Performance Evaluation of Improved Awareness Probability-based Crow Search Algorithm for Breast Cancer Detection

Rajeshwari S. Patil[1], Ambaji S. Jadhav[2] and Nagashettappa Biradar[3]

Department of Electronics and Communication Engineering

Dr. P.G.H College of Engg. & Tech. Vijayapur, India[1, 2], B K I T, Bhalki, India[3]

E-mail:- *rspatil1272014@gmail.com[1], ec.ambaji@bldeacet.ac.in[2], nmbiradar@gmail.com[3]*

DOI: - https://doi.org/10.47531/MANTECH/ECC.2021.34

Abstract

Breast cancer is a major disease, which is usually seen in women. Early researches have shown that early detection and suitable treatment might increase the life span. Those researches have also proven that detection of small lesions at an early stage improves prognosis, which leads to a decrease in the mortality rate. Mammography is the best approach used for screening the disease. The present paper plans to introduce an automatic breast cancer detection approach using four phases as "pre-processing, segmentation, feature extraction, and classification." Here, the median filtering approach is used for eliminating the noise present in the mammogram image. Later, the segmentation of the tumor is done by the optimized region growing approach, which is the advanced version of the traditional region growing algorithm. Furthermore, the features like "Grey Level Co-occurrence Matrix (GLCM)" and Gray-Level "Run-Length Matrix (GRLM)" are extracted from the segmented tumor during feature extraction. Once the feature extraction is done, the features are subjected to a classifier named Fuzzy logic classifier. The threshold of the region growing algorithm and the membership function of the fuzzy classifier is optimally tuned with the help of the Crow Search Algorithm (CSA) named as "Improved Awareness Probability-based CSA" (IAP-CSA). The analysis shows that the proposed IAP-CSA is acquiring the best results in breast cancer detection and classifying the normal, benign, and malignant images.

Keywords: *- Mammogram Image, Breast Cancer Detection, Optimized Region Growing Algorithm, Optimized Fuzzy Classifier, Improved Awareness Probability-based Crow Search Algorithm (IAP-CSA).*

Nomenclature

Abbreviations	**Descriptions**
GLCM	Grey Level Co-occurrence Matrix
GLRM	Gray-Level Run-Length Matrix
CSA	Crow Search Algorithm
IAP-CSA	Improved Awareness Probability-based CSA
FPR	False Positive Rate
DCT	Discrete Cosine Transform
MCC	Matthew's Correlation Coefficient
DST	Discrete Shearlet Transform
LWT	Lifting Wavelet Transform
FDR	False Discover Rate
SURF	Speed-Up Robust Features
MSVM	Multiclass SVM
FNR	False Negative Rate
DNN	Deep Neural Network
CAD	Computer Aided Diagnosis
NPV	Negative Predictive Valu
LDA	Linear Discriminant Analysis
PCA	Principal Component Analysis
ELM	Extreme Learning Machine

MFO-ELM	Moth Flame Optimization approach-ELM
SVM	Support Vector Machine
CNN	Convolutional Neural Network
GOA	Grasshopper Optimization Algorithm
FCM	Fuzzy C-means Clustering
HGRE	“High Grey Level Run Emphasis”
LGRE	“Low Grey Level Run Emphasis”

INTRODUCTION

Tumor is one form of the disease and the significant features of breast cancer disease that lead to uncontrolled cell growth in the specific region of the body. This type of growth of the cell is termed as tumor. When breast tissue produces disease, breast malignancy is formed [6], which is the major problem suffered by the current medical field. The early diagnosis of tumor might increase the recuperation rate. For early detection, identifying the position, and for suitable treatment of breast cancer, mammography is employed. Mammography has the ability to recognize the tumor cells, which are small and incredibly complex to detect, and it is the best technique used for breast cancer detection. In mammography, breast imaging is primarily done with the help of Xbeam’s low measurement by having more targets and huge divergence [7] [8] [9]. Moreover, this mammography technique is employed for both diagnosing and screening breast cancer.

The screening of mammograms manually by the radiographer is expensive, complex, and consumes more time, which leads to huge FPR. However, the variation in the tissue and less knowledge of the disease will make the process of detection complex. With the help of automatic computer-aided models, the problems that occurred during breast cancer detection can be solved. Moreover, automated mammogram detection models need to be introduced, which helps the radiologist for providing accurate measures in order to treat the patients for detecting the disease at an early stage[10]. Many models have been introduced for developing discrete hybrid CAD models for effective classification of mammogram images in the previous years [10] [11] [12] [13] [14].

However, the traditional models need to attain more accuracy and less computational time for enhancing the system performance, which assists the radiographer in effective breast cancer detection.

Many of the conventional CAD mammogram methods are based on various transformations to frequency domain procedures such as DCT, DST, and DCT for feature extraction [11] [15] [20]. This type of wavelet transform has benefits, which covers other transformation models, which is similar to spatial data preservation. However, the traditional wavelet transforms experiences from the overhead of computation and memory. Therefore, in [2], an effective wavelet model called LWT is employed for feature extraction from the region of interest's related to mammogram images. The major benefit of LWT has quick computation by acquiring less memory space when compared to conventional wavelet transforms [19]. In earlier contributions, the LWT model is employed, such as image watermarking [16], human fall detection model [17], and audio watermarking [18]. However, the efficiency of the LWT model is not yet analyzed for cases of detecting breast cancer in mammogram images.

The contribution of the entire paper is mentioned below.

- To perform effective detection of the signs of breast cancer through mammogram images using an optimized region growing and fuzzy classifier.
- The analysis of the proposed IAP-CSA is performed by varying the flight length *Fl* from 1.0 to 3.0.

The organization of the paper is designed in the following manner: The review on conventional breast cancer detection models are shown in Section II. Section III describes the implemented breast cancer discovery model through mammogram images. Moreover, the proposed IAP-CSA for breast cancer recognition is described in Section IV. The results and discussions of the whole paper are discussed in Section V. The conclusion of the entire paper is mentioned in Section VI.

LITERATURE REVIEW

A. Related Works

In 2019, Kaur *et al.* [1] have introduced a novel technique, which was implemented on the Mini-MIAS dataset consisting of 322 images, in which the K-mean clustering model was utilized for inbuilt feature extraction in order to select SURF. In the classification level, a new layer was added that performed a ratio of 70% training, and the remaining 30% was used for testing the MSVM and DNN. The results have been revealed that the proposed automated deep learning model attained huge accuracy by K-mean clustering with MVSM, which was superior to the decision tree model. The test outcomes have shown that the average precision rate of 3 classes like "normal, benign, and malignant" with the developed model were acquired, respectively.

In 2020, Muduli et al. [2] suggested an improved CAD method for breast mass classification into normal or diseased and categorizing it into “benign or malignant.” In order to mine the features from the mammogram image's ROI, the developed model employed LWT. With the help of the combination of LDA and PCA approaches, the feature vector dimension was then reduced. At last, the classification was done with the fusion of an ELM and MFO-ELM. In this method, for hidden node parameter optimization of ELM, MFO was utilized. Later, 5-fold stratified cross-validation was utilized for enhancing the performance generalization of the method. The evaluation of the proposed model was done on two benchmark datasets such as DDSM and MIAS. It has been observed that the developed CAD method acquired the best outcomes for both MIAS and DDSM datasets.

In 2019, Vijayeshwari et al. [3] developed a new model for mammogram classification with the feature extracted by Hough transform, a two-dimensional transform. This was employed for feature isolation for a specific shape of an image. There were two significant markers of threat, namely "miniatured scale characterization and masses" and their mechanized recognition was significant for diagnosing breast cancer in early stages. From the encircling parenchymal, automated mass location and collection was majorly complex, as masses were usually not determined. Moreover, the authors demonstrated the methods used for feature extraction as well as classification. In order to detect the mammogram image features, Hough transform was used, and classification was performed by SVM. The analysis has demonstrated that the developed model has good performance over conventional models for accurate classification of mammogram images.

In 2019, Hossain [4] had developed an automated model and segmented the mammogram image's micro-calcification. Initially, for enhancing the quality of the image, the pre-processing applications of images were implemented. Next, the region of the breast was segmented from the pectoral region. By using the “fuzzy C-means” clustering model, the doubtful regions were recognized and split into positive and negative patches. It was used for removing the manual labeling of ROI. In order to train the modified U-net segmentation network, the positive patches that have micro-calcification pixels were considered. Furthermore, for segmenting the region of micro-calcification from the mammogram images automatically, the trained network was employed. In order to improve the segmentation accuracy and early diagnosis, the proposed model was useful for guiding the radiologist.

In 2020, Sha et al. [5] have recommended a comprehensive model for locating the cancerous region present in the mammogram image. The proposed model used image noise reduction, optimal image segmentation on the basis of CNN, and a grasshopper optimization model, and the feature selection, as well as optimized feature extraction, was done on the basis of GOA; thus, the precision was improved, and the computational cost was reduced. In order to screen the mammography breast cancer datasets, the proposed model was implemented on the "Mammographic Image Analysis Society Digital Mammogram Database and Digital Database". The numerical outcomes were compared over ten existing discrete models for assessing the efficiency of the developed model.

B. Review

Although there are many advantages with the existing breast cancer detection using mammogram images, but there are few conflicts with the existing methodologies. A new model needs to be introduced. Table I show a few advantages and disadvantages with the existing methodologies. Among them, K-means clustering + MSVM [1] produce best outcomes when compared over the decision tree model, and it is used for classifying the patterns into groups. But, it needs to be evaluated on large datasets in less time for acquiring accurate results. MFO-ELM [2] is mostly used for solving the classification issue, and it enhanced the generalization capacity, avoids local minima and produces the best results, as well as it has a faster learning rate. Yet, it is required to attain the best classification accuracy. SVM [3] attains huge classification accuracy, and it performs well when the data is implemented outside the training set. However, when the dataset has more noise, it doesn’t perform well. FCM [4] is used to detect the suspicious regions, and it is used to define the “optimal threshold for the frequency histogram” approach. But, Euclidean distance computes can unequally weight underlying factors. CNN [5] is a robust classification tool, and GOA+CNN [5] is used to segment the cancerous regions from the background. Still, the configuration has a huge computational cost. Hence, the challenges that are mentioned above might help the researcher in developing an efficient model for breast cancer detection.

DEVELOPED BREAST CANCER DETECTION MODEL USING MAMMOGRAM IMAGES

A. Proposed Methodology

In the proposed breast cancer detection model, four phases are included, namely "image pre-processing, tumor segmentation, feature extraction, and classification". In the pre-processing phase, the noise is removed from the original image with the help of median filtering. By using the optimized region growing model, segmentation of the tumor is done. Later, the feature extraction is done, in which the features such as "GLCM and GLRM" are extracted. Moreover, the GLCM features include "contrast, energy, entropy, homogeneity, variance, sum average, correlation, sum variance, sum entropy, difference entropy, difference variance, IMC1, IMC2, maximum correlation coefficient, etc.," whereas GLRM features include "Grey level Non-uniformity, Long run emphasis, Run percentage, Run-length Non-uniformity, HGRE, LGRE etc.". In the classification part, the optimized fuzzy classifier is used for image classification. Here, it is classified whether the tumor is normal, benign, or malignant. The proposed breast cancer detection model using mammogram images is shown in Fig. 1.

TABLE I "Features and Challenges of Existing Models for Breast Cancer Detection"

Author [citation]	Methodology	Features	Challenges
Kaur *et al*. [1]	K-Means Clustering+ MSVM	• It produces best outcomes when compared over decision tree model. • It is used for classifying the patterns into groups.	• Needs to be evaluated on large datasets in less time for acquiring accurate results.
Muduli *et al*. [2]	MFO-ELM	• It is mostly used for solving the classification issue. • It enhanced the generalization capacity, avoids local minima and produces best results, as well as it has faster learning rate.	• It is required to attain best classification accuracy.
Vijayeshwari *et al*. [3]	SVM	• It attains huge classification accuracy. • It performs well when the data is implemented outside the training set.	• When the dataset is having more noise, it doesn't perform well.
Hossain [4]	FCM	• It is used to detect the suspicious regions. • It is used to define the optimal threshold for the frequency histogram approach.	• Euclidean distance computes can unequally weight underlying factors.
Sha *et al*. [5]	GOA+CNN	• CNN is a robust classification tool. • It is used to segment the cancerous regions from the background.	• This configuration has huge computational cost.

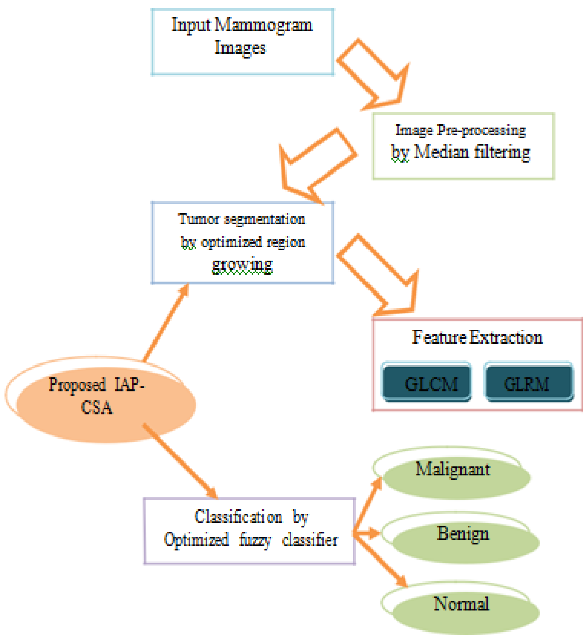


Fig. 1 Architectural representation of proposed breast cancer detection model using mammogram images

The present work focuses on the segmentation and classification of tumor. The optimized region growing algorithm is employed for segmentation, and an optimized fuzzy classifier is employed for classification. With the help of the developed meta-heuristic algorithm named IAP-CSA, the optimization of tolerance of region growing algorithm is done. In addition to this, the same proposed IAP-CSA algorithm uses its advantages for membership function optimization by optimizing the membership function limits.

B. Pre-processing

Here, a median filter is used for removing the noise present in the mammogram image by conquering the low or high frequencies that recognize the image edges. Median filter [22] is the non-linear filter, which is used for substituting the noisy pixels with the neighbourhood pixel's median value, which is sorted on the basis of the image's gray level. The median filter, which is used for substituting the noisy pixels with the neighbourhood pixel's median value, which is sorted on the basis of image's gray level. The median filtered image is denoted as A_a^{pre} represented in Eq. (1). Here, the input image is given by A_a, and the 2D mask is denoted as B

$$A_a^{pre}(i,j) = med\, A_a(i-x, j-y)\quad x, y \in B \tag{1}$$

Therefore, the term A_a^{pre} refers to the pre-processed image, which is further applied for segmentation.

C. Segmentation by Optimized Region Growing

The median filtered image A_a^{pre} is given to the segmentation procedure for tumor segmentation from the image. The advanced version of single seeded region growing algorithm is optimized region growing algorithm, which is used for tumor segmentation. In most of image segmentation models, seeded region growing algorithm is used. It works based on the group of the target pixels into the significant regions of an image. In this, the image segmentation is done by the seed pixel and then merges the newly acquired homogeneous pixels for the seed till the process of segmentation satisfies the enlarged segment. Accordingly, the seed point is selected in the first phase of this algorithm. The inclusion of one pixel to the above set is included in each phase of the seeded region growing algorithm, if the pixel value consists of the condition given in Eq. (2).

$$rg = \frac{1}{e} \sum_{(i,j)\in E} g(i,j) \tag{2}$$

$$\left| g(i,j) - rg \right|_{(i,j)\in E} \tag{3}$$

As per Eq. (2), the image's average is given by rg , in (i, j)) represents the value at the coordinate (i, j). Here, the term is determined as the threshold that must be very close to the image's average. For acquiring the effective automatic segmentation, the optimization of tolerance of seeded region growing algorithm is done by the proposed IAP-CSA algorithm, which must lie in between 0 and 256.

D. Feature Extraction

In this, the features such as GLCM and GLRM are extracted. The description of each is given below.

a) GLCM [24]: It is one of the most famous methods used for texture feature extraction in earlier days. It is mostly used in several texture analysis applications, thus it is remained to be significant feature extraction approach for texture analysis domain. GLCM features such as "energy, contrast, entropy, homogeneity, variance, sum, average, correlation, sum variance, sum entropy, difference entropy, difference variance, information measures of correlation (IMC1, IMC2), and maximum correlation coefficient" are extracted in the proposed breast cancer detection.

b) GLRM [25]: GLRM is the texture feature descriptor, which extracts the data of an image from its gray level runs. GLRM is based on the count of gray level runs computation of various lengths. A gray level run is the accumulation of linearly adjacent picture points by having the equivalent gray level values. Here, the features like "short run emphasis, gray level non uniformity, long run emphasis, run percentage, run length non uniformity, HGRE, and LGRE" are extracted.

A Fuzzy Classifier-based Diagnosis

The concept of the fuzzy set represents and manipulates ambiguity and uncertainty, which is the major beneficiary thing of a fuzzy classifier [26]. The process of the existing fuzzy classifier is to provide the rules and regulations, but here, the accurate image classification is provided with the help of the triangular membership function, and it is mathematically denoted in Eq. (4).

$$ {}_{M}(X) = \begin{cases} 0, & x \;\; lw \\ \dfrac{x-lw}{md-lw}, & lw \;\; x \;\; md \\ \dfrac{hg-x}{hg-md}, & md \;\; x \;\; hg \\ 0, & x \;\; hg \end{cases} \tag{4} $$

In Eq. (4), the term ${}_{M}$ (X) represents the membership function of x in M. The terms lw, md, and hg denotes the low, medium and high operators. The universe of discourse is given by X and the respective element is denoted as x .

Consider, the term $\hat{u_i} = C_r, R^i,$ in which the output in the particular region is given by R^s. The limiting factor ${}_1$ is given in Eq. (5), and the count of linguistic variables is expressed as L_g. The limits of these variables are mathematically represented in Eq. (7), Eq. (8), and Eq. (9), respectively.

$$ {}_1 = \max(\hat{u}_1) - \min(\hat{u}_1) \tag{5} $$

$$ {}_1 = \frac{1}{L_g} \tag{6} $$

$$ \begin{aligned} lw^{min} &= \min(\hat{u}_i) \\ lw^{max} &= \min(\hat{u}_1) + {}_1 \end{aligned} \tag{7} $$

$$ \begin{aligned} md^{min} &= lw^{max} + 0.1 \\ md^{max} &= \min(\hat{u}_1) + 2{}_1 \end{aligned} \tag{8} $$

$$ \begin{aligned} Hg^{min} &= \min(\hat{u}_i) + 2{}_i + 0.1 \\ Hg^{max} &= \max(\hat{u}_1) \end{aligned} \tag{9} $$

Here, the membership function limits play a key role in defining the resultant degree, which needs to be optimized properly. Therefore, the proposed IAP-CSA is used for proper optimization, and it must provide a positive impact on the last membership function.

PROPOSED IAP-CSA FOR BREAST CANCER DETECTION

A. Solution Encoding

Here, the novel contribution on segmentation and classification is generated for the automatic diagnosis of breast cancer. In order to maximize the segmentation and detection accuracy, the optimization of tolerance from the optimized region growing and the membership limits of the fuzzy classifier is done. The solution encoding for both tolerance and membership limits is diagrammatically represented in Fig. 2.

lw	*md*	*hg*

Fig. 2: Solution encoding of fuzzy membership function

B. Objective Model

In the proposed breast cancer detection model, there are two major objectives, which are mentioned below.

By comparing the segmented image with the ground truth image, the first objective is obtained, which is the maximization of segmentation accuracy. It is evaluated in the developed IAP-CSA, which is processed in the region growing for tolerance value optimization.

The other objective is to maximize the classification accuracy, which is obtained from the optimized fuzzy classifier. The classification performance is improved by the optimized membership function. The formula for accuracy is mathematically expressed in Eq. (10).

$$ Acr = \frac{Tr^{Po} + Tr^{Ng}}{Tr^{Po} + Tr^{Ng} + Fs^{Po} + Fs^{Ng}} \tag{10} $$

In Eq. (10), the variables Tr^{Po} , and Tr^{Ng} refers to the true (5) positive, and true negative elements, respectively, as well as the terms Fs^{Po} , and Fs^{Ng} specify the false positive, and false negative elements, respectively.

C. Proposed IAP-CSA

The traditional CSA [21] is motivated by the crow's behaviour towards searching the food that is already stored in the concealed places. However, the traditional CSA provides some advantages, and some issues are faced difficulties during the search for food. The complex optimization issues are taken into consideration in segmentation and classification, and that are resolved by the proposed IAP-CSA algorithm. In traditional CSA, the awareness probability is constant as 0.1, but here the awareness probability is relied on Eq. (11). In this equation, the random number is denoted as rad , which is ranging from 0 to 1, the average of all the fitness values is given by and

Ft_{ti}^{av} ,

Ft_{ti} .
the fitness value of the present solution is denoted as

$$AWP = rad \frac{Ft_{ti}}{Ft_{ti}^{av}} \tag{11}$$

In conventional CSA [21] algorithm, the crow's position iscomputed based on Eq. (12). Here, the term $C^{a,ti} = C_1^{a,ti}, C_2^{a,ti}, ..., C_c^{a,ti}$, $C^{a,ti}\ (a=1,2,...,T; ti=1,2,...,ti_{max})$ and maximum count of iterations is given by tim ax . Moreover, the term D b, ti indicates that the crow b wants to see the hidden location of the food, and the random positionof crow a is denoted as rd a with uniform distribution among 0 and 1. The length of the flight of crow a at ti iteration is given by Fl a,ti . Moreover, the memory update equation is mathematically expressed in Eq. (13), in which the term F (.) represents the fitness value.

$$C^{a,ti+1} = C^{a,ti} + rd_a \ Fl^{a,ti}\left(D^{b,ti} - C^{a,ti}\right) \tag{12}$$

$$D^{a,ti+1} = \begin{cases} C^{a,ti+1} & if \ f(C^{a,ti+1}) \ is \ better \ than \ F(D^{a,ti}) \\ D^{a,ti} & Otherwise \end{cases} \tag{13}$$

The pseudo code of the developed IAP-CSA is depicted in Algorithm 1.

Algorithm 1: Proposed IAP-CSA

```
The position of T crows is initialized
Evaluate fitness values
Initialization of memory is done
 While ti  tim ax
   For 1: T
     The single crow is selected at random
     The awareness probability is computed based on Eq. (11)
       If rd b AWP b,ti
       The location of crow is measured by Eq. (12)
       else
         Random position needs to be selected
       End if
     End for
   Evaluate the vision of new locations
   Evaluate fitness values
   Memory update is done by Eq. (13)
End while
```

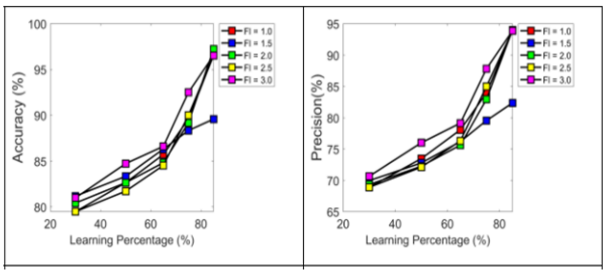
RESULTS AND DISCUSSIONS

A. Experimental Setup

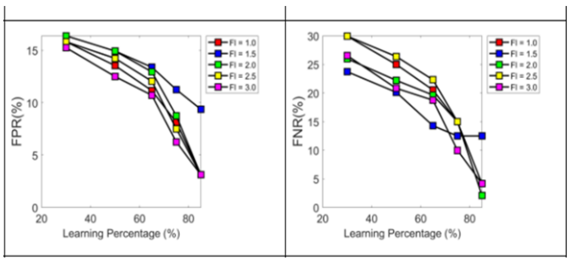
The proposed model for detecting breast cancer was implemented using MATLAB 2018a. In order to evaluate the performance of the proposed model, the mammogram image dataset was downloaded from the URL (https://www.kaggle.com/kmader/miasmammography: Access data 2020-05-07). For analysis, the population size was considered as 10 and the maximum number of iterations considered for segmentation and classification as 25 and 100, respectively. The evaluation measures such as "accuracy, sensitivity, specificity, precision, FPR, FNR, NPV, FDR, F1 score, and MCC" were considered for both segmentation and classification.

B. Performance Analysis

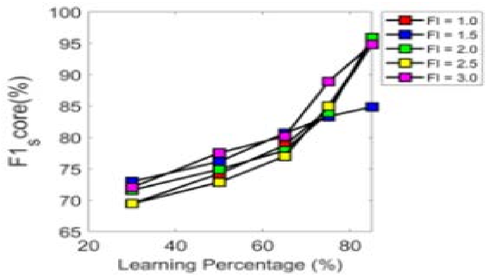
The performance analysis of the proposed model by varying the flight length Fl from 1.0 to 3.0 with respect to learning percentage is shown in Fig. 3. From Fig. 3 (a), the accuracy of the developed IAP-CSA algorithm at learning percentage 50, Fl = 3.0 is acquiring the best performance. It is 2.4% better than Fl = 1.5 , 2.6% better than Fl = 2.0 , and 4.9% better than Fl = 2.5. The precision of the suggested IAP-CSA algorithm is acquiring the best performance at any learning percentage and it is shown in Fig. 3 (b). The recommended IAP-CSA at Fl = 3.0 is 13.2% advanced than Fl = 1.5 when considering the learning percentage as 85. Moreover, the FPR of the developed IAP-CSA is attaining best performance at Fl = 2.0 , which is 3.2% improved than Fl = 2.5, and 6.6% improved than Fl = 3.0. In Table II, the overall performance of the proposed algorithm by varying the flight length Fl is shown. The accuracy of the suggested IAP-CSA acquired best performance when Fl = 3.0. It is 3.2% better than Fl = 1.0, 4.7% better than Fl = 1.5, 3.7% better than Fl = 2.0, and 2.7% better than Fl = 2.5. Moreover, the precision of the developed IAP-CSA algorithm is attained best performance at Fl = 3.0 , which is 4.5% advanced than Fl = 1.0, 10.3% advanced than Fl = 1.5 , 5.8% advanced than Fl = 2.0 , and 3.3% advanced than Fl = 2.5 . Thus, the proposed IAP-CSA is performing well in detecting the breast cancer.



(a) **(b)**



(c) **(d)**



(d)

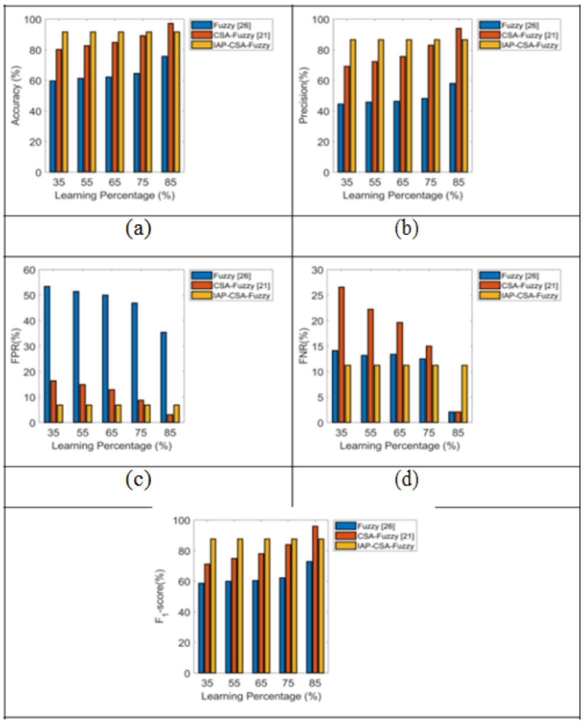
Fig. 3: Analysis on the proposed model by varying the flight length for breast cancer detection with respect to learning percentage concerning the performance measures such as "(a) Accuracy, (b) Precision, (c) FPR, (d) FNR, and (e) F1 score"

Table II "Overall Performance of Proposed Algorithm by Varying the Flight Length for Breast Cancer Detection"

Performance Metrics	$Fl = 1.0$	$Fl = 1.5$	$Fl = 2.0$	$Fl = 2.5$	$Fl = 3.0$
'Accuracy'	0.89583	0.88333	0.89167	0.9	0.925
'Sensitivity'	0.85	0.875	0.85	0.85	0.9
'Specificity'	0.91875	0.8875	0.9125	0.925	0.9375
'Precision'	0.83951	0.79545	0.82927	0.85	0.87805
'FPR'	0.08125	0.1125	0.0875	0.075	0.0625
'FNR'	0.15	0.125	0.15	0.15	0.1
'NPV'	0.91875	0.8875	0.9125	0.925	0.9375
'FDR'	0.16049	0.20455	0.17073	0.15	0.12195
'F1-score'	0.84472	0.83333	0.83951	0.85	0.88889
'MCC'	0.76639	0.7459	0.7579	0.775	0.83244

C. Effect of Optimized Fuzzy

The analysis of the proposed and the traditional fuzzy classifiers is provided with respect to learning percentage, which is shown in Fig. 4. In Fig. 4 (a), the accuracy of the suggested IAP-CSA is attaining the best performance at any of the learning percentages. When considering the learning percentage as 85, the accuracy of the improved IAP-CSA-Fuzzy is 12.5% progressed than CSA-Fuzzy, and 50% progressed than Fuzzy. Moreover, the precision of the introduced IAP-CSA-Fuzzy is detecting breast cancer accurately, and it is shown in Fig. 4 (b). It is 21.4% advanced than CSA-Fuzzy, and 88.8% advanced than Fuzzy at a learning percentage of 55. Table III shows the overall classification analysis of the proposed and the conventional classifiers. In this, the accuracy of the suggested than CSA-Fuzzy, and 50% progressed than Fuzzy. Moreover, the precision of the introduced IAP-CSA-Fuzzy is detecting breast cancer accurately, and it is shown in Fig. 4 (b). It is 21.4% advanced than CSA-Fuzzy, and 88.8% advanced than Fuzzy at a learning percentage of 55. Table III shows the overall classification analysis of the proposed and the conventional classifiers. In this, the accuracy of the suggested IAP-CSA-Fuzzy is 41.9% upgraded than Fuzzy and 2.8% upgraded than CSA-Fuzzy. In addition, the precision of the recommended IAP-CSA-Fuzzy is 79.3% superior to Fuzzy and 4.4% superior to CSA-Fuzzy. Hence, it is concluded that the developed IAP-CSA-Fuzzy is performing well in diagnosing breast cancer.



(e)

Fig. 4: Analysis on the proposed and conventional Fuzzy for breast cancer detection with respect to learning percentage concerning the performance measures such as "(a) Accuracy, (b) Precision, (c) FPR, (d) FNR, and (e) F1 score"

Table III: "Overall Performance of Proposed and Conventional Classifiers for Breast Cancer Detection"

Performance Metrics	Fuzzy [26]	CSA-Fuzzy [21]	IAP-CSA-Fuzzy
'Accuracy'	0.64583	0.89167	0.91667
'Sensitivity'	0.875	0.85	0.8875
'Specificity'	0.53125	0.9125	0.93125
'Precision'	0.48276	0.82927	0.86585
'FPR'	0.46875	0.0875	0.06875
'FNR'	0.125	0.15	0.1125
'NPV'	0.53125	0.9125	0.93125
'FDR'	0.51724	0.17073	0.13415
'F1-score'	0.62222	0.83951	0.87654
'MCC'	0.39161	0.7579	0.81381

CONCLUSION

This paper has presented a new model for breast cancer detection automatically. Here, the mammogram image was considered as an input for the pre-processing phase, in which the median filtering model was utilized for noise removal existing in the image. Further, the optimized region growing algorithm was employed for segmenting the tumor. Next, from the segmented tumor, the features like GLCM and GLRM were extracted. Later, the features were subjected to the optimized fuzzy logic classifier. The proposed IAP-CSA was used to optimize the threshold of the region growing algorithm as well as the fuzzy classifier's membership function.

REFERENCES

1. Prabhpreet Kaur, Gurvinder Singh, and Parminder Kaur, "Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification," Informatics in Medicine Unlocked, vol. 16, pp. 1-19, March 2019.
2. Debendra Muduli, Ratnakar Dash, and Banshidhar Majhi, "Automated breast cancer detection in digital mammograms: A moth flame optimization based ELM approach," Biomedical Signal Processing and Control, vol. 59, pp. 1-12, May 2020.
3. R.Vijayarajeswari, P.Parthasarathy, S.Vivekanandan, and A. AlavudeenBasha, "Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform," Measurement, vol. 146, pp. 800-805, November 2019.
4. Md ShamimHossain, "Microc alcification Segmentation Using Modified U-net Segmentation Network from Mammogram Images," Journal of King Saud University - Computer and Information Sciences, pp. 1-9, November 2019.
5. Zijun Sha, Lin Hu, and Babak Daneshvar Rouyendegh, "Deep learning and optimization algorithms for automatic breast cancer detection," International Journal of Imaging systems and Technology, vol. 30, no. 2, pp. 495-506, June 2020.
6. S. Rahmatinia and B. Fahimi, "Magneto-Thermal Modeling of Biological Tissues: A Step Toward Breast Cancer Detection," IEEE Transactions on Magnetics, vol. 53, no. 6, pp. 1-4, June 2017.
7. P. Parthasarathy & S. Vivekanandan, "A typical IoT architecture-based regular monitoring of arthritis disease using time wrapping algorithm," International Journal of Computers and Applications, vol. 42, no. 3, pp. 222-232, May 2020.
8. J. Xu et al., "Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images," IEEE Transactions on Medical Imaging, vol. 35, no. 1, pp. 119-130, Jan. 2016.
9. Parthasarathy P, and Vivekanandan S, "Investigation on uric acid biosensor model for enzyme layer thickness for the application of arthritis disease diagnosis," Health Inf Sci Syst, vol. 6, no. 1, 23 April 2018.
10. H.D.Cheng, Xiaopeng Cai, Xiaowei Chen, Liming Hu, and Xueling Lou, "Computer-aided detection and classification of microcalcifications in mammograms: a survey," Pattern Recognition, vol. 36, no. 12, pp. 2967-2991, December 2003.
11. Shradhananda Beura, Banshidhar Majhi, and Ratnakar Dash, "Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer," Neurocomputing, vol. 154, pp. 1-14, 22 April 2015.
12. J.Dheeba, N.Albert Singh, and S.Tamil Selvi, "Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach," Journal of Biomedical Informatics, vol. 49, pp. 45-52, June 2014.
13. Rahimeh Rouhi, Mehdi Jafari, Shohreh Kasaei, and Peiman Keshavarzian, "Benign and malignant breast tumors classification based on region growing and CNN segmentation," Expert Systems with Applications, vol. 42, no. 3, pp. 990-1002, 15 February 2015.