World Applied Sciences Journal 31 (6): 1056-1064, 2014
ISSN 1818-4952

DOI: 10.5829/idosi.wasj.2014.31.06.203

Mathematical Morphological Approach for Mammogram Image Segmentation and Classification

[1]S. Pitchumani Angayarkanni and [2]Nadira Banu Kamal

[1]Department of Computer Science, Lady Doak College, Madurai, Tamil Nadu, India
[2]Department of M.C.A. TBAK College, Kilakarai, Ramnad,Tamil Nadu, India

Abstract: This paper presents the mathematical morphological and rough set based approach in detection and classification of cancerous masses in MRI mammogram images. The main objective behind this approach is to build a CAD system with good accuracy and computational speed in detection of cancerous masses compared to the existing system. The ROI(Region of Interest) is segmented using Graph cut method.and the fourteen features including morphological,shape and novel features are calculated for this region. Best rules used for classification are generated using ID3 algorithm. Automatic classification based on the rules generated are determined using Artificial bee colony based Multi Layered Perceptron model.The sensitivity, the specificity, positive prediction value and negative prediction value of the proposed algorithm accounts to 98.79%, 98.8%, 92% and 96.6% which rates very high when compared to the existing algorithms. The area under the ROC curve is 0.89. A GUI based tool was developed for the proposed methodology. An android application using simulator was developed to make the doctor and patient to view the image with appropriate information like Patient Name, age,Size of tumor, Nature of tumor and type of treatment.

Key words: Fuzzification · Graph cut · ID3 and Artificial Bee colony technique

INTRODUCTION

The Population Based Cancer Registry evidently shows from the various statistics, that the incidence of breast cancer is rapidly rising, amounting to a significant percentage of all cancers in women. Breast cancer is the commonest cancer in urban areas in India and accounts for about 25% to 33% of all cancers in women. Over 50% breast cancer patients in India present in stages 3 and 4, which will definitely impact the survival [1]. The survival rate can be increased only through the early diagnosis. Image processing technique together with data mining is used for extraction and analysis of the ROI. Tumor can be classified into three category normal, benign and malignant. A normal tumor is a mass of tissue which exists at the expense of healthy tissue. Malignant tumor has no distinct border. They tend to grow rapidly increasing the pressure within the breast cells and can spread beyond the point from they originate. Grows faster than benign and cause serious health problem if left unnoticed. Benign tumors are composed of harmless cells, have clearly defined borders, can be completely removed and are unlikely to recur. In MRI mammogram images after the appropriate segmentation of the tumor, classification of tumor into malignant, benign and normal is difficult task due to complexity and variation in tumor tissue characteristics like its shape, size, grey level intensities and location. Feature extraction is an important aspect for pattern recognition problem. A Hybrid rough set based mathematical approach for automatic detection and classification of cancerous masses in mammogram images is proposed in this paper.

MATERIALS AND METHODS

The data set used for research were taken from Mammogram Image Analysis Society (MIAS) [2]. The database contains 320 images out of which 206 are normal images, 63 benign and 51 malignant cases.

Corresponding Author: S. Pitchumani Angayarkanni, Department of Computer Science, Lady Doak College, Madurai, Tamil Nadu, India.

Image Enhancement and Preprocessing Using Fuzzy Histogram Equalization: Mammogram images are enhanced using histogram equalization method. The approach developed by Ella Hassanien is implemented for Image preprocessing[3].An image I of size M x N and L gray levels can be considered as an array of fuzzy singletons, each having a value of membership denoting its degree of brightness relative to some brightness levels. For an image I, we can write in the notation of fuzzy sets [4]:

$I = U\mu_{mn}/g_{mn}$ m = 1,2,... M and n = 1,2,... N, Where μ_{mn} its membership function g_{mn} is the intensity of (m, n) pixel. The membership function characterizes a suitable property of image (e.g. edginess, darkness, textural property) and can be defined globally for the whole image or locally for its segments. In recent years, some researchers have applied the concept of fuzziness to develop new algorithms for image enhancement [3]. The principle of fuzzy enhancement scheme is illustrated in Figure 1.

The histogram equalization of the gray levels in the original image can be characterized using five parameters: (α, ß1, γ, ß2, max) as shown in Figure 2.

where the intensity value γ represents the mean value of the distribution, α is the minimum and max is the maximum. The aim is to decrease the gray levels below ß1 and above ß2. Intensity levels between ß1 and γ and ß2 and γ are stretched in opposite directions towards the mean γ.The fuzzy transformation function for computing the fuzzy plane value P is defined as follows:

α = min;
$\beta 1$ = $(\alpha + \gamma)/2$;
$\beta 2$ = $(max + \gamma)/2$;
γ = mean; max;

The following fuzzy rules are used for contrast enhancement based on Figure (1).

Rule-1: If $\alpha \le ui < \beta 1$ then $P = 2 ((ui - \alpha) / (\gamma - \alpha))^2$
Rule-2: If $\beta 1 \le ui < \gamma$ then $P = 1- 2 ((ui - \gamma)/(\gamma - \alpha))^2$
Rule-3: If $\gamma \le ui < \beta 2$ then $P= 1- 2((ui - \gamma) / (max - \gamma))^2$
Rule-4: If $\beta 2 \le ui < max$ then $P = 2 ((ui - \gamma)/(max - \gamma))^2$
Where ui = f(x,y) is the ith pixel intensity.

Pseudocode:
Step 1: Parameter Initialization

Set β1 = (min + mean)/2
Set β2 = (max + mean)/2

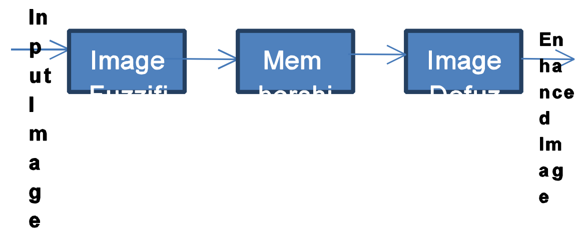

Fig. 1: Fuzzification technique

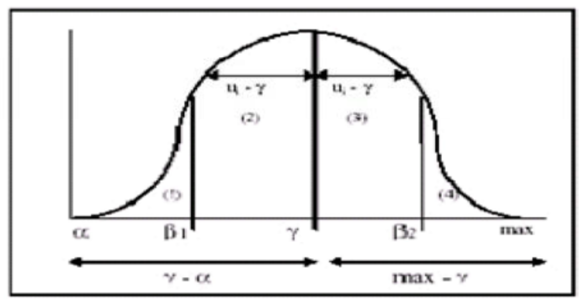

Fig. 2: Histogram Equalization curve with *a* , ß1, γ, ß2, max

Step 2: Fuzzification

For all pixels (i,j) within the image Do
If ((data[i][j]>=min) && (data[i][j]< β1))
Compute NewGrayLevel =2*(pow(((data[i][j]-min)/(mean-min)),2))
If ((data[i][j]> = β1) && (data[i][j] < mean))
Compute NewGrayLevel=1-(2*(pow(((data[i][j]-mean)/(mean-min)),2)))
If ((data[i][j]>=mean)&&(data[i][j]< β2))
Compute NewGrayLevel=1-(2*(pow(((data[i][j]-mean)/(max-mean)),2)))
If ((data[i][j] >= β2) && (data[i][j] <max))
Compute NewGrayLevel=2*(pow(((data[i][j]-mean)/(max-mean)),2))

Step 3: Fuzzification Modification

Compute FuzzyData[i][j]= pow(NewGrayLevel,2)

Step 4: Defuzzification

For all pixels (i,j) within the image Do
Compute Enhanced Data [I][j] = Fuzzy Data [I][j] * data [i][j];

The resultant image contrast enhancement and grayness ambiguity is checked using Peak Signal to noise ratio, Index of fuzziness and Fuzzy entropy technique.

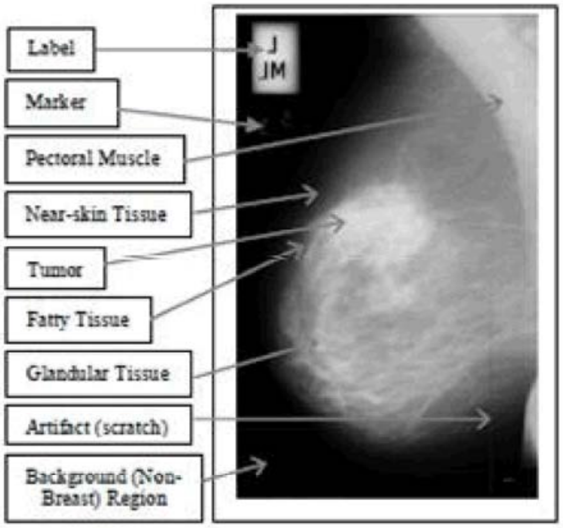

Fig. 3: Mammogram MRI image

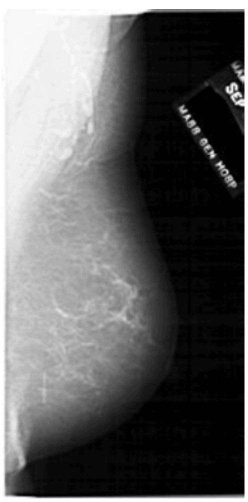

Fig. 4: Input image

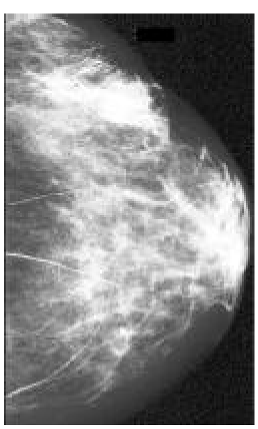

Fig. 5: Output Image

Mean Square Error(MSE) is calculate using the formula:

Table 1: Image Enhancement Prediction of the proposed method

PSNR	RMS	H	γ	MSE	Nature of Filter
87.65	2.97	0.2111	0.0086	8.83	Gabor

$$1/\mathrm{MN}\sum_{i=1}^{m}\sum_{j=1}^{n}\left|x(i,j)-x'(i,j)\right|^2$$

Peak Signal to Noise ratio:
PSNR=20log10 [(2n-1)/vMSE]

where x(i,j) is the input image x'(i,j) is the output image and n is the number of bits used in representing the pixel of the image.

Index of Fuzziness & Fuzzy Entropy are the measures of global grayness ambiguity of a n image. They are defined as a degree of difficulty in decising whether a pixel would be treated as black(dark) or white(bright). Index of fuzziness is the amount of fuzziness present in an iamge which determines the amount of vagueness by measuring the distance between the fuzzy property plane and nearest ordinary plane. Entropy (H) is a measure of quality of information in an image in the fuzzy domain based on Shanon's function.

Index of Fuzziness: $\gamma = 2/MN\sum_{M}\sum_{N}\min\left(\mu mn, 1-\mu mn\right)$

Entropy: H=1/MN ln

$$\sum_{m}\sum_{n}\min In\left(\mu mn\right)-\left(1-\mu mn\right)In\left(1-\mu mn\right)$$

The tabular column clearly shows that the PSNR value is high therefore the image is enhanced and the Index of fuzziness and Entropy decreases with enhancement.

Image Segmentation and ROI Extraction: The region of Interest ie) the tumor region is segmented using the Graph cut method. The main purpose of using this method for segmentation is that it segments the mammogram into different mammographic densities. It is useful for risk assessment and quantitative evaluation of density changes. Apart from the above advantage it produces the contour (closed region) or a convex hull which is used for analyzing the morphological and novel features of the segmented region. The above technique results in efficient formulation of attributes which helps in classification of the ROI into benign, malignant or normal.Graph Cuts has been used in recent years for interactive image segmentation[4]. The core ideology of Graph Cuts is to map an image onto a network graph and construct an energy function on the labeling and then do energy minimization with dynamic optimization techniques. This study proposes a new segmentation method using iterated Graph Cuts based on multi-scale smoothing. The multi-scale method can segment mammographic images with a stepwise process from global to local segmentation by iterating Graph Cuts. The Graph cut approach used by K. Santle Camilus [4] is implemented in this paper

Steps involved in Graph Cut Segmentation are:

- Form a graph
- Sort the graph edges
- Region merging

A graph G=(V,E) is constructed from the mammogram image such that V represent the pixel values of the 3*3 image and E the edges defined between the neighboring pixels.The weight of any edge W(Vi,Vj) is measure of dissimilarity between the pixels Vi and Vj. The weight for an edge is by means of considering the Euclidian distance between the two pixels Vi and Vj. It is represented by the equation

$$W(Vi,Vj)=\sqrt{}(xi - xj)\ 2 + (yi - yj)2 \qquad (1)$$

Vi=(xi,yi) Vj=(xj,yj)

Edges are sorted in ascending order of their weights such that w(e1)<=w(e2).Pick one edge ei in the sorted order from ei to en where ei is between two groups of pixels.This edge determines whether to merge the two groups of pixel to form a single group or not. Each vertex is considered as a group.. If the merge criterion is satisfied, then the two groups are merged. The multiple groups of pixels representing different regions or objects are obtained.

Intra-Region Edge Average: The intra-region edge average (IRA) which is a single-valued function and represents the homogeneity of a group is defined as

$$IRA(R)=\sum_{(\mathrm{vi,vj})\in E}^{\mathrm{vi,vj}\in \mathrm{R}} W(\mathrm{Vi,Vj}) \div |\mathrm{Va}| \qquad (2)$$

IRA for a region "*R*" is a measure of the mean of the weight of edges in "*R*." V_a is a set of edges in the region "*R*" which can be represented as

$$Va=\{(V_{i,}V_j)^a E|V_i,V_j^a R\} \qquad (3)$$

Inter-Region Edge Mean: The discussion for computing IRA can be applicable to compute the homogeneity representative for intermediate edges between the two regions. Hence, we define inter-region edge mean (IRM) as

$$\mathrm{IRM(R1,R2)}=\sum_{(\mathrm{Vi,Vj})\in E}^{\substack{\mathrm{Vi}\in \mathrm{R2} \\ \mathrm{Vi}\in \mathrm{R2}}} W(\mathrm{Vi,Vj})/|\mathrm{Vb}| \qquad (4)$$

IRM is a measure of the mean of the weight of edges between the two regions, R_1 and R_2. V_b is a set of edges between the two regions, R_1 and R_2, which can be represented as

$$Vb=\{(V_{i,}V_j)^a E|V_i{}^a\ R_1,V_j{}^a R_2\} \qquad (5)$$

Dynamic Thresholds: The degree of variations among pixels of a region can be inferred from the homogeneity representative of the region IRA. To merge pixels of two regions, the IRM must be above a threshold value. This threshold must be computed based on the IRA and other parameters to control the merge operation adaptive to the properties of regions. Hence, we define a dynamic threshold (DT) for merging between two regions R_1 and R_2 as given in the following equation

$$DT(R1,R2)=Max(IRA(R1)+\delta 1,IRA(R2)+ \delta 2) \qquad (6)$$

The appropriate selection of the parameters (δ_1 and δ_2) plays a vital role as it determines the merge of regions. Values of these parameters are chosen to achieve the following desirable features: (1) when more number of regions is present, then the choice of the values (of δ_1 and δ_2) should favor merging of regions; (2) the choice of the values should allow the grouping of small regions than large regions. The large regions should be merged only when their intensity values are more similar; (3) it is preferred to have values that are adaptive to the total number of regions and the number of vertices in the regions that are considered for merge. Hence, the parameters δ_1 and δ_2 are defined as

$$\delta 1=C \times N_R/|R_1|$$
$$\delta 2=C \times N_R/|R_2| \qquad (7)$$

N_R - the total number of regions. Initially, $N_R=|V|$, the size of the vertices and each merge decreases the value of N_R by one. $|R_i|$ indicates the number of vertices in the region "R_i." C is a positive constant and is equal to two for the pectoral muscle segmentation which is found through experimental studies.

Merge Criterion: When the pixels of a group have intensity values similar to the pixels of the other group, then intuitively the calculated IRM between these groups should be small. The expected smaller value of the IRM to

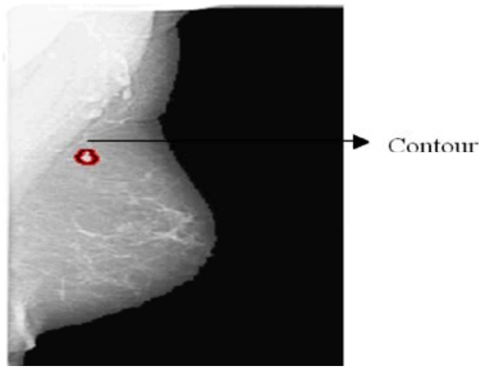

Fig. 6: ROI Extracted using Graph Cut method

merge these two regions is tested by comparing it with the dynamic threshold. Hence, the merge criterion, to merge the two regions, R_1 and R_2, is defined as

$$\text{Merge}(R_1,R_2), \text{ if } IRM(R_1,R_2) \leq DT(R1,R2) \tag{8}$$

Feature Extraction and Selection: The main objective of feature extraction and selection is to maximize the distinguished performance capability of them on interested data base. Consideration of significant feature is based on discrimination, optimality, reliability and independence. In recent years, many kinds of features have been reported for breast mass classification like texture based features, region based features, image structure features, position related features and shape-based features which are described and utilized in the CAD systems [4]. In our study a set of six novel features based on shape characteristics (Novel Feature) proposed by Rangaraj M. Rangayyan [5]were extracted from the original images which are not applied in breast mass classification before and are defined as follows:

Fourier descriptors and Fourier coefficient A0	A_0=1/N $\sum_{i=1}^{360} d(i)$ Where d(i)=v(X(i)-X_c)^2 +(Y(i)-Y_c)^2 where (x_c, y_c) are the coordinates of the centroid, (x(i), y(i)) are the coordinates of the boundary pixel at the i-th location.
Convexity	The measure of convexity of an object is the ratio of the perimeter of the convex object to the original perimeter of the object.Convexity=convexperimeter/perimeter
Aspect Ratio	Aspect Ratio is the ratio between height and width of the corresponding object
Circularity	$4\pi S/D^2$
Compactness- The compactness is a simple measure of circularity of the ROI ∂M and is given byCO=1-4π¦M¦/diameter^2(M) This feature is zero iff M is a disk and grows towards unity with rising complexity ∂M	Eccentricity: It is a scalar that describes the eccentricity of the ellipse which has the same second moment as the mass region.
Fractional Concavity - The fractional concavity is given by the Euclidean length of all concave segments in the polygonal approximation of p divided by the total Euclidean length of all segments	Solidity: It is a scalar that describes theproportion of the pixels in the mass region to the pixels in the convex hull including the mass. It is computed as (Mass's area)/ (Convex hull's area).
Spicularity Index: A spicule is defined as a sequence of segments Iik , . . . , Iik+w, ≥ with w = 1, between a pair of consecutive inflection points p(jk),p(jk+1). If w = 1, then the spicule is merged with its neighbor ,so we can assume that every spicule consists of at least two segments (w ≥ 2). Let T1, . . .,Tw-1 be the angles between consecutive segments, and let, ΘTth = (T1 + . . . + Tw-1)/(w - 1) be the average angle. Given the spicule,narrowness Θk =£{i Ti=Tth}Ti/|{i Ti = Tth}|, and the spicule lengthSk ||Iik|| + . . . + ||Iik+w_||, we calculate the spicularity index by the formula The SI tends to zero for less curvilinear contours, and grows with the appearance of spicules.	
The SI can reach the maximum of 2. Unfortunately, this feature is sensitive to the order in which spicules are merged. Moreover, it is fragile to small contour variations. For example, adding a small and narrow protrusion to the contour may change the SI drastically.	Cross-Correlation left: It is the cross correlation value between the RectangularX region and the Left-side region.
Acutance: The acutance is widely known in signal theory. In image processing it is defined as follows. At first the bundle of normals to the ROI p is computed: for every i and for all k = K = 5 the normals Wk(i) to the	DifferenceAreaMass-RectangualrX: It is absolute value of convex RectangularX's area minus mass's area
Roughness	Perimeter-Area Method: The perimeter L is related to the enclosed area A for a non fractal closed curve in the Plane L=cvA. where C is a constant for a given shape. C=1 for square and c=2v? for circleD is computed as L=C(vA)D

The convex hull which is the smallest convex polygon that contains the mass region is displayed. Mass's area is the actual number of pixels in the mass region and convex hull's area is the actual number of pixels in the convex hull region.

The shape features are calculated for the ROI segmented using Graph cut method for 320 images. The results are plotted in an Excel sheet for classification using WEKA tool.

Feature Selection Using ID3 Algorithm: An ID3 algorithm was implemented to generate the rules based on the attributes which play a vital rule in classificationusing Entropy and Information gain.Mathematical algorithm for building the decision tree. Builds the tree from the top down, with no backtracking. Information Gain is used to select the most useful attribute for classification [9].This method was implemented in the paper "Automatic detection and classification of cancerous masses in mammogram", 2012 Third International Conference on Computing Communication and Networking Technologies (ICCCNT 12),2012.

The attributes used were the eight parameters with the class as benign, malignant and normal. Based on the rule derived by testing 320 mammogram images the rules are applied in classifying the new cases without prior knowledge of whether they are benign, malignant or normal.

Table 2: Nodes representing the 7 attributes

Node1	Compactness
Node2	Fractional Concavity
Node3	Fourier Descriptor
Node4	Convexity
Node5	Aspect Ratio
Node6	Circularity
Node 7	Compactness
Node 8	Fractional Concavity
Node 9	Spicularity Index
Node 10	Acutance
Node 11	Eccentricity
Node 12	Solidity
Node 13	Cross Correlation
Node 14	Difference Area-Mass Rectangle
Class	Benign/Malignant/Normal

Cross validation 10 fold

=== Detailed Accuracy By Class ===

Class	TPRate	FPRate	Precision	Recall	FMS	ROC
Benign	0.595	0.272	0.523	0.595	0.557	0.677
Malig	0.121	0.104	0.368	0.121	0.182	0.542
Normal	0.678	0.427	0.442	0.678	0.536	0.637
Wgt.	0.465	0.268	0.445	0.465	0.425	0.618

=== Confusion Matrix ===

a	b	c	<-- classified as
172	30	87	| a = Benign
94	35	160	| b = Malignant
63	30	196	| c = Normal

The classification accuracy obtained from these fixed-order trees can be compared with those from trees of different feature orders, as well as with those from trees of different feature combinations. The decision tree with the optimal feature combination and order for this task can thus be identified. Examples of the training and test results obtained in this study will be discussed below. It should be noted that, to use a trained decision-tree classifier, one has to choose the specific tree structure with the set of decision thresholds corresponding to the desired sensitivity (TPF) and specificity (~-FPF).The structure and the thresholds will be fixed during testing or application. The decision rule is indicated below

Rule 1:

- Compactness< -0.0450 then Class = Benign (64.89 % of 262 examples)
- Compactness>= -0.0450 then Class = Malignant (62.34 % of 316 examples).
- Compactness>= -0.0050 then Class = Normal (42.29 % of 525 examples)

Rule 2:

- Roughness<2.4 then class=Benign
- Roughness>=2.4 and <=2.7 then class=Malignant
- Roughness>2.7 then class=normal

ID3 Parameters:

- Size before split 200
- Size after split 50
- Max depth of leaves10
- Goodness of split threshold 0.0300

- Features Extracted by ID3 algorithm
- Roughness- Plays a vital Role
- Compactness- Plays a vital Role
- Circularity
- Spicularity Index
- Fractional Concavity

Multilayer Feed Forward Neural Network (MLN) Optimized by Artificial Bee Colony Optimization Technique: Artificial bee colony algorithm(ABC) was proposed for optimization, classification and neural network problem solution based on the intelligent foraging behavior of honey bee. ABC is more powerful and most robust on multimodal functions. It provides solution in organized form by dividing the bee objects into different tasks such as employed bees, onlooker bees and scout bees. It finds global optimization results, optimal weight values by bee agents. Successfully trained MLN improves the classification precision of MRI mammogram images into benign, malignant and normal cases.

The ANN is the best method for pattern recognition. The seven attribute values together with the class are fed as input to the eight neurons at the input layer. The hidden layer has sixteen neurons or nodes. The output layer has one node. The error rate is calculated by subtracting the actual output with the desired output value. The evaluation of the food sources that are discovered by the randomly generated populations is done through the back-propagation procedure. Now after completion of the first cycle the performance of each solution is kept in an index. In order to generate the population of back-propagation, the weight must be created and each value of the weight matrix for every layer is randomly generated within the range of [1,0]. Each set of weight and bias is used to create one element in a whole population. The size of the population in the proposed method is fixed to N.

The classifier employed in this paper is a three-layer Back propagation neural network. The Back propagation neural network optimizes the net for correct responses to the training input data set. More than one hidden layer may be beneficial for some applications, but one hidden layer is sufficient if enough hidden neurons are used. Initially the attributes values (rules) from the Mathematical morphological and novel approach analysis method which plays a vital role in classification between benign,malignant and normal cases extracted from the ID3 algorithm are normalized between [0,1]. That is each value in the feature set is divided by the maximum value from the set. These normalized values are assigned to the input neurons. The number of hidden neurons is equal to the number of input neurons and only one output neuron.

The procedure employed byHabib Shah in his work was implemented for training the mammogram data set using ANN.

Employed Bees: It uses multidirectional search space for food source and get information to find food source and solution space. The employed bees shares information with onlooker bees. They produce a modification on the sources position in her memory and discover a new food sources position. The employed bee memories the new source position and forgets the old one when the nectar amount of the new source is higher than that of the previous source.

Onlooker Bees: It evaluates the nectar amount obtained by employed bees and chooses a food source depending on the probability values calculated using the fitness values. A fitness based selection technique can be used for this purpose. Watches the dance of hive bees and select the best food source according to the probability proportional to the quality of that food source.

Scout Bees: Select the food source randomly without experience. If the nectar amount of the food source is higher that that of the previous source in their memory, the new positions are memorized and forgets the previous position. The employed bee becomes scout bees whenever they get food source and use it very well again. These scout bees find the new food source by memorizing the best path.

- Initialize the population of solutions Xi where i=1…..SN
- Evaluate the population and calculated the fitness function for each employed bee Initialize weights, bias for Multilayerd Feed forward network
- Cycle=1
- Repeat from step 2 to step 13
- Produce new solutions (food source positions) Vi,j in the neighbourhood of Xi,j for the employed bees using the formula

$$V_{i,j}=X_{i,j}+\varphi_{ij}(X_{i,j}-X_{k,j})$$

where k is a solution in the neighborhood of i, φ is a random number in the range [-1, 1] and evaluate them.

- Apply the Greedy Selection process between process
- Calculate the probability values pi for the solutions Xi by means of their fitness values by using formula.

$$Pi=fit_i / \sum_{k=1}^{SN} fit_n$$

The calculation of fitness values of solutions is defined as

$$\mathrm{Fit}_i = \begin{Bmatrix} \dfrac{1}{1+fi} & \left| fi \geq 0 \right. \\ 1+abs(fi) & \left| fi < 0 \right. \end{Bmatrix}$$

Normalize pi values into[0,1]

- Produce the new solutions (new positions) õi for the onlookers from the solutions xi, selected depending on Pi and evaluate them
- Apply the Greedy Selection process for the onlookers between xi and vi
- Determine the abandoned solution (source), if exists, replace it with a new randomly produced solution xi for the scout using the following equation

$X^j_i = x^j_{min} + \mathrm{rand}(0,1)\ (x^j_{max} - x^j_{min})$

- Memorize the best food source position (solution) achieved so far
- cycle=cycle+1
- until cycle= Maximum Cycle Number (MCN)

The classification accuracy is determined using Mean square error and standard variance of mean square error on 30 cycles

The Dataset with best attribute generated by ID3 algorithm is fed as input for benign,normal and malignant cases.

Table 3: Parameter Values

Mass type	Training data set	Testing
Benign	600	425
Malignant	600	425
Normal	600	425

Table 4: Classification Accuracy

SN			20
MaxEpoch			400
Fitness value			10
Algorithm	Mean Square Error	Std. Var of Mean Square Error	Rank
ABC+MLP	1.67 X 10-4	0.0001	1

GUI Based Tool: An android application was developed using emulator to view the mammogram image in "GIF" format together with the segmented region affected by cancer indicated in red circle followed by the Patient Name,Age,Type of tumor,Size of Tumor and nature of treatment. The entire mathematical approach was implement using Scilab and Matlab.The Data base used was mysql and R2. The mobile screen and resolution compatible application was developed.

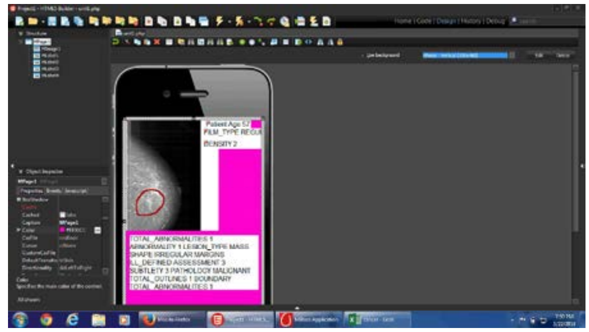

Fig. 7: Android Apps Showing Patient Details

CONCLUSION

The proposed mathematical approach yields a high level of accuracy in minimum period of time that shows the efficiency of the algorithm. The GUI based CAD system was developed using Scilab and R2. The training speed accounts to 6 ms using ANN.The main goal of classifying the tumors into benign,malignant and normal is achieved with a great accuracy compared to other techniques.

Table 5: Accuracy Details

Specificity	98.8%
Sensitivity	98.79%
Positive Prediction value	92%
Accuracy	98.9%
Area under Curve	0.98
Negative Prediction Value	96.6%

Table 6: Comparative Analysis

Methods	Author and References	Computational Time
Rough Set Approach	Hassanien [1]	2'19"
Methematical Morphological	Konrad Bojar	2'50"
Shape and Texture Feature	Fahimeh Sadat Zakeri	8'21"
Proposed Method	S.Pitchumani Angayarkanni, Nadhira Banu Kamal	0'03"

REFERENCES

1. Hassanien,, Aboul Ella Ali and M.H. Jafar, 2001. Enhanced Rough Sets Rule Reduction Algorithm for Classification Digital Mammography, Journal of Intelligent Systems. 13: 151-171.
2. Konrad Bojar and Mariusz Nieniewski, 2008. Mathematical Morphology (MM) Features for Classification of Cancerous Masses in Mammograms, Information Technologies in BiomedicineAdvances in Soft Computing, Springer, 47: 129-138.

3. Fahimeh Sadat Zakeri, Hamid Behnam and Nasrin Ahmadinejad, 2012. Classification of Benign and Malignant Breast Masses Based on Shape and Texture Features in Sonography Journal of Medical Images Systems, 36: 1621-1627.
4. Aboul Elaa Hassanien and Amr Badr, A comparative Study on Digital A Mamography Enhancement Algorithms Based on Fuzzy Theory,Science in Informatics and Control, 12(1): 21-31.
5. Santle Camilus, K., V.K. Govindan and P.S. Sathidevi, 2010. Computer-Aided Identification of the Pectoral Muscle in Digitized Mammograms, J Digit Imaging, 23(5): 562-580.
6. Tingting Mu, Asoke K. Nandi and Rangaraj M. Rangayyan, 2008. Classification of Breast Masses Using Selected Shape, Edge-sharpness and Texture Features with Linear and Kernel-based Classifiers, J. Digit. Imaging, 21(2): 153-169.
7. Ertas, G., H.O. Gulcur, E. Aribal and A. Semiz, 2001. Feature Extraction From Mammographic Mass Shapes And Development Of A Mammogram Database,Engineering in Medicine and Biology Society, Proceedings of the 23rd Annual International Conference of the IEEE, 3: 2752 - 2755.
8. Habib Shah, Rozaida Ghazali and Nazri Mohd Nawi, 2011. Using Artificial Bee Colony Algorithm for MLP Training on Earthquake Time Series Data Prediction, Journal Of Computing, 3: 6.
9. Angayarkanni, S. Pitchumani, Nadira Banu Kamal and V. Thavavel, 2012. "Automatic detection and classification of cancerous masses in mammogram", 2012 Third International Conference on Computing Communication and Networking Technologies (ICCCNT 12).