Hindawi
Journal of Healthcare Engineering
Volume 2023, Article ID 1491955, 18 pages
https://doi.org/10.1155/2023/1491955

Research Article

Ensemble Learning-Based Hybrid Segmentation of Mammographic Images for Breast Cancer Risk Prediction Using Fuzzy C-Means and CNN Model

Sudan Jha,[1] **Sultan Ahmad**,[2,3] **Anoopa Arya**,[4,5] **Bader Alouffi**,[6] **Abdullah Alharbi**,[7] **Meshal Alharbi**,[2] **and Surender Singh**[4]

[1]*Department of Computer Science and Engineering, School of Engineering, Kathmandu University, Banepa, Kathmandu, Nepal*
[2]*Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, P.O. Box. 151, Alkharj 11942, Saudi Arabia*
[3]*University Center for Research and Development (UCRD), Department of Computer Science and Engineering, Chandigarh University, Gharuan, Mohali 140413, Punjab, India*
[4]*Department of Computer Science and Engineering, Chandigarh University, Gharuan, Mohali 140413, Punjab, India*
[5]*Maharishi Markandeshwar (Deemed To Be University), Mullana, Ambala, Haryana 133207, India*
[6]*Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*
[7]*Department of Information Technology, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*

Correspondence should be addressed to Sudan Jha; sudan.jha@ku.edu.np

Received 1 November 2021; Revised 23 July 2022; Accepted 25 November 2022; Published 31 January 2023

Academic Editor: Agostino Forestiero

The research interest in this field is that females are not aware of their health conditions until they develop tumour, especially when breast cancer is concerned. The breast cancer risk factors include genetics, heredity, and sedentary lifestyle. The prime concern for the mortality rate among females is breast cancer, and breast cancer is on the rise, both in rural and urban India. Women aged 45 or above are more vulnerable to this disease. Images are more effective at depicting information as compared to text. With the advancement in technology, several computerized techniques have come up to extract hidden information from the images. The processed images have found their application in several sectors and medical science is one of them. Disease-like breast cancer affects most women universally and it happens due to the existence of breast masses in the breast region for the development of breast cancer in women. Timely breast cancer detection can also increase the rate of effective treatment and the survival of women suffering from breast cancer. This work elaborates the method of performing hybrid segmentation techniques using CLAHE, morphological operations on mammogram images, and classified images using deep learning. Images from the MIAS database have been used to obtain readings for parameters: threshold, accuracy, sensitivity, specificity rate, biopsy rate, or a combination of all the parameters and many others under study.

1. Introduction

Cancer is a disease that causes abnormal changes in the body's tissues and cells, as well as growth that is out of control. One of the types of cancer is breast cancer. The prognosis assessment of breast cancer can help patients with breast cancer improve their chances of survival. The idea behind the segmentation is to segment out the region of interest, which gives more meaning due to which analysis is more effective and precise. In females, breast cancer is quite frequent compared to other cancers and is the most prominent reason for cancer death in the world [1]. The reason behind the cause of the disease is still a mystery, and researchers are still working on the same.

Few factors learned which lead to or increased the probability of developing cancer are radiation, dense breast cells, consumption of alcohol, improper living styles, etc. The way to reduce the mortality rate caused by cancer is through early detection and examination at the initial stage of cancer. Segmentation in image processing is an essential step in image processing. In this phase of image processing, we segment out the selected region for extracting the desired information to infer the conclusion. The data fetched out using ROI will be used further for accurate feature measurements. As discussed above, breast density leads to breast cancer, and it is not easy to detect cancer in dense breasts. Mammography is one of the modalities to detect masses, especially in dense breasts; it is the best suitable technique for the same [2]. Despite a few shortcomings, mammography holds a sensitivity of approximately 90% in the detection of tumours [3].

Segmentation becomes robust in noisy, blurred images, and low contrast images. Images need to be preprocessed before segmentation. Multiple techniques for segmentation to segment out masses, microcalcification, pectoral muscles, and lesions have been discussed in the paper. All these first filter the image by removing a patient's information and other extra information. The noise and contrast of the image are also modified according to the standards to get appropriate and accurate results for distinguishing benign from malignant. Various features are studied such as shape and size of tumour, texture, intensity, and grey level histogram to figure the growth [4, 5]. Mammographic images have poor contrast and noise. The image may carry both benign and malignant tissue, and the threshold (Otsu image segmentation) technique alone may not be sufficient to distinguish between both of them.

The following are the main points of the paper, based on the novelty and contributions:

(i) To conduct the segmentation of mammograms with the help of different phases such as "2D median filter, CLAHE, FCM on images, removing connected components having less than x pixels."

(ii) To improve the segmentation accuracy by developing the algorithm which optimized the threshold value and specificity of each threshold between the data points.

(iii) For displaying the relevant captions, calculate the best value for threshold position, sensitivity, specificity, area under curve, accuracy, and all false and true positives and negatives.

(iv) To make use of the same algorithm for hybrid segmentation of mammographic images with integration of fuzzy C-Means and CNN model for optimization, which improves the accuracy.

(v) To perform segmentation of mammograms and the readings obtained on sixteen different parameters: distance, Sensitivity, Specificity, ARoC, Accuracy, PPV, NPV, FNR, FPR, FDR, FOR, F1 Score, MCC, BM, and MK.

2. Literature Review

Researchers have done fabulous work in the field of cancer and have learned that if the disease is detected in the early stages, then the mortality rate can be reduced much and the ratio can be improved. The best modality for early detection is mammography, especially in low-contrast and dense breast images. Different authors have worked in this field for the early detection of cancer using various modalities and segmentation techniques that have been listed in this section for better future research and implementation. The study by Bick et al. [6] implements different procedures, such as thresholding, filtering, and region-growing. The mammogram reduces noise from the image, improves the contrast, and then the texture operator fetches the features. All the pixels in the image are traversed, and then the histogram is used to differentiate between an object and nonobject regions. The region-growing technique is implemented to segment out different areas and then label them, and then morphological filtering is performed on the resultant part to remove the irregularities on curve boundaries. A comparative approach was formulated for various feature extraction methods by Nithya et al. [7] to get a better technique for the identification of tumours. For classification, a supervised neural network was used to select a few features for the study as intensity-based, histogram-based, and grey level co-occurrence matrix features. To segment out doubtful lumps from 70 mammographic images taken from database Mini-MIAS, Anitha et al. [8] worked by updating cellular strength to maximum using cellular automata. Seed selection is made using automation along with histogram peak analysis. The appropriateness of the segmented region is studied. The preprocessing of the image is done before carrying out segmentation. The sensitivity is primarily focused upon during the work. GLCM-based sum average features learned to fetch the seed point automatically, which is considered far better than other GLCM-based texture features. The paper also discussed the importance of extracting the mass boundary more precisely to understand the severity of the tumour. Eltoukhy et al. [9] proposed a technique using a multiscale curvelet transform for the recognition of tumours in the early stage. The coefficient value of the input evaluated and based on the result, and the maximum amount used to alter the information into different scales. The different levels used for the study are 2, 3, 5, 6, and 7, and these are all plotted in vector form. In addition to segmentation, supervised classification method (Euclidean distance measure])is used for better feature classification results. The MIAS database was used for validation purposes. The accuracy of 98.59% raised in a 2- scale and 99% built-in 5- scale. Hariraj et al. [10] have worked on the Mini- MIAS database; preprocessing of the images is done to remove noise and spurious content from the image to improve the quality using the Wiener filter method. K-means cluster techniques used to segment out ROI and KNN and SVM techniques are used to classify the attributes among benign and malignant tissue. The data mining technique is widely used in the paper. The rigorousness of the cancer stage predicted, which may further help in the early detection of cancer. Vala and Baxi [11] discussed the benefits of the Otsu

Table 1: Related works in tabular form.

Author and year	Algorithm implemented	Working	Dataset	Research outcome	Advantages
Anitha and Peter, 2015 [8]	MCSU in cellular automata	Mammographic image preprocessed to remove noise, markers, pectoral muscles, and unwanted information. Peak analysis is done and enhanced by CLAHE followed by automatic selection of seed and finally updating the cellular strength using cellular automata	Mini-MIAS, 70 samples	The specificity of the dataset is maintained and improved	The initial seed selection needs no manual interception; the automatic pick of seed point is done
Gupta and Tiwari, 2017 [14]	HM-GRA and CLAHE	The histogram of the mammographic image was generated and the selection of parameters for enhancement was performed. The modification of the histogram is done using the uniform histogram. The grey relational analysis was used to improve the contrast and further normalization and segmentation of ROI was presented	Mini-MIAS 322 samples	The contrast of the image is enhanced to improve minute calcification by decreasing the ratio of false positives	Global and local contrast is improved, and sensitivity and specificity both are taken care of at the same time
Taghanaki et al., 2017	Geometry-based model	The maximum area in the breast contour is covered along with the boundary to be marked for early detection	INbreast, DDSM, MIAS, 197, 353, and 322 samples	The precision of ROI segmentation increased	It works perfectly with multilayered samples having a lot of variations in intensity and edge boundary
Shi et al., 2018 [16]	Gradient weight map, pixel wise clustering, and local text filter	Artifacts removed from original image further segmentation are done based on pixel-wise clustering, followed by detection of the boundary of breast muscles. Finally, a local texture filter is used to detect calcification	MIAS, BCDR, INbreast, 322, 100, and 201 samples	It is immune to noises and can detect calcification in dense breasts too. With few settings, FFDM images can also analyze effectively	The distinction of skin air boundary marked by the proposed algorithm gradient weight map in compassion to another threshold-based algorithm

Table 1: Continued.

Author and year	Algorithm implemented	Working	Dataset	Research outcome	Advantages
Shen et al., 2018 [17]	A genetic algorithm for threshold and segmentation and morphological selection	The genetic algorithm was implemented to study multilevel threshold, segmentation, and classification based on pectoral muscle segmentation done to classify between successful, acceptable, and unacceptable	MIAS, DDSM, and INbreast	The precision of segmentation is higher in comparison to other existing algorithms	The classification between acceptable, unacceptable, and successful. The sensitivity was then checked for unacceptable samples
Hazarika and Mahanta, 2018 [18]	Pectoral muscle removal using region growing	A suppression algorithm is applied, and further, the samples whose results come comparable and close with the hand-drawn segmented mask are distinguished as accepted	Mini-mias and 150 samples	86.67% is the accuracy of acceptable and 5.33% for partially fair	Hand-drawn segmentation mask compare the accuracy of segmentation algorithm given
Alam, et al., 2018 [19]	Segmentation using morphology	Morphological and interpolation operations are used to segment ROI, and further splitting is done based on intensity value. For creating clusters of microcalcification area, ranking is used on the differenced image	DDSM, MIAS,248, and 24 samples	The highest classification accuracy is 94.48% approximate	Dice metric similarity score was calculated to measure the evaluation, and further reference masks were also used
Anitha and Peter, 2015 [20]	KFLS (kernel-based fuzzy clustering)	One the preprocessing of the mammogram is done then ROI, which is segmented out using fuzzy C-means clustering segmentation method	DDSM and 300 samples	94% segmentation precision in terms of sensitivity	As it is based on an intelligent system, i.e., fuzzy clustering, it provides high precision
Touil and Kalti, 2016 [21]	IFBS (iterative fuzzy breast segmentation algorithm)	The image is divided into k clusters to remove the over-segmentation background region and extract perfect ROI	MIAS and 200 samples	As compared to the manual ROI curve, its performance is 60% better	It reduces the over-segmentation of the background

Table 1: Continued.

Author and year	Algorithm implemented	Working	Dataset	Research outcome	Advantages
Kozegar et al., 2018 [22]	DRLSE and OBNLM filter	Region growing with combination with GMM is used, followed by despeckling and fine segmentation. DRLSE was modified in the paper	Ultrasound images and 50 samples	It assumed that seed position is known before as it does not work on images with edges	Also, work where the variance is different
Aggarwal and Chatha, 2019 [23]	Edge detection algorithm is designed on 8-bit grayscale image	Binarization is done to reduce the data reduction step using an edge detection algorithm	MIAS 50 random samples	It reduced the difference between region of interest and background	Data reduction leads to loss of information can be reduced by using multilevel thresholding
Tembhurne et al., 2021 [24]	Computer-aided transfer learning-based deep model for binary classification for breast cancer detection	Multichannel merging methods for making a dual-ensemble architecture	Break-his dataset is used	Ensemble architectures by using pretrained models like Xception and InceptionV3 results in an accuracy of 97.5% is achieved	Combining different algorithms gives better accuracy over measuring accuracy from one algorithm
Malathi et.al., 2021 [25]	The algorithm uses breast CAD scheme feature fusion using CNN deep features network	The abnormality in breast images is scrutinized through deep belief network	CAD images are used	The outcome shows random forest algorithm is giving an accuracy of around 97.51% over the CNN classifier	The algorithm removes the point spread function where low-dose medical CT image restoration and recovers the reconstructed image quality, efficiency, and speed through sparse transform
Fang et al., 2021 [26]	Configuration of the multilayer perceptron (MLP) neural network multilayer perceptron network using backpropagation network	A new training algorithm is proposed based on whale optimization for MLP network	MIAS 332 digitized mammography images	Detection performance is detected using detection rate and identification with the false percentage	Accuracy is achieved as compared to other methods

Table 2: Comparative analysis of various segmentation techniques for mammographic images.

Segmentation techniques	Overview	Advantages	Drawbacks
Thresholding [4]	This method is based on the threshold maximum and the minimum value, corresponding to different peaks depicting different regions. Various techniques have emerged from high threshold values like a balanced histogram, k-means, and otsu maximum variance	No prerequisites are required about image and computation is fast with less complexity	Information with low peaks is not considered and ignored; hence, continuous value is not obtained. In the presence of noise and poor contrast, performance is not up to mark
Region-based [16, 27]	Identical regions are grouped using techniques like region growing, splitting, merging, etc	Better than edge detection about noise immunity. It works better in homogeneous regions	Quite expensive in the context of both time and memory
Clustering [10]	It creates different clusters in the spatial domain. Groups are homogeneous	Best results on overlapped data. Results best for classification. Suitable for real-life applications as it uses fuzzy logic	Clustering validity is a challenging task to be determined. Expensive and sensitive to primary clusters
Edge detection [22]	Work on discontinuity principle and locate regions with minimum sudden change. It is of two types sequential and parallel	Results are better with images having a high contrast value	Give unexpected results with images having many edges or improperly defined boundaries and noisy images
Contour-based segmentation [25]	In this work computer-aided diagnosis (CAD) and fusion via CNN is done for recognition, analysis, and further treatment with the help of RF giving maximum accuracy of 97.51% and minimum error via CNN classifier 95.65%. After segmentation using sparse transform, the algorithm eliminates the attributes of the point spread function	Diagnosis of cancer and accuracy detection is now upgraded with machine learning techniques. This work offers enhanced performance and better implementation results, with accuracy and lesser time in medical CT image restoration, new image recovery with quality and speed	The present CAD system for measuring the accuracy is not recognized and acceptable
Energy function-based technique [28]	It is based on parameterizing the curve, taking some sampling values	Minimum processing is required and if flexible. Fast and efficient	It is not useful in the case of higher dimensions, selection of sampling strategy, topology changes, etc
Breast-region segmentation [29]	This segmentation has three divisions: Classical segmentation, which includes region, threshold, edge-based segmentation utilizing supervised and unsupervised methods using deep learning	U-net was chosen because, unlike other deep learning models, it does not require annotated photos	The method is easily simulated by AI, ML, or DNN persons rather than physicians or biologists

Table 2: Continued.

Segmentation techniques	Overview	Advantages	Drawbacks
Optimized-region growing technique [30]	The optimal features are selected by a hybrid optimization algorithm and classified using a neural network	The velocity updated lion algorithm combines the lion algorithm (LA) with PSO to achieve the best feature selection and weight optimization using NN (VU-LA) and VU-LA is also compared to existing models such as the whale optimization algorithm, grey wolf optimization, firefly, PSO, and LA in terms of performance	The algorithm cannot work well on hazy or faint images
Region of interest [31]	After segmentation to obtain a region of interest (ROI), the images were cleaned up using a median filter and compared using ANN, SVM, and reduced features of SVM	The statistics and grey level cooccurrence matrix are utilized to classify to extract the features from enhanced images using the hybrid SVM-ANN	A better algorithm can be made for enhanced feature detection
Supervised segmentation [32]	The work proposes to constrain the segmentation output when morphological operations to measure performance which uses top-hat and closing operations to evaluate on high-resolution images from the INBreast dataset	It achieves an increase in F1 and in the recall if compared to the training without morphology loss	The evaluations cannot be justified sometimes when images are taken from different sources
Pseudo-color segmentation [33]	Thermal cameras record radiation images, which are then transformed to images of pseudo-colored. All the colors of the thermogram correspond to a specific temperature. The interpretation of breast thermograms is mostly dependent on color analysis and thermogram asymmetry analysis. The work analyses breast thermograms by segmenting the RoI, extracted as a hot region, and then analyzing the color. Abnormalities are shown in the hottest regions by contours	The results compared to diagnosis to ensure infrared thermography is a reliable tool for detecting breast cancer	Sometimes pathologists fail to identify the raw image received by the radiation
Graph segmentation [34]	The automatic segmentation of stained tissue images aid in the detection of malignant disease, and this is done after the separation of contacting cells. Traditional segmentation algorithms face numerous challenges. They present a novel automatic approach for segmenting clustered cancer cells in this work. In the first stage, also use Chan-Vese energy functional to determine cell areas by a modified geometric active contour	By determining high concavity locations along with the cell outlines, contacting cell areas are recovered from the presegmented image	Image profiles can sometimes fail to subclassify breast tumours into additional subtypes, which can help in diagnosis and survival
Variant feature transformation [35]	To subclassify extremely aggressive breast malignancies, the researchers used public transcriptomics datasets in breast cancer cell lines and breast cancer tumours, and associated splice variants	Splicing is becoming more often used as a biomarker for grading tumours	Sometimes splicing may not be an exact method to classify their variants

Table 2: Continued.

Segmentation techniques	Overview	Advantages	Drawbacks
Region –growing algorithm [36]	The work provides a new CRG (conditional region growth) approach for determining correct MC bounds beginning from seed points selected, and they are determined by detecting regional maxima and analyzing superpixels. The region-growing stage is then maintained by the criteria set tuned to MC detection for contrast and shape variation obtained from prior knowledge, which determines the size of the searching area in the neighbourhood. To do qualitative and quantitative analysis for detection of MC and delineation many experiments are done on MC of multiple types	The importance of utilized criteria in the context of MC delineation for better management of breast cancer is demonstrated by a comparison of the proposed technique state-of-the-art	Microcalcifications with their morphology are the indicators of breast cancer when shape and size are considered when malignity degree is to be found out and therefore delineation of MC is done for diagnosis of cancer
Graph-cut algorithm [37]	The mammograms can be classified with BIRADS which subjects to qualitative assessments and face inter-reader and intra-reader variations. Comparative analysis of various articles is done for classification accuracy and computational complexity to design an algorithm for the measurement of breast density using machine learning or deep learning	Breast density measurement via machine learning and deep learning is increasing the rapid development	If the density of the breast increases, then chances of breast cancer also increase which reduces the sensitivity of measuring mammographic density
Watershed algorithm [38]	A novel approach for detection of breast cancer automatically from histopathological images that are composed of binary and eight-class based on a convolutional-LSTM learning model trained on BreakHis dataset and preprocessing using marker-controlled watershed segmentation algorithm and optimized SVM classifier. When MWSA is used with the optimized SVM classifier with Bayesian optimization processed HPIs improve when compared to the CLSTM model's softmax classifier	When compared to present approaches utilizing BreakHis dataset, the methodology achieves great performance for both classifications	Sometimes histopathological images cannot be classified and may not detect breast cancer, therefore, AI and deep learning-based applications are used for automated breast cancer detection with high performances
Fuzzy –C means [39]	This method uses computer-aided cancer diagnosis in its various stages. The first stage performs a reduction of noise and preprocessing contrast enhancement, while the 2nd stage segments the ROI and, the kernel fuzzy C-means method is used for segmentation, and features are extracted and feature selection is employed. For final identification, these selected features are fed into an SVM	The proposed version of the metaheuristic neural network optimization approach improves feature selection and SVM classifier performance. When the methodology was compared against five state-of-the-art methodologies, results revealed the approach was superior	Sometimes Computer-aided design lacks when input is not appropriate

Table 2: Continued.

Segmentation techniques	Overview	Advantages	Drawbacks
Otsu's optimal thresholding [40]	The paper creates a framework that scans the stage of cancer by using the optimized kernel fuzzy clustering algorithm to determine cancer and identify segmented regions in mammogram images. The mammographic images are preprocessed noise-free images obtained by using the hybrid denoising filter algorithm. Data clustering is done to classify data of similar types in one group and of dissimilar types in another group	Results give accuracy & efficiency of the proposed system compared to methods such as K-means, OKFCA, and otsu	Issues like mammograms deviate artifacts, are similar breast tissues and contrast issues on the boundary between skin and air
Fusion of K-means and region growing algorithms [41]	An automated system used by radiologists for diagnostic decisions involving detection of breast masses by using an optimized region growing method, optimal seed point selection, and threshold generation were achieved using grey wolf optimization. The features for both global and local are extracted for shape features, grey level: co-occurrence matrix, run length matrix, texture feature: Local binary pattern, and scale-invariant feature transform	An amalgamation of local and global features with an SVM classifier differentiates benevolent or malignant images and an accuracy of 96% is achieved by GLCM and LBP	The existence rate after detection of cancer affected person cannot be sometimes predicted

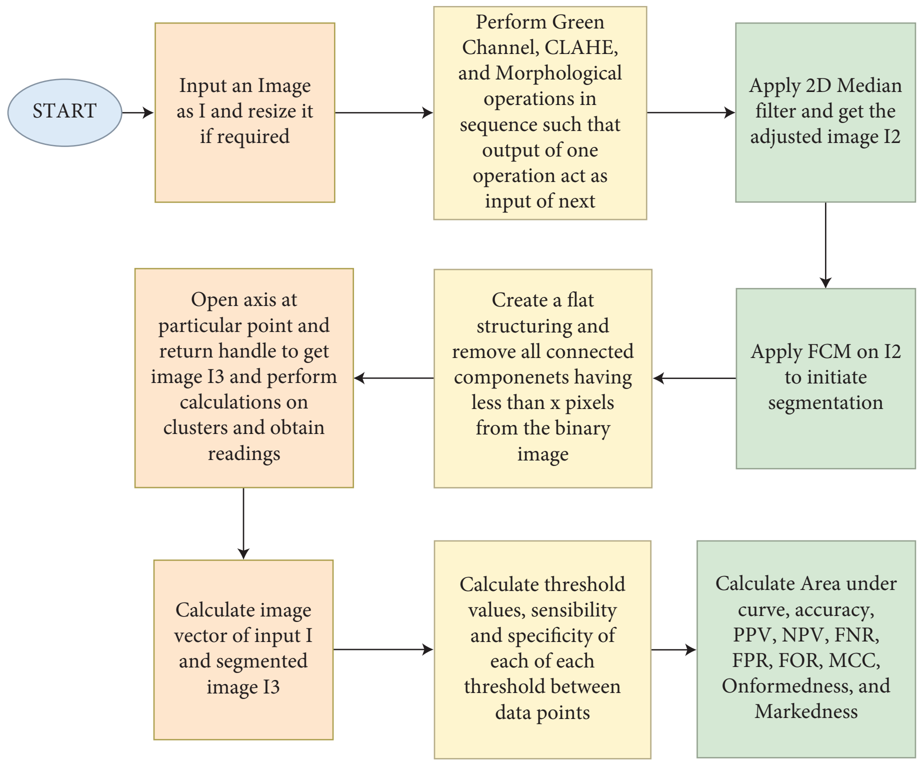

Figure 1: Proposed method for doing segmentation on mammograms.

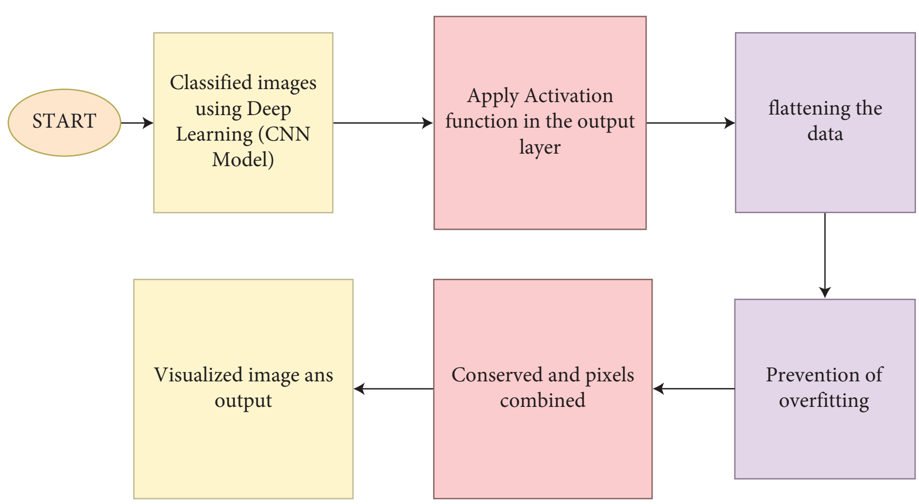

Figure 2: Diagram of the proposed method of classification using CNN models on mammograms.

image segmentation method for thresholding the image for automatic ROI segmentation. On paper this method proves to be simple and easy for calculations. The various Otsu methods discussed as thresholding-based improvised histogram, K-means, etc., along with their advantages and disadvantages. This method is mostly used to reduce the complexity of 1-D and 2-D. Agbley et al. [12] and Singh and Veenadhari [13] gave hybrid technology for segmenting out ROI by merging the region and global thresholding applied to the mammographic images. To eliminate Gaussian noise, Wiener filters were used, and then the resulting image was normalized using the histogram to enhance the quality of input images. Among the above two technologies, a global threshold is used to segment ROI, and the segmented region is extracted by region merging. The implementation and testing was done on 50 mammographic images and the specificity of the research was 82%. The related works in tabular form are shown in Table 1.

```
To design a Graphical User Interface
IM = imresize(I,.3)//Input an image by renaming it as IM and resizing it if required.
GR = ImageIn(:,:,3);//Apply Complement using the green channel on input image IM.
GRC = imgcomplement(GR);
axxes(handlesi.greenimg_channel);
set(imgshow(GRC));
CLH = adaptohisteq(GRC);//To apply CLAHE on GRC to receiveCLH % contrast limited image.
set(imshow(CLH));
SE = strel("ball",8,8);//Perform structuring of an element with the specified neighbourhood (8).
mgopen = imgopen(CLH,SE);//Do morphological on binary image CLH with structuring element 'sse'.
gordisk = CLH - gimgopen;//Replace optic disk by appling a 2D Median Filter.
medfillt = medfillt2(gordisk);
backgroundimg = imgopen(medfillt, strell("disk",160));
IM2 = GC1;//Eliminate background for adjustment of image to retrieve GC1.
IM2 = double(IM2);//The I/P image (GC1) by using Fuzzy C-Means, does image segmentation.
//Execute the above segmentation to construct a flat structure element with in the specified neighbourhood.
backgroundimg = imopenimg(IMMM, strell("disk", 46));
//Eliminate all connected components having less than 40 pixels to create new binary image I5 from a binary image and is called as an area opening
I5=IMMMM-backgroundimg;
I5 = bwareaopenimg(I5,30);
axess(handless.segmented_img);//Open an axis at the specified position and return a handle to it.
set(imshowimg(I5));
set(LTprojectt.segmented_img,"Userdata",I5);
ffcmm1=(['The value of Cluster1 = 'num2str(cccc1)]);//Retrieve the final image I5 to find cluster.
ffcmm2=(['The value of Cluster2 = 'num2str(cccc2)]);
classIM_1 = Imgg(:);//Find image vectors of input image (IM) and segmented image (I5)
  classI5_2 = Imgg1(:);
//To detect errors set all default parameters
//Evaluate the threshold values among the data points.
% Sort data points %
  ss_data = unique(sort([class_1; class_2]));
% Del NaN values %
  ss_data(isnan(ss_data)) = [];
% Cal difference between consecutive points %
  dd_data = diff(ss_data);
% Cal last point %
  dd_data(length(d_data)+1,1) = dd_data(length(d_data));
% Cal first point %
  thresh(1,1) = ss_data(1) - dd_data(1);
% Cal Threshold %
  thres(2:len(s_data)+1,1) = s_data + d_data./2;
  cur = zeross(sizeof(thresh,1),2);//Find sensibility and specificity of every threshold value
  dis = zeross(sizeof(thresh,1),1);
  for idd_t = 1:1:len(thresh)
    TruePositive = len(find(class2≥thresh(idd_t)));
    FalsePositive = len(find(class1 ≥ thresh(idd_t)));
    FalseNegative = len(find(class2 <thresh(idd_t)));
    TrueNegative = len(find(class1 <thresh(idd_t)));
    S = TruePositive/(TruePositive + FalseNegative);
    SP = curve(idd_t1,2) = TrueNegative/(TrueNegative + FalsePositive);
  //Calculate distance between every point and optimum point ranging [0,1]
  distance(idd_t1) = sqrt((1-curve(idd_t1,1))^2+(curve(idd_t1,2)-1)^2);
Calculate the best value for threshold position, Sensitivity, Specificity, Area under curve, Accuracy, all false and true positives and negatives.
```

Algorithm 1: Process to design a Graphical User Interface of the proposed method

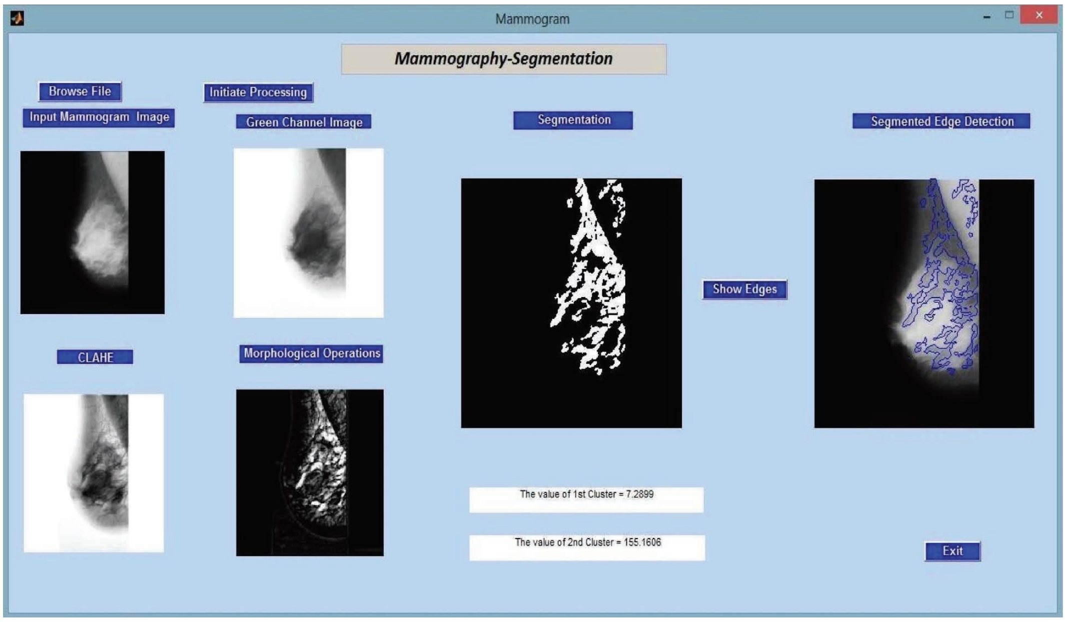

FIGURE 3: First image of the MIAS database "mdb001.pgm" undergoing segmentation.

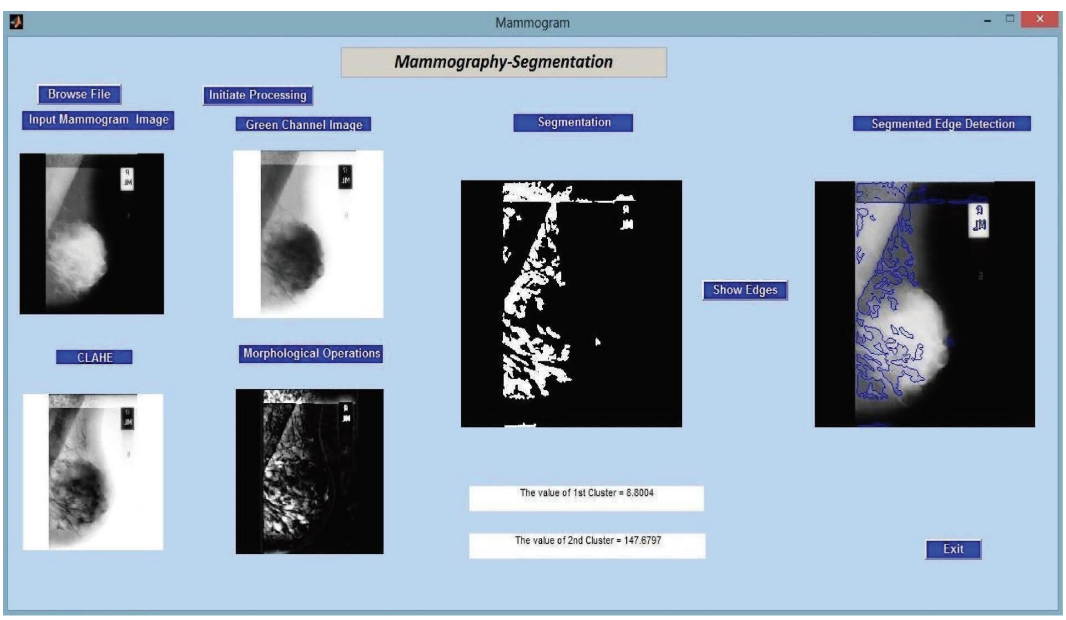

FIGURE 4: Second image of the MIAS database "mdb002.pgm" undergoing segmentation.

3. Comparision of Segmentation Techniques for Mammographic Images

There are many works that follow segmentation techniques of masses in mammographic images. Table 2 is highlighting the key-points and overview and advantages and major drawbacks of various works. The key objective is to point out the advantages and disadvantages of the various approaches.

4. Proposed Methodology

Image segmentation refers to the techniques of dividing an image into different regions. The most effective method to

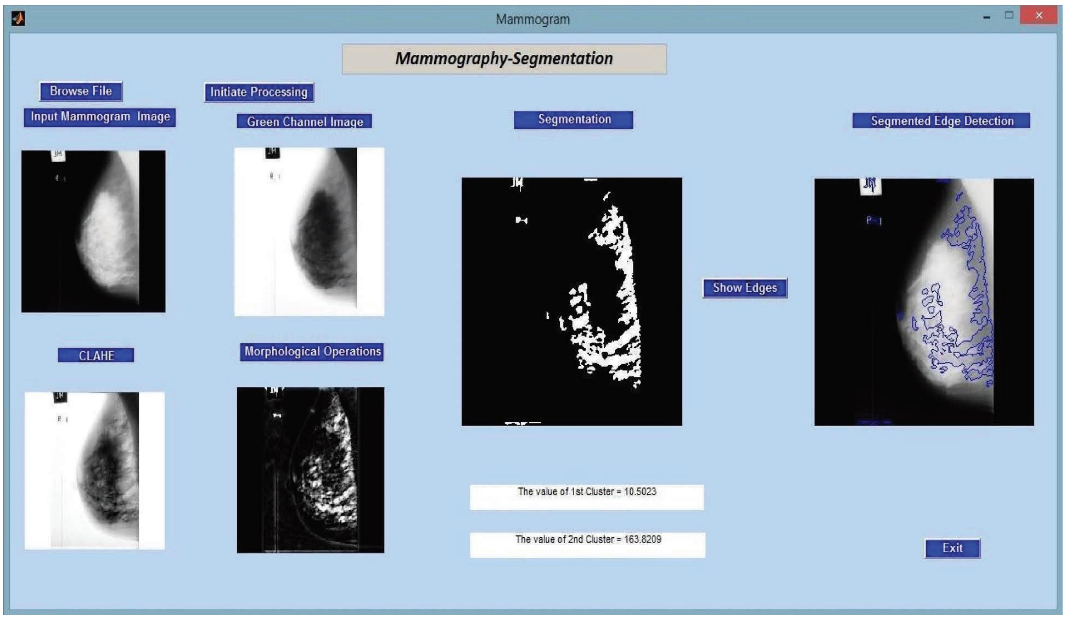

FIGURE 5: Third image of the MIAS database "mdb003.pgm" undergoing segmentation.

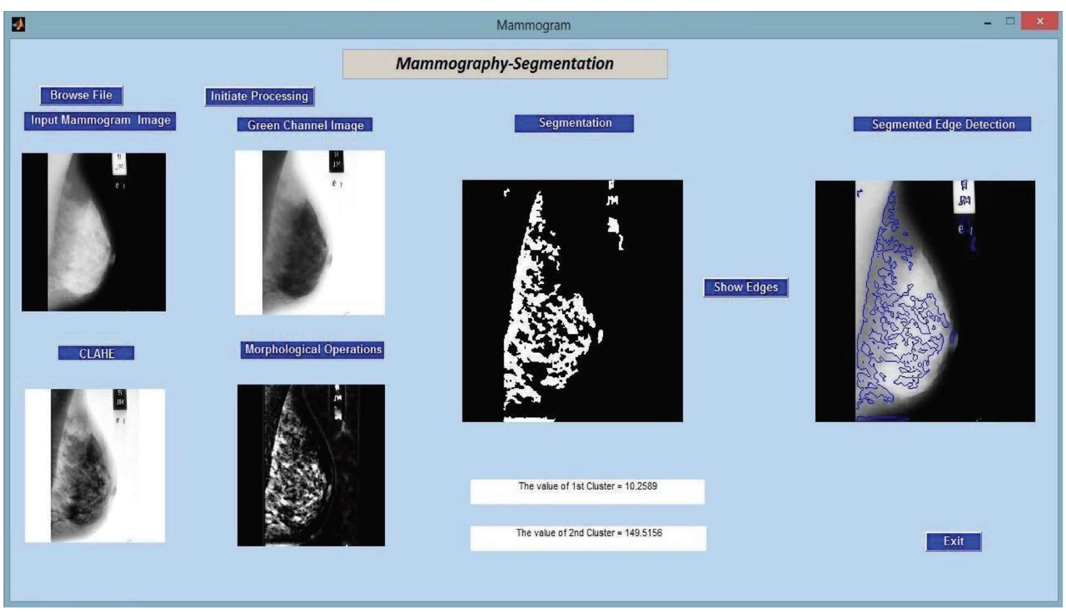

FIGURE 6: Fourth image of the MIAS database "mdb004.pgm" undergoing segmentation.

analyze anatomical structure in medical is "region growing method" [42, 43]. But it does not give proper and more accurate results if it directly applies to the input images that are having noisy and low contrast. We have proposed algorithm could be applied on the mammographic images more effectively in such condition.

The proposed method developed to conduct the segmentation of mammograms is detailed in the flowchart shown in Figures 1 and 2.

The algorithm of the implemented work, is as below (Algorithm 1):

5. Experiment and Results

The mean based region growing segmentation (MRGS) method [44] is presented which has the improvement over ordinary region growing (RG) method with regard to the selection of threshold.

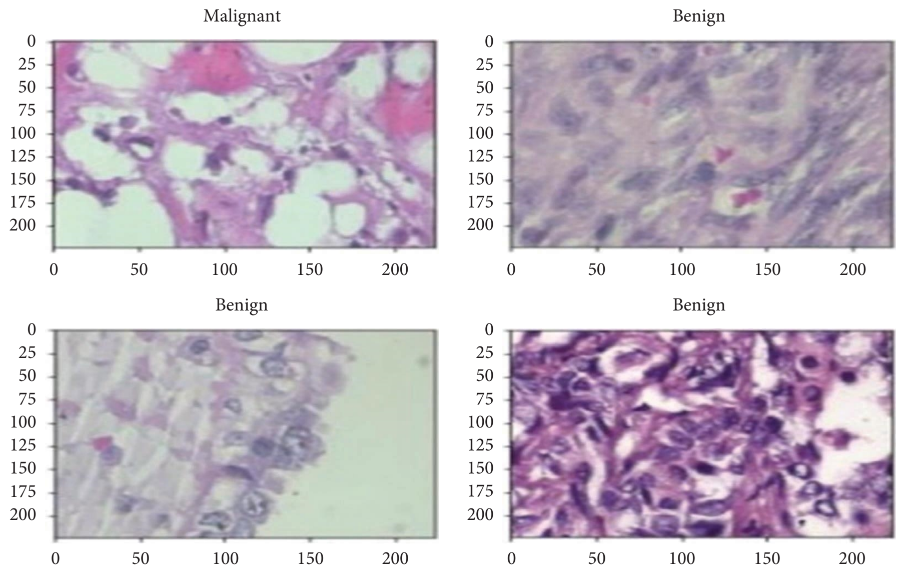

FIGURE 7: Image of the MIAS database undergoing classification for finding out the benign and malignant images.

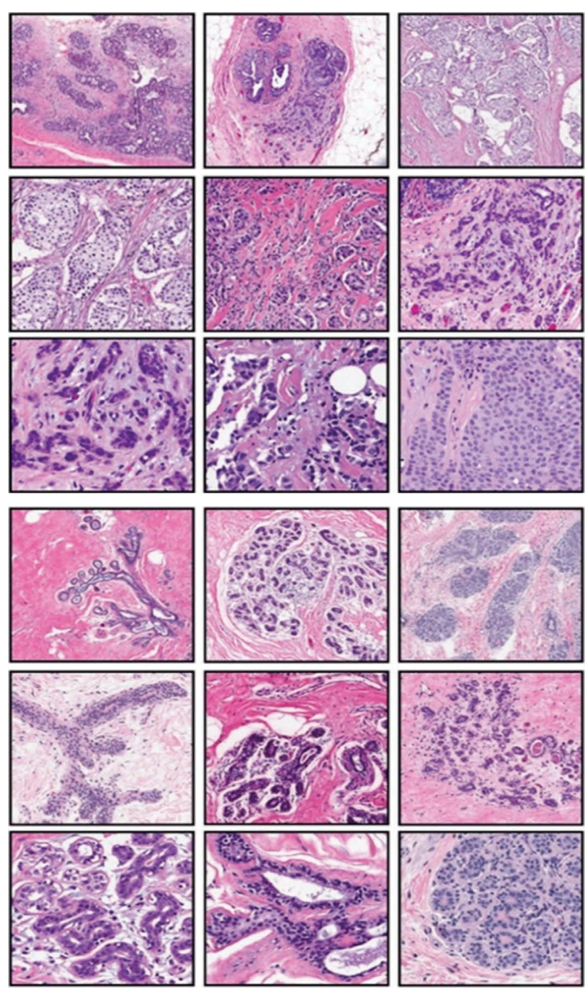

FIGURE 8: Malignant images, as well as benign images of all image sets of the MIAS database undergoing classification to find the benign and malignant image.

Table 3: Observations of various parameters on different images from the MIAS database.

Image name	Distance	Threshold	Sensitivity	Specificity	ARoC	Accuracy	PPV	NPV	FNR	FPR	FDR	FOR	*F*1 score	MCC	BM	MK
mdb001.pgm	0.9101	219	0.0899	1.0000	0.3710	0.9245	0.9948	0.9239	0.9101	0.000	0.0052	0.0761	0.1649	0.2873	0.0898	0.9187
mdb002.pgm	0.8915	217	0.1085	0.9998	0.3893	0.9259	0.9804	0.9254	0.8915	0.0002	0.0196	0.0746	0.1953	0.3132	0.1083	0.9057
mdb003.pgm	0.9193	212	0.0807	0.9970	0.3636	0.9210	0.7113	0.9230	0.9193	0.0030	0.2887	0.0770	0.1449	0.2220	0.0777	0.6343
mdb004.pgm	0.8771	217	0.1229	0.9998	0.4006	0.9270	0.9826	0.9265	0.8771	0.0002	0.0174	0.0735	0.2184	0.3339	0.1227	0.9091
mdb005.pgm	0.7749	216	0.2251	0.9996	0.4808	0.9354	0.9822	0.9345	0.7749	0.0004	0.0178	0.0655	0.3662	0.4539	0.2247	0.9166
mdb006.pgm	0.7525	216	0.2475	0.9996	0.4930	0.9372	0.9837	0.9362	0.7525	0.0004	0.0163	0.0638	0.3954	0.4768	0.2471	0.9200
mdb007.pgm	0.8649	215	0.1351	0.9993	0.4084	0.9276	0.9441	0.9274	0.8649	0.0007	0.0559	0.0726	0.2364	0.3422	0.1344	0.8715
mdb008.pgm	0.8454	211	0.1546	0.9958	0.4227	0.9260	0.7697	0.9287	0.8454	0.0042	0.2303	0.0713	0.2575	0.3241	0.1504	0.6984
mdb009.pgm	0.8504	218	0.1496	0.9999	0.4158	0.9294	0.9931	0.9286	0.8504	0.0001	0.0069	0.0714	0.2601	0.3713	0.1496	0.9216
mdb010.pgm	0.8933	218	0.1067	0.9999	0.3817	0.9258	0.9903	0.9252	0.8933	0.0001	0.0097	0.0748	0.1926	0.3123	0.1066	0.9155
mdb011.pgm	0.8481	218	0.1519	0.9999	0.4175	0.9295	0.9932	0.9287	0.8481	0.0001	0.0068	0.0713	0.2635	0.3741	0.1518	0.9219
mdb012.pgm	0.8531	216	0.1469	0.9996	0.4142	0.9289	0.9729	0.9283	0.8531	0.0004	0.0271	0.0717	0.2553	0.3634	0.1466	0.9012
mdb013.pgm	0.8569	211	0.1431	0.9958	0.4111	0.9251	0.7557	0.9228	0.8569	0.0042	0.2443	0.0722	0.2406	0.3081	0.1389	0.6835
mdb014.pgm	0.8517	216	0.1483	0.9996	0.4179	0.9290	0.9732	0.9284	0.8517	0.0004	0.0268	0.0716	0.2574	0.3653	0.1480	0.9016
mdb015.pgm	0.8825	214	0.1175	0.9987	0.3879	0.9256	0.8929	0.9260	0.8825	0.0013	0.1071	0.0740	0.2076	0.3084	0.1162	0.8189

Figure 3 shows the first image of MIAS database "mdb001.pgm" given as input to the developed method. The obtained images by applying different approaches displayed under relevant captions within the frame.

Figure 4 shows the second image of the MIAS database "mdb002.pgm" being given as an input to the developed method.

Figure 5 shows the third image of the MIAS database "mdb003.pgm" been given as an input to the developed method.

Figure 6 shows the fourth image of the MIAS database "mdb004.pgm" been given as an input to the developed method.

The images in Figures 7 and 8 are classified images out of the pixels, combined using PYPLOT and bypassing the input data through 3 layered CNN models with alternated max pool layers to combine the pixels of similar density. They used ReLu activation function in the output layer after flattening the dataset with the dropout to prevent the NN overfitting in the predictions. On the predictions, the original input is conserved, and pixels are combined using PYPLOT to create a visual image of the flattened input data to review the visualized image and the output.

The first fifteen images from the MIAS database were taken for performing segmentation of mammograms, and the readings were obtained on sixteen different parameters, as shown in Table 3.

6. Conclusion and Future Scope

This paper discussed the method for performing the segmentation of mammograms. More than fifteen images of the MIAS database are tested to assure the worth of the conducted research work. The undertaken research work proved that the combined approaches provide improved segmentation accuracy. Accuracy related to segmentation has a vital role in categorizing cancer as benign or malignant. The adopted preprocessing methods assist in procuring enhanced segmentation outcomes. In future work, images from different databases are used to perform segmentation, and the number of relevant parameters (distance, sensitivity, specificity, ARoC, accuracy, PPV, NPV, FNR, FPR, FDR, FOR, *F*1 Score, MCC, BM, and MK) increased. Even other types of breast images bearing different properties are used, such as ultrasound and thermography. The model can be more optimized with PCA or applying SVM at the output layer for confident results, and we can say that images can produce a huge number of dimensions. So, we can limit the dimensions with PCA with a minute compromise in accuracy but optimize code. The proposed model underwent different steps detect all the errors, evaluate the threshold values among the data points, find the sensibility and specificity of every threshold value, and also calculate the best value for threshold position, sensitivity, specificity, area under curve, accuracy, and all false and true positives and negatives. Later, the classification is done for finding out the benign and malignant images. The proposed model helps in detecting breast cancer, which reduces the need for breast removal and also the need of chemotherapy, saving the lives at earlier stage.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This study is supported via funding from Prince Sattam Bin Abdulaziz University project number (PSAU/2023/R/1444).

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics," *CA: A Cancer Journal for Clinicians*, vol. 69, no. 1, pp. 7–34, 2019.

[2] M. Mustra, M. Grgic, and R. M. Rangayyan, "Review of recent advances in segmentation of the breast boundary and the pectoral muscle in mammograms," *Medical, & Biological Engineering & Computing*, vol. 54, no. 7, pp. 1003–1024, 2016.

[3] C. Dromain, B. Boyer, R. Ferré, S. Canale, S. Delaloge, and C. Balleyguier, "Computed-aided diagnosis (CAD) in the detection of breast cancer," *European Journal of Radiology*, vol. 82, no. 3, pp. 417–423, 2013.

[4] J. Dabass, S. Arora, R. Vig, and M. Hanmandlu, "Segmentation techniques for breast cancer imaging modalities- A review," in *Proceedings of the 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 658–663, Noida, India, January 2019.

[5] A. U. Haq, J. P. Li, A. Saboor et al., "Detection of breast cancer through clinical data using supervised and unsupervised feature selection techniques," *IEEE Access*, vol. 9, Article ID 22090, 2021.

[6] U. Bick, M. L. Giger, R. A. Schmidt, R. M. Nishikawa, D. E. Wolverton, and K. Doi, "Automated segmentation of digitized mammograms," *Academic Radiology*, vol. 2, no. 1, pp. 1–9, 1995.

[7] R. Nithya and B. Santhi, "Comparative study on feature extraction method for breast cancer classification," *Journal of Theoretical and Applied Information Technology*, vol. 33, no. 2, pp. 220–226, 2011.

[8] J. Anitha and J. D. Peter, "Mammogram segmentation using maximal cell strength updation in cellular automata," *Medical, & Biological Engineering & Computing*, vol. 53, no. 8, pp. 737–749, 2015.

[9] M. M. Eltoukhy, I. Faye, and B. B. Samir, "Breast cancer diagnosis in digital mammogram using multiscale curvelet transform," *Computerized Medical Imaging and Graphics*, vol. 34, no. 4, pp. 269–276, 2010.

[10] V. Hariraj, K. Wan, Vikneswaran, and I. Zunaidi, "An efficient data mining approaches for breast cancer detection and segmentation in a mammogram," *J. Adv. Res. Dyn. Control Syst.*vol. 9, no. 3, pp. 185–194, 2017.

[11] H. J. Vala and A. Baxi, "A review on Otsu image segmentation algorithm," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, vol. 2, no. 2, pp. 387–389, 2013.

[12] B. L. Agbley, J. Li, A. U. Haq et al., "Multimodal melanoma detection with federated learning," in *Proceedings of the 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing*

(ICCWAMTIP), pp. 238–244, IEEE, Chengdu, China, December 2021.
[13] N. Singh and S. Veenadhari, "Breast cancer segmentation using global thresholding and region merging," *International Journal on Computer Science and Engineering*, vol. 6, no. 12, pp. 292–297, 2018.
[14] B. Gupta and M. Tiwari, "A tool supported approach for brightness preserving contrast enhancement and mass segmentation of mammogram images using histogram modified grey relational analysis," *Multidimensional Systems and Signal Processing*, vol. 28, no. 4, pp. 1549–1567, 2017.
[15] S. A. Taghanaki, Y. Liu, B. Miles, and G. Hamarneh, "Geometry-based pectoral muscle segmentation from MLO mammogram views," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 11, pp. 2662–2671, 2017.
[16] P. Shi, J. Zhong, A. Rampun, and H. Wang, "A hierarchical pipeline for breast boundary segmentation and calcification detection in mammograms," *Computers in Biology and Medicine*, vol. 96, pp. 178–188, 2018.
[17] R. Shen, K. Yan, F. Xiao, J. Chang, C. Jiang, and K. Zhou, "Automatic pectoral muscle region segmentation in mammograms using genetic algorithm and morphological selection," *Journal of Digital Imaging*, vol. 31, no. 5, pp. 680–691, 2018.
[18] M. Hazarika and L. B. Mahanta, "A novel region growing based method to remove pectoral muscle from MLO mammogram images," *Lecture Notes in Electrical Engineering*, vol. 443, pp. 307–316, 2018.
[19] N. Alam, A. Oliver, E. R. E. Denton, and R. Zwiggelaar, "Automatic segmentation of microcalcification clusters," *Communications in Computer and Information Science*, vol. 894, pp. 251–261, 2018.
[20] J. Anitha and J. D. Peter, "Mass segmentation in mammograms using a kernel-based fuzzy level set method," *International Journal of Biomedical Engineering and Technology*, vol. 19, no. 2, pp. 133–153, 2015.
[21] A. Touil and K. Kalti, "Iterative fuzzy segmentation for an accurate delimitation of the breast region," *Computer Methods and Programs in Biomedicine*, vol. 132, pp. 137–147, 2016.
[22] E. Kozegar, M. Soryani, H. Behnam, M. Salamati, and T. Tan, "Mass segmentation in automated 3-D breast ultrasound using adaptive region growing and supervised edge-based deformable model," *IEEE Transactions on Medical Imaging*, vol. 37, no. 4, pp. 918–928, 2018.
[23] A. Aggarwal and D. Chatha, "To design a mammogram edge detection algorithm using an artificial neural network (ANN)," *International Journal of Distributed Artificial Intelligence (IJDAI)*, vol. 11, no. 1, pp. 34–43, 2019.
[24] J. V. Tembhurne, A. Hazarika, and T. Diwan, "BrC-MCDLM: breast Cancer detection using Multi-Channel deep learning model," *Multimedia Tools and Applications*, vol. 80, no. 21-23, Article ID 31647, 2021.
[25] M. Malathi, P. Sinthia, F. Farzana, and G. Aloy Anuja Mary, "Breast cancer detection using active contour and classification by deep belief network," *Materials Today Proceedings*, vol. 45, no. 2021, pp. 2721–2724, 2021.
[26] H. Fang, H. Fan, S. Lin, Z. Qing, and F. R. Sheykhahmad, "Automatic breast cancer detection based on optimized neural network using whale optimization algorithm," *International Journal of Imaging Systems and Technology*, vol. 31, no. 1, pp. 425–438, 2021.
[27] Z. Chen and R. Zwiggelaar, "A combined method for automatic identification of the breast boundary in mammograms," in *Proceedings of the 2012 5th International Conference on BioMedical Engineering and Informatics*, pp. 121–125, Chongqing, China, October 2012.
[28] R. Poonguzhali, S. Ahmad, P. Thiruvannamalai Sivasankar et al., "Automated brain tumor diagnosis using deep residual U-net Segmentation㯀odel," *Computers, Materials & Continua*, vol. 74, no. 1, pp. 2179–2194, 2023.
[29] E. Michael, H. Ma, H. Li, F. Kulwa, and J. Li, "Breast cancer segmentation methods: current status and future potentials," *BioMed Research International*, vol. 2021, Article ID 9962109, 29 pages, 2021.
[30] E. Shivhare and V. N. Saxena, "Breast cancer diagnosis from mammographic images using optimized feature selection and neural network architecture," *International Journal of Imaging Systems and Technology*, vol. 31, no. 1, pp. 253–269, 2021.
[31] T. S. Lim, K. G. Tay, A. Huong, and X. Y. Lim, "Breast cancer diagnosis system using hybrid support vector machine-artificial neural network," *International Journal of Electrical and Computer Engineering*, vol. 11, p. 3059, 2021.
[32] M. Tardy and D. Mateus, "Morphology-based losses for weakly supervised segmentation of mammograms," in *Proceedings of the Medical Imaging with Deep Learning*, Lübeck, Germany, July 2021.
[33] A. T. Wakankar and G. R. Suresh, "Automatic diagnosis of breast cancer using thermographic color analysis and SVM classifier," in *Intelligent Systems Technologies and Applications 2016*, J. Corchado Rodriguez, S. Mitra, S. Thampi, and ES. El-Alfy, Eds., vol. 530, New York, NY, USA, Springer Cham, 2016.
[34] A. Mouelhi, M. Sayadi, and F. Fnaiech, "Automatic segmentation of clustered breast cancer cells using watershed and concave vertex graph," in *Proceedings of the 2011 International Conference on Communications, Computing and Control Applications (CCCA)*, pp. 1–6, Hammamet, Tunisia, March 2011.
[35] J.-P. Villemin, C. Lorenzi, M. S. Cabrillac, A. Oldfield, W. Ritchie, and R. F. Luco, "A cell-to-patient machine learning transfer approach uncovers novel basal-like breast cancer prognostic markers amongst alternative splice variants," *BMC Biology*, vol. 19, no. 1, pp. 70–19, 2021.
[36] A. Touil, K. Kalti, P.-H. Conze, B. Solaiman, and M. A. Mahjoub, "A new conditional region growing approach for microcalcification delineation in mammograms," *Medical, & Biological Engineering & Computing*, vol. 59, pp. 1–20, 2021.
[37] S. D. Pawar, K. Kr Sharma, and S. G. Sapate, "Advances in machine learning and deep learning approaches for mammographic breast density measurement for breast cancer risk prediction: an overview," *Design of Intelligent Applications Using Machine Learning and Deep Learning Techniques*, Taylor and Francis, Oxford, UK, pp. 125–143, 2021.
[38] F.. Demir, "DeepBreastNet: a novel and robust approach for automated breast cancer detection from histopathological images," *Biocybernetics and Biomedical Engineering*, vol. 41, 2021.
[39] J. Huaping, Z. Junlong, and A. M. Norouzzadeh Gil Molk, "Skin cancer detection using kernel fuzzy C-means and improved neural network optimization algorithm," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9651957, 12 pages, 2021.
[40] S. Safdar, M. Rizwan, T. R. Gadekallu et al., "Bio-imaging-based machine learning algorithm for breast cancer detection," *Diagnostics*, vol. 12, no. 5, p. 1134, 2022.
[41] A. Ashok, D. Vijayan, and R. Lavanya, "Computer-aided mass segmentation in mammogram images using Grey wolf

Optimized Region growing technique," in *Proceedings of the 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 1082–1087, IEEE, Tirunelveli, India, June 2021.

[42] S. Ahmad, H. A. M. Abdeljaber, J. Nazeer, M. Y. Uddin, V. Lingamuthu, and A. Kaur, "Issues of clinical identity verification for healthcare applications over mobile terminal platform," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 6245397, 10 pages, 2022.

[43] J. Srivastava, S. Routray, S. Ahmad, and M. M. Waris, "Internet of medical things (IoMT)-based smart healthcare system: trends and progress," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7218113, 17 pages, 2022.

[44] S. Ahmad, S. Khan, M. Fahad AlAjmi et al., "Deep learning enabled disease diagnosis for secure internet of medical things," *Computers, Materials & Continua*, vol. 73, no. 1, pp. 965–979, 2022.