Expert Systems With Applications 185 (2021) 115580

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

An automatic and efficient technique for tumor location identification and classification through breast MR images

Poonam Jaglan [*], Rajeshwar Dass, Manoj Duhan

ECED, DCRUST, Murthal, Haryana 131039, India

ARTICLE INFO

Keywords:
De-noising
Breast Tumor
Tumor Segmentation
Tumor Classification
Support Vector Machine

ABSTRACT

Aim: The basic objective of this paper is to develop a single structured algorithm to classify the breast tissues into normal or abnormal.

Material & Methodology: For this study, the breast MR images dataset of 448 images collected from Healthmap diagnostics centre, PGIMS, Rohtak, India. The proposed algorithm consists of several steps i.e. an integrated (Median Wiener & Median) filtering technique is used for de-noising; breast boundary region extraction via selection of nipple and mid- sternum points to make the image rotation invariant; determined the tumor region intensity by using morphological operations & hole filling; classify the normal and abnormal breast tissues by SVM using 14 texture features extracted through GLCM & 13 morphological or kinetic features; evaluated the exact location as well as area of abnormal tissues.

Results: The proposed algorithm has been evaluated statistically as well as visually. The quality parameters achieved are accuracy, sensitivity and specificity with values 0.937, 0.956 and 0.872 respectively. The Jaccard Index coefficient achieved is 0.921, which indicates promising overlap between the predicted tumor and the manually done image by the radiologist so called ground truth image.

Conclusion: This work may be taken as a second opinion by the radiologists. The evaluated results may give a basic foundation for optimization by selecting the features more precisely and also different evolutionary algorithms using multi-classifiers can be designed in future.

1. Introduction

MRI is the safest technique for the scanning and early detection of breast tumors as it does not involve any kind of radiation (Jaglan, Dass, & Duhan, 2019b). MRI uses both magnetic fields and radio-frequency pulses to stimulate the signal from the object (Losurdo et al., 2018). It reads the data through magnetic gradients and places the same into k-space which is further translated into spatial domain for image formation (Warner et al., 2001). Earlier detection of breast cancer especially in high risk women reduces the mortality rate and even the need of biopsy/ chemotherapy (Saadatmand et al., 2019). MRI can be used both for diagnostic (tumor, bone damage, assess and surgery planning) and research purposes (Schoub, 2018). Image segmentation i.e. finding the tumor ROI (Region of Interest) is central in breast MRI analysis (Cai, Liu, Peng, Wu, & Li, 2014). The breast anatomy contains fatty tissues as well as dense fibro-glandular tissues which makes it more heterogeneous so analysis of breast MR Images is a quite challenging (Bouchebbah & Slimani, 2019). The breast tumor extraction through MR images is the most tedious and time taking task performed by experts/ radiologists whereas their experience plays the key role to manually rule out the abnormal regions (Jaglan, Dass, & Duhan, 2019c). Even well-experienced physician may sometimes get inter-observer variation rates during the image interpretation. Therefore, there is a need to investigate a lesion detection method with high sensitivity and low false positive detections in order to achieve a fully automated CAD system (Cheng, Shan, Ju, Guo, & Zhang, 2010). However, the higher value of sensitivity is presented by the fusion of MRI and mammography methods shows in the *meta*-analysis (Saadatmand et al., 2019). Few studies available in investigating fully automated lesion detection systems in breast MRI are given below in Table 1.

This paper is further arranged as: Section II explains the material & methodology used. Section III explains the experimental results & analysis. Section IV gives the qualitative analysis of the proposed work and presented the conclusion in section V.

* Corresponding author.
E-mail address: jaglanpoonam@gmail.com (P. Jaglan).

https://doi.org/10.1016/j.eswa.2021.115580
Received 23 January 2020; Received in revised form 16 July 2020; Accepted 7 July 2021
Available online 12 July 2021

Table 1
Related work along with key findings.

Author	Description	Key Findings
Cui et al. (2009)	Determined the external as well as internal markers automatically for marker controlled watershed segmentation using Gaussian mixture modelling.	Achieved 62.6%±9.1% & 61.0%±11.3% values of overlap ratios by comparing with the results of manual segmentations respectively.
Nie et al. (2008)	Predicted the breast MRI diagnosis using quantitative morphologic and texture features analysis method.	High accuracy, AUC is observed as 86%.
Sinha et al. (1997)	Improved the specificity of breast MR by using LDA of shape, contrast uptake, and texture descriptors.	Recall and specificity are 93% & 95% respectively (when 10 features were combined)
Elif Derya Ubeyli (Übeyli, 2007)	Four neural network (NN) methods i.e. Probabilistic NN, Multilayer Perceptron NN, Recurrent NN, Combined NN were compared with SVM.	SVM achieved the highest classification accuracy of 99.54%.
Maglogiannis, Zafiropoulos, and Anagnostopoulos (2009)	Compared SVM classifier with ANN (Artificial NN) and Bayesian classifiers for prognosis & diagnosis of breast tumor.	The obtained accuracies were 97.54%, 92.80% and 97.90%, respectively while SVM exhibiting high values of specificity i.e. 97.67% and sensitivity i.e. 97. 84% than other methods.
Karabatak and Ince (2009)	Presented a CAD system built of AR (Association Rules) & Neural Network for detecting breast cancer.	95.6% was the correct classification rate.
Chen, Yang, Liu, and Liu (2011)	Proposed an algorithm named RS_SVM based on rough set & SVM classifier	The reported average accuracy was 96.87%.
Kele, Kele, and Yavuz (2011)	Expert system for diagnosis of breast cancer abbreviated as Ex-DBC is developed.	Specificity, sensitivity, PPV & NPV were 97%, 76%, 96% & 81% respectively.
Marcano-Cedeño, Quintanilla-Domínguez, and Andina (2011))	Presented biological *meta*-plasticity property based ANN (AMMLP) for breast cancer classification.	Total classification accuracy was 99.26%.
Şahan, Polat, Kodaz, and Güneş (2007)	Hybridization of fuzzy-artificial immune system with KNN is implemented.	The accuracy value of 99.14% is obtained.
Abonyi and Szeifert (2003)	A technique called (SFC) Supervised Fuzzy Clustering is applied.	95.57% accuracy obtained.
Muthu Rama Krishnan, Banerjee, Chakraborty, Chakraborty, and Ray (2010)	SVM Classifier	Overall accuracies evaluated for two datasets are 99.385% and 93.726%.
Stoean and Stoean (2013)	Proposed two-step hybridized methodology based on Evolutionary Algorithms (EA) and SVM.	Correct classification for diagnostics & prognostic were 97% and 79% respectively.
Mu and Nandi (2007)	Given the SVM with various parameters of tuning system and the SOM–RBF classifier.	98.6%. is the value given by the L2-SVM classifier by GDSEE parameter of tuning system.
Moftah et al. (2014)	Intensity distribution used as feature.	90.83% is the given accuracy of ROI.
Al-faris, Ngah, Ashidi, Isa, and Lutfi (2012)	Applied a modified technique by SRG (Seeded Region Growing) based on PSO.	AUC is reported as 0.95.
Huang (2014)	Proposed a Level Set Method (LSM) along with shape model for lesion segmentation of breast MRI.	The sensitivity & specificity values are 0.952 and 0.999 respectively.
Al-Faris, Kalthum Ngah, Ashidi Mat Isa, and Lutfi Shuaib (2012)	An integration method is presented consists of level set active contouring method with morphological thinning operation.	An accuracy of 0.9607 is obtained. AUC is 0.9905, Jaccard Index is 0.9275.
Al-faris, Ngah, Isa, and Shuaib (2015)	Skin border regions are excluded from Breast MRI images through an automatic combined approach.	The average of sensitivity is 86% and specificity is 97%.
Fooladivanda, Shokouhi, and Ahmadinejad (2017b)	Presented a decision-making framework using Localized atlas based Segmentation and SVM for classification the breast tissues into complex or simple.	Jaccard Index Coefficient and Dice Similarity Coefficient are given as 92.9 & 96.3 respectively whereas total overlap was 97.4%. The values of false negative, and false positive were 2.61% and 4.77% respectively.

2. Material & methodology

2.1. Breast MRI data acquisition

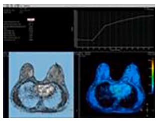
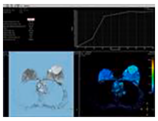
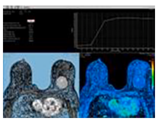
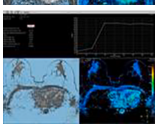
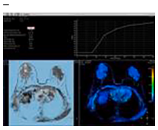
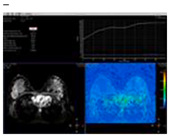
In this study, Breast MR Image dataset of 448 cases (318 abnormal and 130 normal) collected from the Healthmap diagnostics centre, PGIMS, Rohtak. MRI sequences obtained using a 1.5 Tesla super conducting MRI unit (GE Medical Systems) to acquire multi-parametric breast images. Dynamic contrast MRI of both breasts is performed in open array coil after giving contrast IV using a M3D-FSPGR sequence with post contrast subtracted and MIP images. The dynamic data sets were acquired by high resolution T1W-FSPGR fat saturated and SSFSE-T2W breathhold section in axial & sagittal planes and T2W/ Tirm coronal planes. Table 2 gives the detail of twelve cases as per the reports generated by the experts where the BIRAD staging is done on the basis of signal–Intensity curve. The manually segmented tumor ROI done by the experts becomes the ground truth image & the statistical parameters like area, perimeter etc. are evaluated given in Table 2 while the normal breast MR images contains no statistical data. There is no curve plot even for mass or tuberculosis but other parameters are evaluated.

Data processing: The breast MR dataset is processed using MATLAB through Intel inside CORE i3 processor with 64 GB RAM and Windows 10. Fig. 1 displayed the framework of the proposed tumor detection algorithm made up of three phases; Pre-processing, Breast region segmentation & tumor detection and feature extraction & classification. The detailed description of all the phases is given as:

2.2. Phase 1: Pre-processing

Pre-processing improves the quality of image & make the image quite appropriate for further processing by mankind or even computer vision systems. It helps in removing the artefacts in the background, and sharpening/smoothing the required region (Jaglan, Dass, & Duhan, 2019a). Initially, the acquired breast MR scanned images are read in DICOM format so converted into 8-bit grey level scale, thus each pixel of image has intensity value (0 to 255). Gray scale value is the range of shades of gray which is required to assign a specific intensity value per pixel. Hence, the MRI image is intended to pre-process in gray-scale level.

Table 2
Real Time Breast MR Image Dataset Description.

Case	Signal - Intensity Curve	Type	Tumor location	Stage	Tumor Area (mm^2)	Mean	Standard Deviation	Perimeter
1		Benign	Right	BIRAD-II	1035.61	1532.9	377.1	131.13
2		Benign	Left	BIRAD-III	2817.36	1807.2	420.5	210.16
3	_	Normal	_					
4		Benign	Left	BIRAD-IV	1680.15	1345.9	348.7	159.17
5		DCIS	Left	BIRAD-IV	124.92	765.2	522.7	44.38
6	_	Normal	_					
7	_	Solid mass	Right	_	488.56	928.9	626.6	84.68
8		Benign	Right	BIRAD-III	287.4	1383.8	255.6	65.46
9	_	Normal	_					
10	_	Tuberculosis	left	_	374.76	967.5	390.1	82.05
11		Benign	Right	BIRAD-III	156.58	2203.5	519	49.36
12	_	Normal	_					

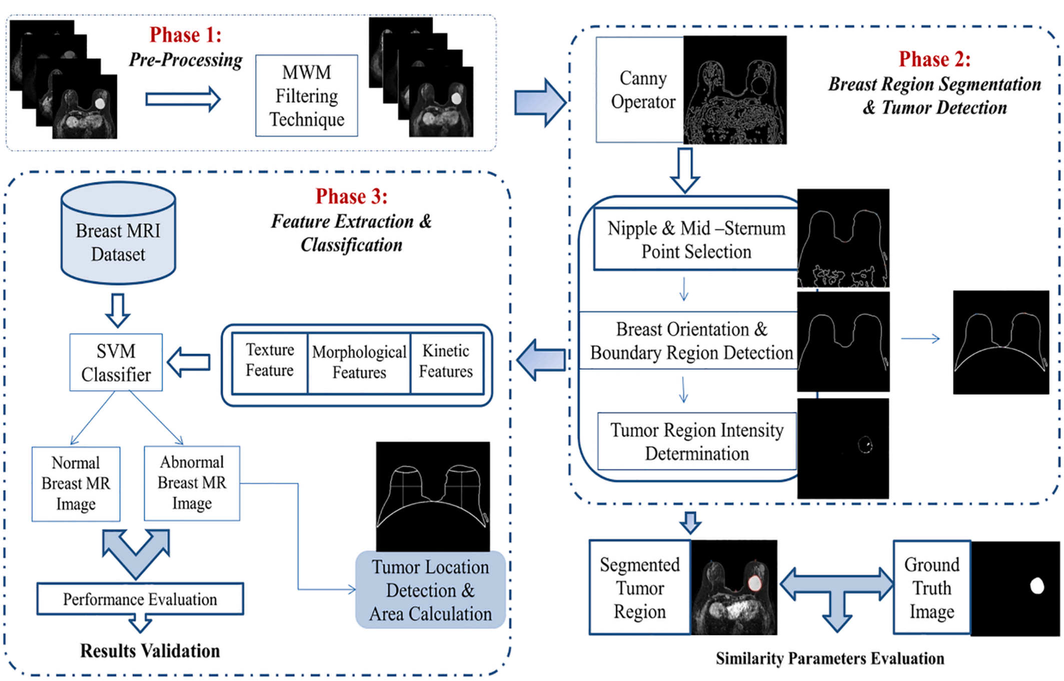


Fig. 1. Framework of Proposed Algorithm.

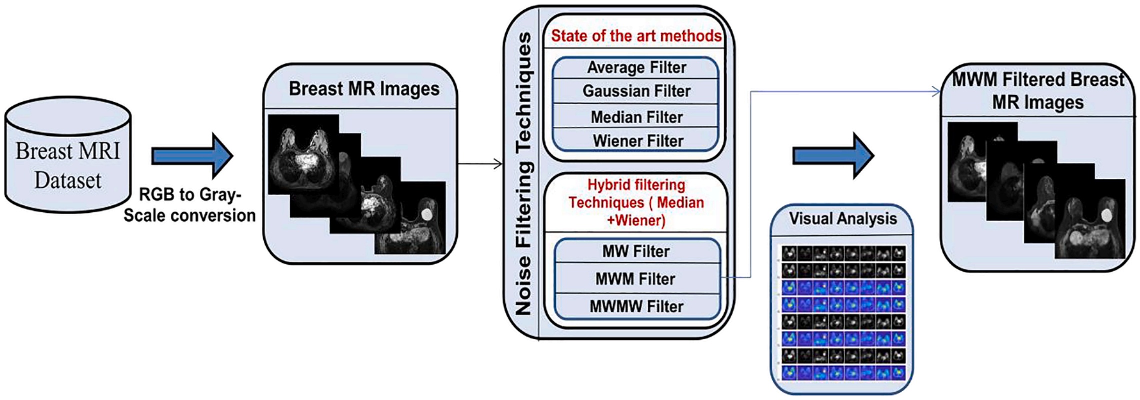


Fig. 2. Workflow of Noise removal Technique.

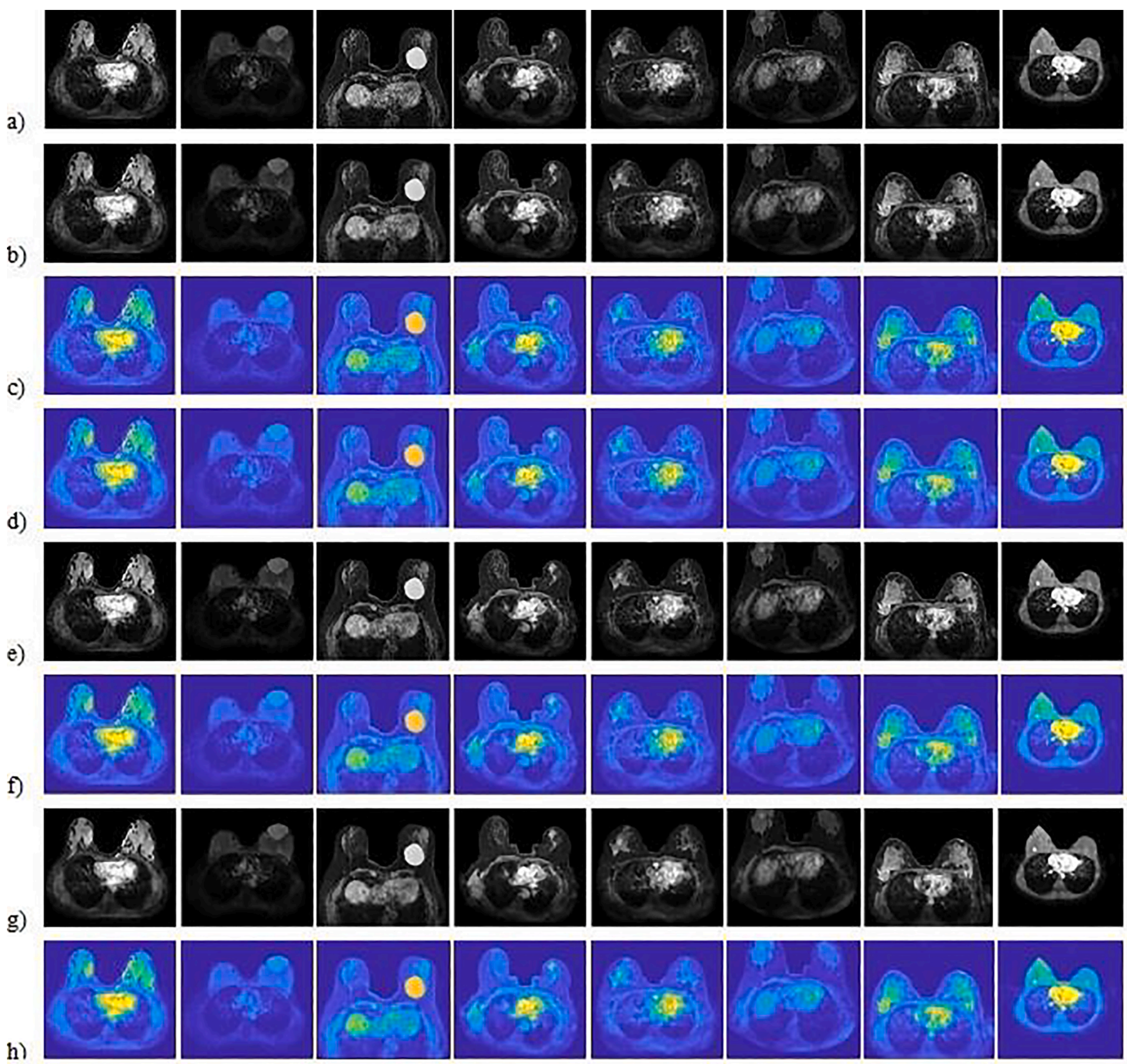


Fig. 3. Qualitative analysis of proposed hybridized filters with other existing filters a) Original breast MR images of eight patients [1–8(from left to right)] (b) Average filtered output (c) Gaussian filtered output (d) Wiener filtered output (e) Median filtered output (f) MW filtered images (g) MWM filtered images (h) MWMW filtered images.

Noise removal: From the literature, the suggested filter used to remove the impulse/Salt & pepper noise is Median filter & Wiener filter is best for removal of Gaussian noise (Kumar & Nachamai, 2017). Therefore, the combination of these two effective filters in a specific order may lead to an efficient hybrid filtering method as shown in Fig. 2.

In this phase, three integrated filter designed by cascading Median & Wiener Filter in a specific order i.e. MW, MWM, MWMW are used for noise removal and proves better than filters i.e. Average, Gaussian, Wiener and Median in terms of quantitative as well as qualitative analysis. The MWM filter is validated as the better one in terms of

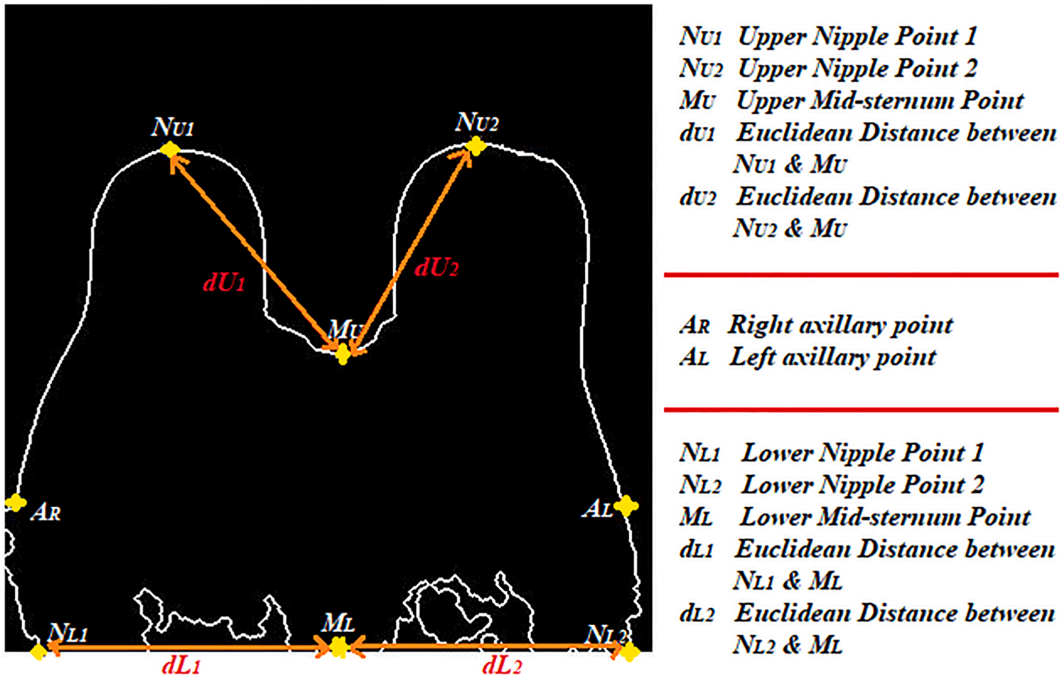


Fig. 4. Point selection criteria for exact breast orientation.

quantitative analysis done through various Performance evaluation parameters like MSE, RMSE, PSNR and MAE and even by the radiologists among all the other filtering techniques as well as the other proposed ones i.e. MW and MWMW as per their visual prospective depicted as qualitative analysis in Fig. 3.

2.3. Phase 2: Breast region segmentation & tumor detection

Filtered breast MR images are further processed through canny edge operator (Jaglan et al., 2019a) to get the sharp edges as it helps in detecting the tumor boundary. It gives very high rated results and used very widely for the object's finer details (Gobindchandra, 2015).

Algorithm 1 gives the detailed description of all the steps included:

Algorithm 1 Proposed Algorithm
Input: I_B (I is breast MR Image where B = 1, 2, 3. . . n)
Output: B_{TR} (Segmented Tumor Region)
procedure 1: Extract Breast Region and Tumor Region
Convert I_N → gray scale I_G.
Noise Filtering using 'MWM' $I_G \rightarrow I_{FG}$ section 2.2.
Edge Detection I_{FG} → 'Canny operator' I_ED section 2.3.
Selection of Nipple and Mid-sternum points N_{U1}, N_{U2}, N_{L1}, N_{L2} and M_U, M_L where N_U and M_U represent peak and valley point from upper side while N_L and M_L represents the same from lower side of the breast MR Image.
Max.(d_U , d_L), where d_U = distance (Euclidian, N_{U1}, N_{U2} and M_U) and d_L = distance (Euclidian, N_{L1}, N_{L2} and M_L) defines the orientation of the breast MR Image.
Breast Region I_{BR} → connecting corner points C_1, C_2 and $M_U \mid M_L$.
Extract Breast region MR Image $I_{BRO} \mid I_{BRO} \rightarrow I_{BR} * I_B$.
Extract Tumor region $B_{TR} \mid B_{TR} \subseteq max(SI_{BR})$.
I_{Tn} → Thinning I_{TR}
Tumor Part I_{Tp} → Largest Connected Component I_{Tn}
procedure 2: Extract Tumor Features Information
Compute Features (Texture, Morphological & Kinetic) → I_{Tp}.
Calculate μ, σ, Contrast, Homogeneity, Correlation, Energy, globalmean, smoothness, uniformity, entropy, skewness, correlation, area, peri, compact → I_{Tp}.
Extract the Region B_{TR} and Presence based on the SVM Classification using features information.
Calculate Tumor Area *Area*(B_{TR}) and Identify *Location*(B_{TR}).

Nipple & mid-sternum point selection: In this, the pixel by pixel scanning of whole image is done to get the peak & valley points for finding out the orientation of the image. The points are selected from both the lower as well as upper image side and then calculated the Euclidean distance between the upper peak points i.e. N_{U1} & N_{U2} and the valley point M_U as shown in Fig. 4. Similar process is done for the lower side to get the Euclidean distance between N_{L1} & N_{L2} and M_L. The Euclidean distance between points ***N & M*** is the connecting length of the line segment. If these N_{U1},N_{U2}, N_{L1}, N_{L2} and M_U, M_L are nipple & mid-sternum points from both side of the breast MRI in Euclidean space, then the distance (d_U) from N_{U1} to M_U or N_{U2} to M_U is compared with the distance (d_L) from $N_{L1}$to M_L or N_{L2} to M_L as shown in Fig. 4. After comparing both the Euclidean distance, the maximum one defines the exact orientation of the image. However, it also makes the process Rotation Invariant(Islam, Wijewickrema, & O'Leary, 2019)which means the system doesn't get affected by the augmentation of the image at any angle. It helps in determining the exact orientation of the breast which is further helpful in the tumor location identification. The right as well as left axillary points are denoted by A_R and A_L respectively.

Tumor intensity determination & segmentation: The next step included morphological operators for hole-filling and leakage removal; closing is needed for filling the holes on the breast boundary line; the inner holes within the breast are filled through hole filling; & the dilation caused by the closing operation is reduced, erosion is performed (Fusco, Di Marzo, Sansone, Sansone, & Petrillo, 2017). This is the final step to extract the tumor from the MR image clearly. Some morphological operators need to be applied to the binary mask obtained after the binarization of the MR image. This is done in order to make visible only that part of the region where the tumor is detected. That region is shown specifically in white colour and that is the part which is the region containing more intensity.

2.4. Phase 3: Feature extraction and classification

Once the lesion has been localized or segmented accurately, various properties are computed to characterize the lesion. The visual details of the breast MR image like shape, contrast, size, and texture are collected through feature extraction. The diagnosis system's accuracy can be improved by selecting leading features. Haralick, Shanmugam, and Dinstein (1973) presented a two step method for medical images' feature extraction in which firstly computed the GLCM and then the texture features are calculated on the basis of GLCM. Kinetic features are extensively investigated and used as a conventional method in the evaluation of the performance of various classifiers. More recently, morphological features are computed and evaluated in the lesion analysis (Fusco et al., 2016). There are 14 texture features i.e. energy, variance, entropy, maximal correlation coefficient, contrast, correlation, difference & sum variance, inverse difference moment, sum & difference entropy, information measurements correlation are extracted through GLCM. In this study, total 27 features are selected such as morphological features, the likelihood of abnormality, the standard deviation (SD) & mean of every kinetic parameter of whole abnormal tissues. After all lesion features have been extracted, various classifiers i.e. CNN (Convolutional Neural Network), LDA (Linear Discriminant Analysis) & SVM can be used to label each tissue as normal or abnormal(Kanchanamani & Perumal, 2016). In this work, the abnormal or normal tissues are categorized by the extracted features using SVM classifier(Nahid & Kong, 2017). SVM has the important attributes like robustness to outliers, high generalization ability and absence of local minima (Fooladivanda, Shokouhi, & Ahmadinejad, 2017). An optimally splitting hyper-plane found as the decision surface which maximizes the stretch between the closest data point by both the sides of the surface (Gal, Mehnert, Bradley, Kennedy, & Crozier, 2009). The classifier is trained by manual annotations (done by the radiologists) of both types of MR scans either normal or abnormal tissues. The advantage of SVM is that it ends with the global minima (Erickson, Korfiatis, Akkus, & Kline, 2019). Also SVMs are less prone to over-fitting and works faster than Artificial Neural Network(Übeyli, 2007). *Negi R., Mathew R.* reviewed different machine learning algorithms breast cancer detection & analysis and showed that the SVM gives an optimal execution in terms of distinctness and accuracy (Negi & Mathew, 2020).

Deep learning is taking command in all kinds of classification tasks in different fields but their applicability to the field of medical imaging is limited by the lack of large imaging datasets as it requires a massive annotated training dataset(Amit et al., 2017). Even though the employment of deep learning methods in the field of medical imaging is captivating but multiple issues need to be resolved before introducing it for radiological practice such as data standardization, limited datasets, uninterpretable black box model and privacy & legal issues(Lee et al., 2017). The prime advantage of the algorithms based on deep learning in contrast with other machine learning based algorithms is to find out the prime features during the search process so feature computation is not needed to be done as its first step(Giger, 2018; Wu et al., 2020).

2.5. Tumor location detection & area calculation

If the segmented region is classified as the abnormal one on the basis of extracted features, then the area of that particular lesion is calculated. The area of the abnormal tissues depends on the absolute sum of pixels lies in that region; therefore the area of a single pixel must be determined. In this, after the tumor region is detected, the maximum intensity pixel values can be obtained. By using this intensity value, the pixels can be converted into square centimetres. The equation (i) & (ii) are used to calculate the area of the abnormal lesion (Lakshmi, Swathi, & Srinivas, 2016) :

$$Area\ of\ Tumor = A \times total\ no\ of\ tumor\ region's\ pixels \tag{1}$$

Therefore,

$$A = V \times H \tag{2}$$

where, A gives the value of single pixel's area; V and H represents the vertical and horizontal dimension of the breast MRI respectively. The unit of area is taken as cm^2.

Next, the tumor location is identified as per the primary location guidelines designed specifically for the axial orientation of breast MR images given below. In literature, only coronal plane is available to locate the tumor (Rummel, Hueman, Costantino, Shriver, & Ellsworth, 2015) but a new perspective is designed in this work to identify the location of tumor in the axial view with the consent of the experts. These guidelines are specifically designed for the axial view of the breast MRI.

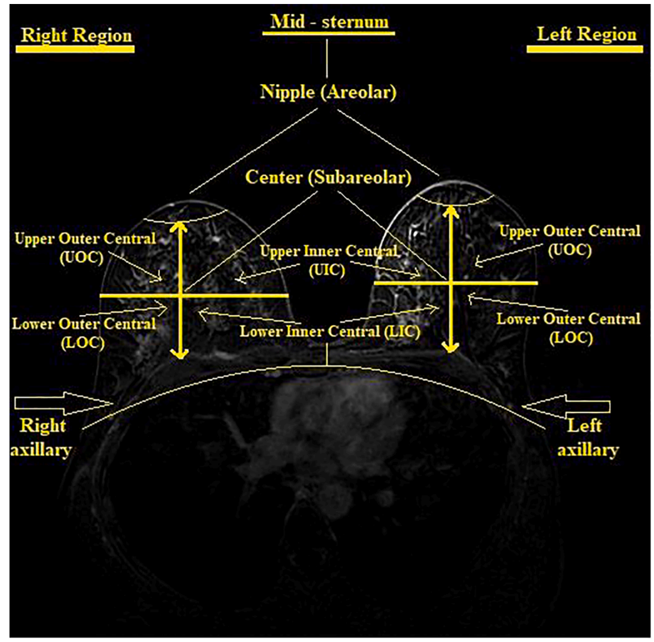


Fig. 5. Pictorial representation of Primary Location Guidelines.

Primary location guidelines for axial plane of breast MRI

- Right Breast Region
- Left Breast Region

A) Nipple (areolar)
i. Left areolar
ii. Right areolar
B) Center location (subareolar) area extending 1 cm around areolar (Left/ Right) complex
i. Upper Inner Central (UIC)
ii. Lower Inner Central (LIC)
iii. Upper Outer Central (UOC)
iv. Lower Outer Central (LOC)
C) Axillary tail of breast
i. Left Axillary
ii. Right Axillary
D) Overlapping lesion of breast
i. Inner Breast
ii. Outer Breast
iii. Upper Breast
iv. Lower breast
v. Medial breast
vi. Midline breast
E) Multiple legion of Breast
i. Left Breast
ii. Right Breast
iii. Entire Breast
iv. Multiple tumors at various sub-locations within breast
v. Third-forth or more of the breast region involved with tumor

Fig. 5 gives the pictorial representation of the primary location guidelines. The whole breast region is initially divided into two regions; left or right by identifying the mid-sternum point in between. Then the nipple/areolar points of both the breast regions are determined at the peak points. Similarly, center/subareolar point is defined by the center of each breast. Each breast is divided into four different central points like Upper Inner Central (UIC), Upper Outer Central (UOC), Lower Inner Central (LIC) and Lower Outer Central (LOC). Lastly, the left /right axillary points are determined which genuinely helps in finding the exact tumor location specifically for the axial view of breast MRI.

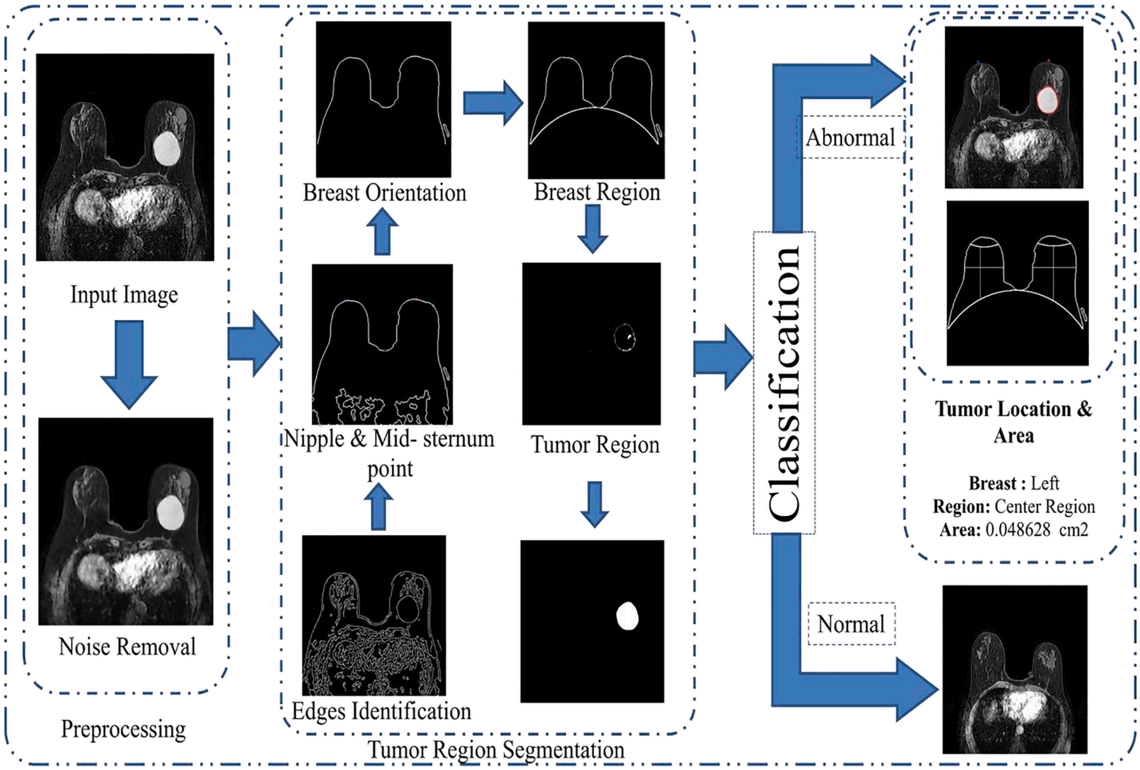


Fig. 6. Workflow adoption of proposed algorithm.

Table 3
Training and testing set.

Breast MR Image Dataset (448 Images)	Partition (%)	Training Set	Testing Set
Abnormal(318 Images)	50–50	158	160
Normal (130 Images)	70–30	83	47

3. Experimental results & discussion

In Fig. 6, the workflow adoption of the proposed algorithm is depicted. Initially, the proposed MWM filter is used to de-noise the input breast MR image. In the next step, the canny operator is used to sharp the edge so the nipple & mid-sternum points are easily selected and determine the exact breast orientation on the basis of Euclidean distance. The maximum Euclidean distance between these points determines the exact orientation. The point selection is done from both upward & downward directions of the breast MR image to make the image rotation invariant which helps in tumor location identification.

The tissues i.e. fatty, fibro-glandular and tumor are extracted from the breast MRI by determining the intensity level of the MR image and then holes' filling is done. The region with maximum intensity value is determined by defining the region properties of the breast MR image. Again, the abnormal tissues are identified from the segmented region through morphological operation. Next, the segmented lesion is classified as abnormal tissue or normal one by SVM on the basis of various features of the manual annotations by which the system is trained. If the extracted lesion is abnormal then the tumor location as well as area of that tissue must be evaluated.

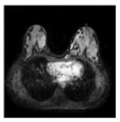
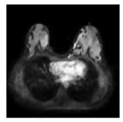
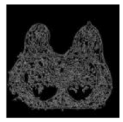
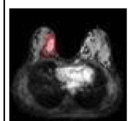
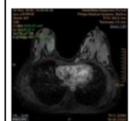
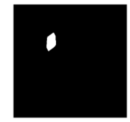
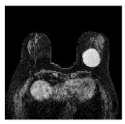
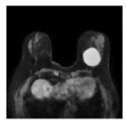
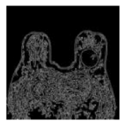
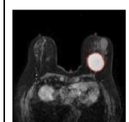
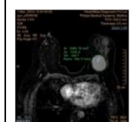
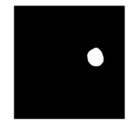
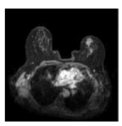
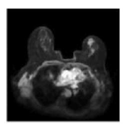
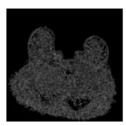
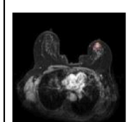
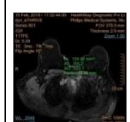
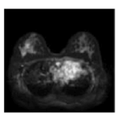
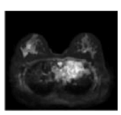
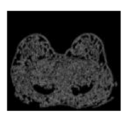
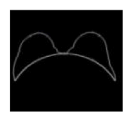
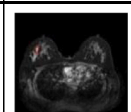
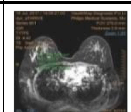
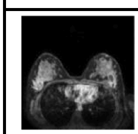
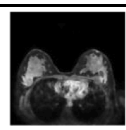
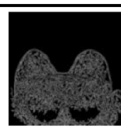
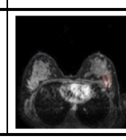
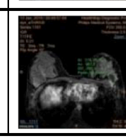
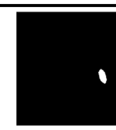
Table 3 gives the detailed description about the training and testing set used by the SVM classifier. Using this system, the statistically analysis of the tumor area and probabilistic classification is successfully done from the MRI images of twelve individual patients.

Out of all the breast MR images with suspicious lesion in the dataset, we have found 318 images which have well-defined as well as visible boundaries of breast outer region in axial view. These images were manually segmented by experts/radiologists to generate the ground truth images for reference. Rest of the images contains many thin structures or patterns which were difficult to segment manually even by radiologists. Therefore, it was difficult to validate the results with manual segmentations. Every 10 consecutive images are allocated with one breast region label due to similarity in consecutive images in MRI sequences. It employed patient-level, 10-fold cross validation so each subset contains approximately 31 MR images. The system is trained with five subsets and tested on remaining subsets to illustrate the robustness of the algorithm. Table 4 shows the average of performance parameters i.e. accuracy, sensitivity, precision, F-measure, Dice & Jaccard index (Sokolova & Lapalme, 2019) of the predicted image generated through proposed algorithm for MR image dataset of breast against the manual segmented image obtained through mean value of the five testing

Table 4
Performance analysis of the predicted images against ground Truth images of abnormal tissues.

S. No.	Statistical Measures	Testing Subset 1	Testing Subset 2	Testing Subset 3	Testing Subset 4	Testing Subset 5	Mean ± Standard deviation
1	Accuracy	0.998671	0.99629	0.999466	0.999885	0.999537	0.9987 ± 0.0014
2	Sensitivity	0.979766	0.7316	1	0.97235	0.95295	0.9273 ± 0.1107
3	Precision	0.998792	1	0.999465	0.999925	0.999925	0.9996 ± 0.0005
4	F-measure	0.838651	1	0.609211	0.95045	0.977029	0.8750 ± 0.1610
5	Dice	0.903733	0.844999	0.757155	0.961276	0.964839	0.8864 ± 0.0872
6	Jaccard	0.989233	0.855336	0.855336	0.986041	0.976118	0.9324 ± 0.0705

Table 5
Result Analysis.

Original Breast MRI	Pre-processing	Edge Detection	Breast Region Segmentation	Tumor detection	Ground truth Image	Predicted Image	Outcomes
							Breast: Right Region: Center Area of Tumor: 0.067541 cm^2
							Breast: Left Region: Center Area of Tumor: 0.048834 cm^2
							Breast: Left Region: Nipple Area of Tumor: 0.004624 cm^2
							Breast: Right Region: Upper Outer Central Area of Tumor: 0.003159 cm^2
							Breast: Left Region: Upper Outer Central Area of Tumor: 0.013807 cm^2

N= 207	Predicted (Normal)	Predicted (Abnormal)	
Actual (Normal)	TN (41)	FP (6)	47
Actual (Abnormal)	FN (7)	TP (153)	160
	48	159	

Fig. 7. Confusion Matrix of classification outcomes.

subsets.

Experts delineated manual segmentation for abnormal tissues which may used as the reference region so called ground truth image against the predicted one generated through the proposed algorithm. Table 5 gives the qualitative analysis of five breast MR image from each testing subset of abnormal cases of breast MRI with tumor of specific size, intensity and shape. This analysis proves the fact that proposed algorithm has accurately extracted the tumor. Also the tumor location and area is exactly determined.

Table 6
Performance Evaluation Parameters by SVM classifier.

S. No.	Performance Evaluation parameters	Formulae/ Calculation	Value (%)
1	Sensitivity	$\frac{TP}{Actual(Abnormal)}$	95.62
2	Specificity	$\frac{TN}{TN+FP}$	87.23
3	Accuracy	$\frac{TP+TN}{Total}$	93.71
4	Misclassification Rate	$\frac{FP+FN}{total}$	6.28
5	FPR	$\frac{FP}{Actual(Normal)}$	12.76
6	F Measure	$2\times\frac{Precision\times Recall}{Precision+Recall}$	86.44
7	Prevalence	$\frac{Actual(Abnormal)}{Total}$	77.29
8	Jaccard Index	$\frac{TP}{TP+FN+FP}$	92.16

4. Classification analysis

The performance evaluation of the work is done through metrics i.e. sensitivity, specificity, accuracy, precision etc. (Dass & Sanjeet, 2013) measured by the equations given below in Table 4. Total 448 images (318 abnormal and 130 normal cases) of breast MR dataset are used for performance analysis of SVM classifier. The system is trained with 241 images in total of both cases while tested the remaining 207 breast MR images. The evaluations are done through various parameters (Jalalian et al., 2017) shown in Fig. 7.

Where TP represents abnormal tissues correctly classified as the abnormal one; FP represents normal images which are wrongly classified as abnormal image; TN represents normal tissues which are correctly unclassified as abnormal image, FN shows abnormal images which are wrongly unclassified as abnormal tissues (Polat & Güneş, 2007). Table 6 gives the other performance evaluation parameters i.e.

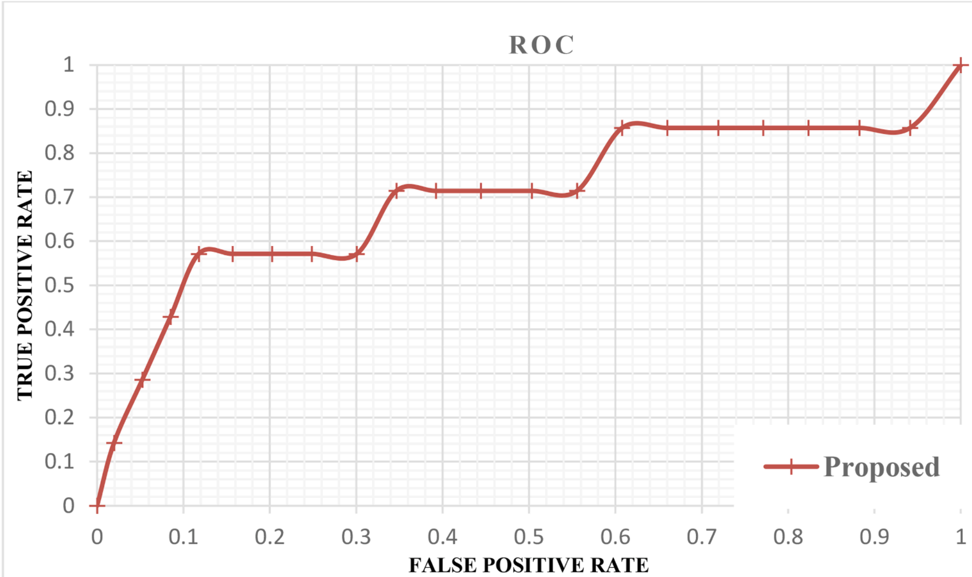


Fig. 8. ROC curve for all tested images of breast MR dataset.

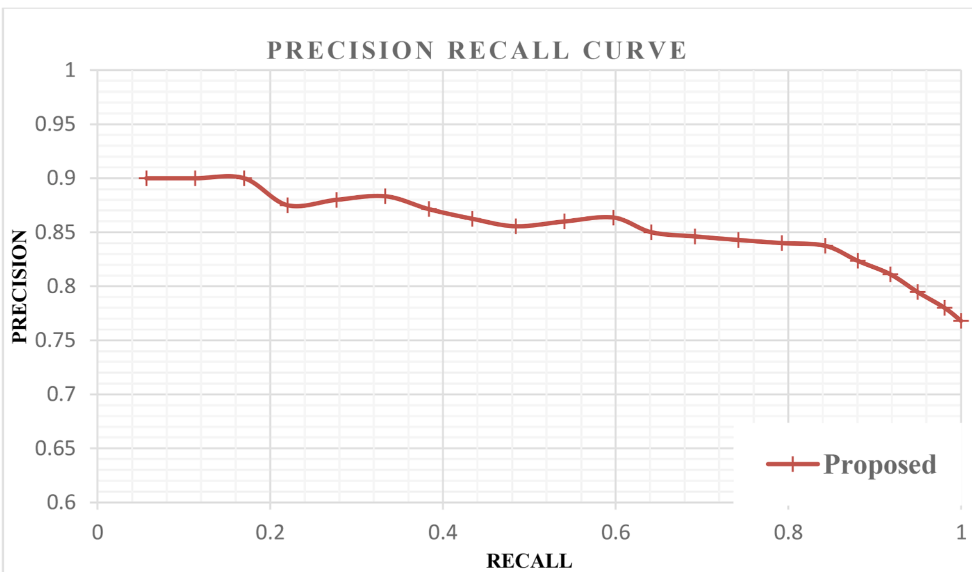


Fig. 9. Precision-Recall curve.

Accuracy, Sensitivity, Specificity etc.

ROC (Receiver Operating Curve) is generated in the two-dimensional plane: Sensitivity known as True Positive Rate (TPR) plotted on Y-axis and False Positive Rate (FPR) lies at X-axis [46]. ROC is considered as a benchmark to analyse the efficacy of classifier. The proposed algorithm shows the value of sensitivity is 95.6% means detection of abnormal tissues is good enough while the specificity is reached at 87.2%. The system attains good accuracy of 93.7% proves its significance in classifying normal or abnormal tissues. ROC curve is shown in Fig. 8 whereas Fig. 9 depicted the Precision-Recall curve.

RIDER breast MR dataset is one among the online available databases of Breast MRI with limited patients' data. The existing methods presented the statistical results of privately collected breast MRI databases from different hospitals. Therefore, comparative analysis between the approaches/methods implemented on different databases is not justified. Hence, the best way to validate the results is to demonstrate different evaluation metrics when compared with manually segmented results by the radiologist. This requires the active involvement of many radiologists to establish a large & even standardized medical imaging database. The research community also needs to share the benchmark datasets to facilitate this revolution.

5. Conclusion

In this study, the breast tissues are segmented and then classified into normal and abnormal tissues through MR image dataset of breast. The pre-processing is done to get a noise-free breast MR image which is quite appropriate to be used for edge detection & segmentation. The intensity of abnormal as well as normal tissues is determined; therefore the abnormal lesion is identified from the segmented region using morphological operations and holes' filling. The SVM classifier is used to classify the breast tissues by analyzing selected feature vectors. The implementation of the proposed algorithm is analyzed through the 207 breast MR images collected from the Healthmap Diagnostic centre, PGIMS, Rohtak. The experimental analysis performed on the different cases shows that the given algorithm is accurate enough when compared with the manual segmentation done by radiologists. Result analysis shows that the proposed one can aid in the early and accurate detection of abnormal tissues along with the identification of exact tumor location. The other evaluation parameters also prove the fact that the proposed algorithm provides improvements in certain parameters such as sensitivity, specificity, accuracy, Jaccard Index etc. The overlap coefficient termed as Jaccard Index is 92.1% and 93.7% accuracy demonstrated the efficacy of the system to classify the breast tissues into normal or abnormal one. The proposed work can be used like initial screening for the radiologists or may be taken as a second opinion in highly critical cases. The results of this study give a basic foundation for optimizing this work by analyzing more value added criterion of feature extraction & selection. Also, other classifiers along with different evolutionary or optimization algorithms can be designed in future.

Funding

This research is funded by the UGC, New Delhi under the Junior Research Fellowship Scheme (UGC-NET/JRF).

Ethical Statement

Due to the retrospective nature of this study, the Institutional Review Board has exempted the ethics approval.

CRediT authorship contribution statement

Poonam Jaglan: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing. **Rajeshwar Dass:** Conceptualization, Project administration, Resources, Writing - original draft, Writing - review & editing. **Manoj Duhan:** Conceptualization, Project administration, Resources, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the MRI Unit of Healthmap diagnostic centre, PGIMS, Rohtak for providing us with breast MRI scanning images dataset. Also thankful to Dr. Vinay Malik & Dr. Ajit for patient referrals and manual segmentation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eswa.2021.115580.

References

Abonyi, J., & Szeifert, F. (2003). Supervised fuzzy clustering for the identification of fuzzy classifiers. *Pattern Recognition Letters, 24*(14), 2195–2207. https://doi.org/10.1016/S0167-8655(03)00047-3

Al-Faris, A. Q., Kalthum Ngah, U., Ashidi Mat Isa, N., & Lutfi Shuaib, I. (2012). MRI breast skin-line segmentation and removal using integration method of level set active contour and morphological thinning algorithms. *Journal of Medical Sciences (Faisalabad), 12*(8), 286–291. https://doi.org/10.3923/jms.2012.286.291

Al-faris, A. Q., Ngah, U. K., Ashidi, N., Isa, M., & Lutfi, I. (2012). Breast MRI Tumour Segmentation using Modified Automatic Seeded Region Growing Based on Particle Swarm Optimization Image Clustering. *Online Coference on Soft Computing in Industrial Applications, 2012*, 1–11.

Al-faris, A. Q., Ngah, U. K., Isa, N. A. M., & Shuaib, I. L. (2015). Automatic Exclusion of Skin Border Regions from Breast MRI Using Proposed Combined Approach Automatic Exclusion of Skin Border Regions from Breast MRI Using Proposed Combined Approach. 2nd International Conference on Biomedical Engineering (ICoBE), (September). https://doi.org/10.1109/ICoBE.2015.7235873.

Amit, G., Ben-Ari, R., Hadad, O., Monovich, E., Granot, N., & Hashoul, S. (2017). Classification of breast MRI lesions using small-size training sets: Comparison of deep learning approaches. *Medical Imaging 2017: Computer-Aided Diagnosis*. *Proc. of SPIE, 10134*, 101341H–101345H. https://doi.org/10.1117/12.2249981

Bouchebbah, F., & Slimani, H. (2019). Levels Propagation Approach to Image Segmentation: Application to Breast MR Images. *Journal of Digital Imaging, 32*(3), 433–449. https://doi.org/10.1007/s10278-018-00171-2

Cai, H., Liu, L., Peng, Y., Wu, Y., & Li, L. (2014). Diagnostic assessment by dynamic contrast-enhanced and diffusion-weighted magnetic resonance in differentiation of breast lesions under different imaging protocols. *BMC Cancer, 14*(1), 1–12. https://doi.org/10.1186/1471-2407-14-366

Chen, H. L., Yang, B., Liu, J., & Liu, D. Y. (2011). A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis. *Expert Systems with Applications, 38*(7), 9014–9022. https://doi.org/10.1016/j.eswa.2011.01.120

Cheng, H. D., Shan, J., Ju, W., Guo, Y., & Zhang, L. (2010). Automated breast cancer detection and classification using ultrasound images: A survey. *Pattern Recognition, 43*(1), 299–317. https://doi.org/10.1016/j.patcog.2009.05.012

Cui, Yunfeng, Tan, Yongqiang, Zhao, Binsheng, Liberman, Laura, Parbhu, Rakesh, Kaplan, Jennifer, … Schwartz, Lawrence H. (2009). Malignant lesion segmentation in contrast-enhanced breast MR images based on the marker-controlled watershed. *Medical Physics, 36*(10), 4359–4369. https://doi.org/10.1118/1.3213514

Dass, R., & Sanjeet. (2013). Effect of Feedforward Back Propagation Neural Network for Breast Tumor Classification. *International Journal of Computer Science and Technology, 4*(2), 731–735.

Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, Timothy L. (2019). Machine Learning for Medical Imaging. *Radio Graphics, 37*(2), 505–515. https://doi.org/10.1155/2019/9874591

Fooladivanda, A., Shokouhi, S. B., & Ahmadinejad, N. (2017). Breast-region segmentation in MRI using chest region atlas and SVM. *Turkish Journal of Electrical Engineering & Computer Sciences, 25*, 4575–4592. https://doi.org/10.3906/elk-1512-40

Fooladivanda, Aida, Shokouhi, Shahriar B., & Ahmadinejad, Nasrin (2017). Localized-atlas-based segmentation of breast MRI in a decision-making framework. *Australasian Physical and Engineering Sciences in Medicine, 40*(1), 69–84. https://doi.org/10.1007/s13246-016-0513-3

Fusco, R., Di Marzo, M., Sansone, C., Sansone, M., & Petrillo, A. (2017). Breast DCE-MRI: Lesion classification using dynamic and morphological features by means of a multiple classifier system. *European Radiology Experimental, 1*(10), 1–7. https://doi.org/10.1186/s41747-017-0007-4

Fusco, R., Sansone, M., Filice, S., Carone, G., Amato, D. M., Sansone, C., & Petrillo, A. (2016). Pattern Recognition Approaches for Breast Cancer DCE-MRI Classification: A Systematic Review. *Journal of Medical and Biological Engineering, 36*(4), 449–459. https://doi.org/10.1007/s40846-016-0163-7

Gal, Y., Mehnert, A., Bradley, A., Kennedy, D., & Crozier, S. (2009). Feature and classifier selection for automatic classification of lesions in dynamic contrast-enhanced MRI of the breast. DICTA 2009 - Digital Image Computing: Techniques and Applications, 132–139. https://doi.org/10.1109/DICTA.2009.29.

Giger, M. L. (2018). Machine Learning in Medical Imaging. *Journal of the American College of Radiology, 15*(3), 512–520. https://doi.org/10.1016/j.jacr.2017.12.028

Gobindchandra, K., & K.L, S. K. (2015). Analysis of Image Segmentation Techniques. International Research Journal of Computer Science (IRJCS), 2(6), 45–53. https://doi.org/10.24128/ijraer.2017.mn89cd.

Haralick, Robert M., Shanmugam, K., & Dinstein, Its'Hak (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics, SMC-3*(6), 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Huang, C. (2014). Breast Mass Segmentation on Breast MRI using the Shape-based Level Set Method. *Biomedical Engineering: Applications, Basis and Communications, 26*(4), 1–11. https://doi.org/10.4015/S1016237214400067

Islam, Kh Tohidul, Wijewickrema, Sudanthi, & O'Leary, Stephen (2019). A rotation and translation invariant method for 3D organ image classification using deep convolutional neural networks. *PeerJ Computer Science, 5*, e181. https://doi.org/10.7717/peerj-cs.18110.7717/peerjcs.181/fig-110.7717/peerjcs.181/fig-210.7717/peerjcs.181/fig-310.7717/peerjcs.181/fig-410.7717/peerjcs.181/fig-510.7717/peerjcs.181/fig-610.7717/peerjcs.181/fig-710.7717/peerjcs.181/fig-810.7717/peerjcs.181/table-110.7717/peerjcs.181/table-210.7717/peerjcs.181/table-310.7717/peerjcs.181/table-410.7717/peerjcs.181/table-510.7717/peerj-cs.181/supp-1

Jaglan, P., Dass, R., & Duhan, M. (2019a). A Comparative Analysis of Various Image Segmentation Techniques. Proceedings of 2nd International Conference on Communication, Computing and Networking, Lecture Notes in Networks and Systems, 359–374. https://doi.org/10.1007/978-981-13-1217-5.

Jaglan, Poonam, Dass, Rajeshwar, & Duhan, Manoj (2019). Breast Cancer Detection Techniques: Issues and Challenges. *Journal of The Institution of Engineers (India): Series B Electrical. Electronics & Telecommunication and Computer Engineering, Springer, India, 100*(4), 379–386. https://doi.org/10.1007/s40031-019-00391-2

Jaglan, P., Dass, R., & Duhan, M. (2019b). Detection of Breast Cancer using MRI: A Pictorial Essay of the Image Processing Techniques. *International Journal of Computer Engineering in Research Trends, 6*(1), 238–245. https://doi.org/10.22362/ijcert/2019/v6/i01/v6i0101.

Jalalian, A., Mashohor, S., Mahmud, R., Karasfi, B., Ramli, A. R. B., Engineering, C. S., … Branch, Q. (2017). Foundation and Methodologies in Computer-Aided Diagnosis system for Breast Cancer Detection. *EXCLI Journal, 16*, 113–137.

Kanchanamani, M., & Perumal, V. (2016). Performance evaluation and comparative analysis of various machine learning techniques for diagnosis of breast cancer. *Biomedical Research (India), 27*(3), 623–631.

Karabatak, Murat, & Ince, M. Cevdet (2009). An expert system for detection of breast cancer based on association rules and neural network. *Expert Systems with Applications, 36*(2), 3465–3469. https://doi.org/10.1016/j.eswa.2008.02.064

Kele, A., Kele, A., & Yavuz, U. (2011). Expert system based on neuro-fuzzy rules for diagnosis breast cancer. *Expert Systems with Applications, 38*(5), 5719–5726. https://doi.org/10.1016/j.eswa.2010.10.061

Kumar, Nalin, & Nachamai, M (2017). Noise Removal and Filtering Techniques used in Medical Images. *ORIENTAL JOURNAL OF COMPUTER SCIENCE & TECHNOLOGY, 10* (1), 103–113.

Lakshmi, T. M., Swathi, R., & Srinivas, A. (2016). MATLAB Implementation of an Efficient Technique for Detection of Brain Tumor by using Watershed Segmentation and Morphological Operation. *GRD Journals- Global Research and Development Journal for Engineering, 1*(4), 80–87.

Lee, J.-G., Jun, S., Cho, Y.-W., Lee, H., Kim, G. B., Seo, J. B., & Kim, Namkug (2017). Deep learning in medical imaging. *Korean Journal of Radiology, 18*(4), 570–584. https://doi.org/10.14245/ns.1938396.198.

Losurdo, L., Basile, T. M. A., Fanizzi, A., Bellotti, R., Bottigli, U., Carbonara, R., … La Forgia, D. (2018). A Gradient-Based Approach for Breast DCE-MRI Analysis. BioMed Research International, Vol. 2018, Article ID 9032408, 10 Pages. https://doi.org/10.1155/2018/9032408.

Maglogiannis, I., Zafiropoulos, E., & Anagnostopoulos, I. (2009). An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers. *Applied Intelligence, 30*, 24–36. https://doi.org/10.1007/s10489-007-0073-z

Marcano-Cedeño, A., Quintanilla-Domínguez, J., & Andina, D. (2011). WBCD breast cancer database classification applying artificial metaplasticity neural network. *Expert Systems with Applications, 38*(8), 9573–9579. https://doi.org/10.1016/j.eswa.2011.01.167

Moftah, Hossam M., Azar, Ahmad Taher, Al-Shammari, Eiman Tamah, Ghali, Neveen I., Hassanien, Aboul Ella, & Shoman, Mahmoud (2014). Adaptive k-means clustering algorithm for MR breast image segmentation. *Neural Computing and Applications, 24* (7-8), 1917–1928. https://doi.org/10.1007/s00521-013-1437-4

Mu, Tingting, & Nandi, Asoke K. (2007). Breast cancer detection from FNA using SVM with different parameter tuning systems and SOM-RBF classifier. *Journal of the Franklin Institute, 344*(3-4), 285–311. https://doi.org/10.1016/j.jfranklin.2006.09.005

Muthu Rama Krishnan, M., Banerjee, Shuvo, Chakraborty, Chinmay, Chakraborty, Chandan, & Ray, Ajoy K. (2010). Statistical analysis of mammographic features and its classification using support vector machine. *Expert Systems with Applications, 37*(1), 470–478. https://doi.org/10.1016/j.eswa.2009.05.045

Nahid, A. Al, & Kong, Y. (2017). Involvement of Machine Learning for Breast Cancer Image Classification: A Survey. Computational and Mathematical Methods in Medicine, 2017(Article ID 3781951), 29 pages. https://doi.org/10.1155/2017/3781951.

Negi, R., & Mathew, R. (2020). In: Pandian A., Senjyu T., Islam S., Wang H. (eds) Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2018). ICCBI 2018. Lecture Notes on Data Engineering and Communications Technologies (Vol. 31). https://doi.org/10.1007/978-3-030-24643-3.

Nie, K., Chen, J. H., Yu, H. J., Chu, Y., Nalcioglu, O., & Su, M. Y. (2008). Quantitative Analysis of Lesion Morphology and Texture Features for Diagnostic Prediction in Breast MRI. *Academic Radiology, 15*(12), 1513–1525. https://doi.org/10.1016/j.acra.2008.06.005

Polat, K., & Güneş, S. (2007). Breast cancer diagnosis using least square support vector machine. *Digital Signal Processing: A Review Journal, 17*(4), 694–701. https://doi.org/10.1016/j.dsp.2006.10.008

Rummel, S., Hueman, M. T., Costantino, N., Shriver, C. D., & Ellsworth, R. E. (2015). Tumour location within the breast: Does tumour site have prognostic ability? *Ecancermedicalscience, 9*, 1–10. https://doi.org/10.3332/ecancer.2015.552

Saadatmand, Sepideh, Geuzinge, H Amarens, Rutgers, Emiel J T, Mann, Ritse M, de Roy van Zuidewijn, Diderick B W, Zonderland, Harmien M, … van Druten, Edith (2019). MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): A multicentre, randomised, controlled trial. *The Lancet Oncology, 20*(8), 1136–1147. https://doi.org/10.1016/S1470-2045(19)30275-X

Şahan, S., Polat, K., Kodaz, H., & Güneş, S. (2007). A new hybrid method based on fuzzy-artificial immune system and k-nn algorithm for breast cancer diagnosis. *Computers in Biology and Medicine, 37*(3), 415–423. https://doi.org/10.1016/j.compbiomed.2006.05.003

Schoub, P. K. (2018). Understanding indications and defining guidelines for breast magnetic resonance imaging. *South African Journal of Radiology, 22*(2), 1–12. https://doi.org/10.4102/sajr.v22i2.1353

Sinha, Shantanu, Lucas-Quesada, Flora Anne, Debruhl, Nanette D., Sayre, James, Farria, Dionne, Gorczyca, David P., & Bassett, Lawrence W. (1997). Multifeature analysis of Gd-enhanced MR images of breast lesions. *Journal of Magnetic Resonance Imaging, 7*(6), 1016–1026. https://doi.org/10.1002/(ISSN)1522-258610.1002/jmri.v7:610.1002/jmri.1880070613

Sokolova, M., & Lapalme, G. (2019). A systematic analysis of performance measures for classification tasks. *Information Processing and Management, 45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Stoean, R., & Stoean, C. (2013). Modeling medical decision making by support vector machines, explaining by rules of evolutionary algorithms with feature selection. *Expert Systems with Applications, 40*(7), 2677–2686. https://doi.org/10.1016/j.eswa.2012.11.007

Übeyli, E. D. (2007). Implementing automated diagnostic systems for breast cancer detection. *Expert Systems with Applications, 33*(4), 1054–1062. https://doi.org/10.1016/j.eswa.2006.08.005

Warner, E., Plewes, D. B., Shumak, R. S., Catzavelos, G. C., Di Prospero, L. S., Yaffe, M. J., … Narod, S. A. (2001). Comparison of breast magnetic resonance imaging, mammography, and ultrasound for surveillance of women at high risk for hereditary breast cancer. *Journal of Clinical Oncology, 19*(15), 3524–3531. https://doi.org/10.1200/JCO.2001.19.15.3524

Wu, Nan, Phang, Jason, Park, Jungkyu, Shen, Yiqiu, Huang, Zhe, Zorin, Masha, … Geras, Krzysztof J. (2020). Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. *IEEE Transactions on Medical Imaging, 39*(4), 1184–1194. https://doi.org/10.1109/TMI.4210.1109/TMI.2019.2945514