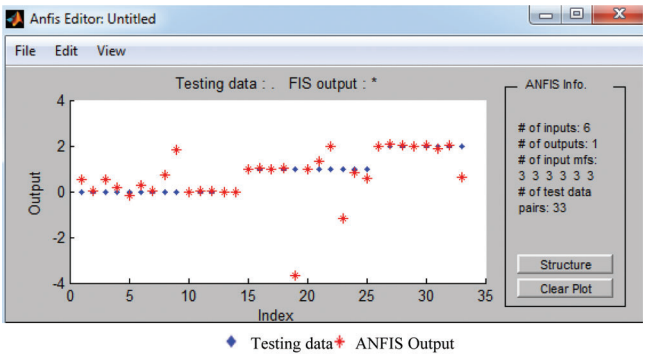


Figure 12: ANFIS classifier output

Table 7: Confusion matrix

Actual Desired	Normal	Benign	Malignant
Normal	13	1	0
Benign	1	8	2
Malignant	0	1	7

From the confusion matrix, the following performance metrics are calculated: sensitivity, specificity, precision, accuracy, and MCC [15]:

$$\text{Sensitivity} = \frac{TP}{TP + FN},$$

$$\text{Specificity} = \frac{TN}{TN + FP},$$

$$\text{Precision} = \frac{TP}{TP + FP},$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN},$$

$$\text{MCC} = \frac{[(\text{TP} * \text{TN}) - (\text{FP} * \text{FN})]}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}.$$

The preceding values are calculated and listed in Tab. 8.

Tab. 8 shows that accuracy is inferred to be 93.93% and MCC is 0.876. Hence, a high accurate classification is made. The proposed method serves as a second opinion for doctors in diagnosing cancer.

Table 8: Performance metrics of the classifier

Classification Parameters	Results
True Positive (TP)	18
False Positive (FP)	1
True Negative (TN)	13
False Negative (FN)	1
Sensitivity (%)	94.73
Specificity (%)	92.85
Precision (%)	94.73
Accuracy (%)	93.93
Mathew's Correlation Coefficient (MCC)	0.876

6 Conclusion

Women are mostly affected by breast cancer, thereby resulting in this disease becoming a major cause of death. Thus, technology for early detection should be identified. Evidently, the proposed method is of high medical significance in detecting early-stage cancer.

This research proposes a new hybrid active contour mass segmentation. The result of the segmentation and existing techniques are compared with the manual segmentation, proving that the proposed iterative active contour technique stays good in mass segmentation. The classification result proves satisfying in terms of accuracy, sensitivity, and MCC.

In the future, numerous images should be considered and examined using the proposed method, and different modalities can also be included. Along with statistical features, other features, such as structures and patterns, can also be considered. Classification may be performed using advance classifiers.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interests: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] A. Mencattini, G. Rabottino, M. Salmeri, R. Lojacono and E. Colini, "Breast mass segmentation in mammographic images by an effective region growing algorithm," in *Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Science*, Berlin, Heidelberg: Springer, vol. 5259, pp. 948–957, 2008.

[2] A. Oliver, J. Freixenet, J. Martí, E. Pérez, J. Pont *et al.,* "A review of automatic mass detection and segmentation in mammographic images," *International Journal on Medical Image Analysis*, vol. 14, no. 2, pp. 87–110, 2010.

[3] B. Charan Patel and G. R. Sinha, "Energy and region based detection and segmentation of breast cancer mammo graphic images," *International. Journal of Image Graphics and Signal Processing*, vol. 6, pp. 44–51, 2012.

[4] F. Akram, J. Heon Kim, C. Gun Lee and K. Nam Choi, "Segmentation of regions of interest using active contours with SPF Function," *Computational and Mathematical Methods in Medicine*, vol. 2015, pp. 710326:1–710326:14, 2015.

[5] Z. Han, H. Chen, J. Li and C. Yao, "A novel method of mass segmentation in mammogram," in *Int. Conf. on Systems and Informatics (ICSAI2012)*, Yantai, China, pp. 1412–1416, 2012.

[6] M. Mustafa, H. N. Omar Rashid, N. R. Hasma Abdullah, R. Samad and D. Pebrianti, "Mammography image segmentation: Chan-vese active contour and localized active contour approach," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 5, no. 3, pp. 577–588, 2017.

[7] N. Petrick, H. Ping, C. Heang, P. ChanBerkman, S. B. Sahiner *et al.,* "Combined adaptive enhancement and region-growing segmentation of breast masses on digitized mammograms," *Medical Physics*, vol. 26, no. 8, pp. 1642–1654, 1999.

[8] B. Sahiner, N. Petrick, H. Ping Chan, L. M. Hadjiiski and C. Paramagul, "Computer-aided characterization of mammographic masses: Accuracy of mass segmentation and its effects on characterization," *IEEE Transaction on Medical Imaging*, vol. 20, no. 12, pp. 1275–1284, 2001.

[9] S. Kumar and B. Padhyay, "Survey on segmentation methods for locating masses in a mammogram image," *An International Journal of Computer Applications*, vol. 9, no. 11, pp. 25–28, 2011.

[10] Y. Jin Lee, "Mammographic mass detection by adaptive thresholding and region growing," *International Journal of Imaging Systems and Technology*, vol. 11, no. 5, pp. 340–346, 2000.

[11] K. Yuvaraj and U. S. Ragupathy, "Automatic mammographic mass segmentation based on region growing technique," in *Proc. ICEBEA*, Singapore, pp. 121–128, 2012.

[12] K. Yuvaraj and U. S. Ragupathy, "Computer aided segmentation and classification of mass in mammographic images using ANFIS," *European Journal of Biomedical Informatics*, vol. 9, no. 2, pp. 37–41, 2013.

[13] A. Oliver, J. Freixenenet, J. Marti, E. Perez, J. Pont *et al.,* "A review of automatic mass detection and segmentation in mammographic images," *Medical Image Analysis*, vol. 14, pp. 87–110, 2010.

[14] A. Mencattini, G. Rabottino, M. Salmeri, R. Lojacono and E. Colini, "Breast mass segmentation in mammographic images by an effective region growing algorithm," in *Int. Conf. on Advanced Concepts for Intelligent Vision Systems*, pp. 948–957, 2008.

[15] S. Maheswaran, S. Nehru and S. Sathesh, "A wearable embedded device for chronic low back patients to track lumbar spine position," *Biomedical Research*, vol. 20, pp. 118–S123, 2018.

[16] S. Maheswaran, S. Sathesh, M. Gayathri, E. D. Bhaarathei and D. Kavin, "Design and development of chemical free green embedded weeder for row based crops," *Journal of Green Engineering*, vol. 10, no. 5, pp. 2103–2120, 2020.

[17] D. Sami, W. Barhoumi, J. Kurek and B. Swiderski, "False-positive reduction in computer-aided mass detection using mammographic texture analysis and classification," *Computer Methods and Programs in Biomedicine*, vol. 160, pp. 75–83, 2018.

[18] C. Chowdhary and D. P. Acharjya, "A hybrid scheme for breast cancer detection using intuitionistic fuzzy rough set technique," *International Journal of Healthcare Information Systems and Informatics*, vol. 11, pp. 38–61, 2016.