Breast cancer detection in mammogram image with segmentation of tumour region

Vikramathithan Andiappan Chinnasamy*

Visvesvaraya Technological University,
Karnataka, India
Email: vikramathithanac0476@gmail.com
*Corresponding author

Dandinashivara Revanna Shashikumar

Department of Computer Science,
Cambridge Institute of Technology, India
and
Department of Computer Science,
Visvesvaraya Technological University (VTU), India
Email: drshashikumar@gmail.com

Abstract: In our proposed breast cancer malignant detection study are performed with the aid of fuzzy min max neural network technique. Majority of women's are affected in this breast cancer at a early stage the mammogram images are mostly play in a vital role. Initially the input mammogram image smoothened with the aid of adaptive median filer from that smoothened image we are segmenting tissues with the aid of Histon based fuzzy c-means clustering. We are extracting features from that segmented image the features are statistical and semantic features. Then we can identify the malignant region with the aid of these features. The segmented region is maligned or benign using an optimal fuzzy min max neural network with grey wolf optimisation algorithm with the aid of these we will identify a breast cancer region.

Keywords: breast cancer; mammogram; fuzzy min-max; grey wolf; optimisation; segmentation.

Reference to this paper should be made as follows: Chinnasamy, V.A. and Shashikumar, D.R. (2020) 'Breast cancer detection in mammogram image with segmentation of tumour region', *Int. J. Medical Engineering and Informatics*, Vol. 12, No. 1, pp.77–94.

Biographical notes: Vikramathithan Andiappan Chinnasamy obtained his Bachelor's in Electronics and Communication Engineering in 1998 and then he obtained his Master's in Microwave and Optical engineering in 2000 from Madurai Kamaraj University, Tamil Nadu, India. He is pursuing his PhD. in Visvesvaraya Technological University, Karnataka, India. Currently, he is working as Associate Professor in Department. of Electronics and Communication Engineering, Sai Vidya Institute of Technology, Bangalore, India. His specialisations are in the field of medical image processing and signal processing.

Dandinashivara Revanna Shashikumar received his BE degree from Mysore University and ME degree from Bangalore University, Bangalore and PhD in Information and Communication Technology of Fakir Mohan University, Balasore, Orissa. He is currently working as Professor and is the Head of the Department of Computer Science, Cambridge Institute of Technology, Visvesvaraya Technological University (VTU). His research interests include microprocessors, pattern recognition, and biometrics, computer networks, data mining and data warehouse. He has published 30 research publications in referred national and international journals. He is the reviewer for some of the international journals.

1 Introduction

In as per cancer frequency measurements in 2012, 14.1 million individuals were determined to have cancer worldwide and 8.2 million individuals kicked the bucket from cancer. In Turkey, 97.000 men, 62.000 ladies and general 159.000 individuals were determined to have cancer consistently (Turkey Public Health Agency). There are numerous variables to influence this, for example, way of life, nourishments, pollution, stress and indistinct things (El Saghir et al., 2011; Çalışkan et al., 2015). There are distinctive sorts of cancer. A standout amongst the most widely recognised malignancies and the principle reason of death in ladies matured 45 to 55 is breast cancer. In 2010, the American cancer society reported that breast cancer has been the most widely recognised kind of cancer amongst ladies (27%) and a noteworthy reason for disease passing among them. Metastatic breast cancer (MBC) is a hopeless infection, which also leads to lead death of the people (Rastghalam and Pourghassem, 2016; Nielsen et al., 2013). These high death rates are likely because of low mindfulness, numerous ladies being determined to have propelled infection, and absence of screening and demonstrative administrations (Kohler et al., 2017).

Breast cancer is the strange development of cells happening in breast tissue. The expression ‘breast cancer’ alludes to a dangerous tumour that either starts in the cells of the lobules, which is the milk creating organs, or the ducts, the entries that channel milk from the lobules to the areola. Less regularly, breast cancer can start in the stromal tissues, which incorporate the greasy and sinewy connective tissues of the breast. The most common breast cancer tumours express the estrogen (ER) and/or progesterone receptor (PgR) and the determination for endocrine treatment depends on the outflow of these hormone receptors (HR). For postmenopausal ladies with ER and/or PgR positive tumours, aromatase inhibitors are compelling and entrenched specialists both in the adjuvant and in the metastatic circumstance. Over expression or enhancement of the human epidermal growth factor receptor 2 (HER) is analyzed in around 15 to 20% of breast cancer patients. Radiological breast thickness has been connected with breast cancer hazard on an individual or familial premise (Huober et al., 2012; Ciatto et al., 2012; Arif et al., 2015). Thus the capacity to segregate and recognise rare cancer cells [i.e., circulating tumour cells (CTCs) in fringe blood offers awesome points of interest in growing new ways to deal with cancer analysis and prognostic forecast. A noteworthy test in CTC recuperation is their great irregularity in the circulatory system, evaluated to be around 1–10 CTCs for each 1 mL of entire blood in metastatic cancer patients. Another drawback is that CTCs display huge phenotypic heterogeneity (Hong et al.,

2016). In spite of the fact that the mechanical advances in restorative imaging promotes the accuracy of elucidation of pictures, the enhanced determination may not encourage the distinguishing proof of breast cancer at an early stage because of the numerous bewildering elements related, for case, to contrasts in instrument settings or breast locating by the administrator (Hong et al., 2016).

- Consequently it is watched that early recognition of danger can help in the determination of the sickness for lady and it can help emphatically to upgrade the anticipation of survival. For the location of breast cancer, different procedures are utilised. Sentinel lymph hub biopsy (SLNB) is the standard methodology for auxiliary arranging in patients with breast cancer worldwide and the reasonable pattern of breast cancer treatment is tending towards minimising auxiliary surgery, even within the sight of sentinel lymph nodes (SLN) association (Gentilini and Veronesi, 2012). Another strategy is the most encouraging mammography system and is utilised by radiologist much of the time. Mammogram pictures are more often of low differentiation and boisterous. The benefit of the subordinate advanced mammographic innovation is its capability to enhance breast imaging by giving a progression of pictures (slim cuts) through the breast which decreases the impact of tissue superimposition. Subsequently, this may permit better perception of danger (by decreasing tissue cover and the concealing impact of tissue thickness) and may likewise encourage separation between ordinary tissue structures and sores (Houssami and Skaane, 2013). But in mammography, ultrasound is viewed (Magna et al., 2016) as an integral technique for further appraisal of mammography suspicious discoveries. The fundamental reasons why ultrasound is not utilised for screening are the determination reliance on the specialist aptitudes, the long checking time and the trouble of imaging and enrolling the entire breast (González-Salido et al., 2016). Finally, all the shortcomings of mammography technique could be overcome by the presentation of dynamic contrast-enhanced magnetic resonance (DCE-MR) imaging in the field of breast imaging. This has enhanced the analysis of breast cancer with reported sensitivities and specificity of 95–99% and 80%, separately and has given urgent data in regards to typical breast tissue (Telegrafo et al., 2016; Singh and Gupta, 2015).

2 Literature survey

Microwave breast screening has been proposed as a reciprocal methodology to the present standard of X-beam mammography. Li et al. (2015) planned three ensemble classification forms which wired data from numerous sensors to recognise anomalies in the breast. A principled Neyman–Pearson methodology was produced to permit control of the exchange off between false positive rate and the false negative rate. They assessed execution utilising information got from estimations of heterogeneous breast phantoms. They likewise utilised information gathered in a clinical trial that checked 12 solid patients month to month over an eight month period. So as to evaluate the viability of the proposed algorithms they showed sweeps of breast with threatening injuries by falsely adding re-enacted tumour reactions to existing outputs of sound volunteers. Tumour reactions were developed in light of measured properties of breast tissues and genuine breast estimations, in that manner the re-enactment model considered the heterogeneity of

the breast tissue. The algorithms they showed exploited breast checks from different patients or tissue-mirroring breast phantoms to find out about breast content and what constituted a 'tumour less' and 'tumour-bearing' arrangement of estimations. They showed that the ensemble choice based, algorithm which developed an ensemble of the most useful classifiers, altogether outflanked other detection methods for the clinical trial information set.

Cai et al. (2016) introduced a signal amplification electrochemical aptasensor for the detection of breast cancer cell via free running DNA walker. Hypothetically, only one DNA walker, discharged by target cell-responsive response, could naturally divide all D-RNA (a chimeric DNA/RNA oligo nucleotide with a separation point rArU) tied down on cathode into smaller produced, offering ascended to significantly perceivable signal at long last. Under the ideal conditions, the electrochemical signal diminished directly with the grouping of MCF-7 cell. The direct range was from 0 to 500 cells mL 1 with a detection breaking point of 47 cells mL 1. In a word, that methodology might have preferences over conventional reported DNA machines for bioassay, especially regarding simplicity of operation, cost productivity, free of naming and of complex track plan, which might hold incredible potential for wide uses.

Circling miRNAs are rising as novel dependable biomarkers for early detection of cancer maladies. Through joining the upsides of electrochemical strategies and nano-materials with the selectivity of the oligo-hybridisation-based biosensors, Azimzadeh et al. (2016) proposed a novel electrochemical nano-biosensor for plasma miR-155 detection there, in view of thiolated test functionalised gold nano-rods (GNRs) finished on the graphene oxide (GO) sheet on the surface of the smooth carbon electrode (GCE). The decreased signs of a novel intercalating name oracet blue (OB) were measured by differential pulse voltammetry (DPV) strategy. The transmission electron microscope (TEM) imaging, UV-unmistakable spectrophotometry, cyclic voltammetry (CV), field emission scanning electron microscope (FE-SEM) imaging and energy dispersive spectroscopy (EDS) were demonstrated the right combination of the GNRs and right get together of the changed cathode. The electrochemical sign had a straight association with the centralisation of the objective miRNA running from 2.0 fM to 8.0 pM, and as far as possible was 0.6 fM. Moreover, the nano biosensor demonstrated high specificity, and could segregate forcefully between integral target miRNA, single-, three-base bungle, and non-corresponding miRNA. That nano biosensor had awesome capacity, reproducibility, and demonstrated a conventional reaction in the genuine example investigation with plasma. The proposed electrochemical nano biosensor could clinically be utilised as a part of the early detection of the breast cancer, by direct identification of the plasma miR-155 in genuine clinical specimens, without a requirement for test planning, RNA extraction and/or intensification.

To assess the execution of a mechanised computer-aided detection (CAD) framework to detect breast cancers that were ignored or confused in a breast MRI screening program for women at expanded danger. Gubern-Mérida, et al. (2016) distinguished 40 patients that were determined to have breast cancer in MRI and had an earlier MRI detection reported as negative accessible. In that earlier detection, 24 injuries could reflectively be recognised by two breast radiologists in agreement: 11 were scored as obvious and 13 as negligibly unmistakable. Furthermore, 120 ordinary outputs were gathered from 120 women without history of breast cancer or breast surgery taking an interest in the same MRI screening program. A completely automated CAD framework was connected to that dataset to distinguish dangerous sores.

Chen et al. (2015) produced a combined detection measure for assurance of cancer antigen 15-3 (CA15-3) and copper level in breast cancer serum by utilisation of gold nanorod (GNR)- based plasmonic sensor. Fast recognition of tumour marker (CA15-3) could be achieved using a basic transduction of optical sign actuated by safe response between CA15-3 and GNR@CA15-3 neutraliser. Then again, the copper induced change of restricted surface plasmon reverberation of GNRs resulted in an immediate readout of optical sign. In that manner, pragmatic application for combined assurance of CA15-3 and copper levels of human serum tests of breast cancer was performed. The outcomes were close concurrence with those given by the healing centre or other standard methodology, affirming that test with high exactness. The combined detection technique gave potential applications to early diagnosing and checking the advancement of dangerous tumour.

Breast cancer is known as the most widely recognised obtrusive cancer sort among women and programmed breast cancer detection frameworks are popular. Therefore, different machine learning and example acknowledgment strategies have been proposed to detect breast cancers. One of these methods is the Bayes classifier. Naïve Bayesian (NB) is known not a straightforward classifier, which depends on the Bayes hypothesis. There have been such a variety of utilisations utilised as a part of writing. Karabatak (2015) conveyed another NB (weighted NB) classifier and its application on breast cancer detection. A few examinations were directed to assess the execution of the weighted NB on the breast cancer database. The trials were acknowledged with 5-fold cross approval test. In addition, different execution assessment strategies specifically affectability, specificity and precision were considered. As indicated by the analyses, the weighted NB acquired the accompanying assessment values. The computed affectability, specificity and the exactness qualities were 99.11%, 98.25%, and 98.54% separately. In addition, an examination with the current techniques in the writing was displayed. Thus, the execution of weighted NB was superior to general NB and numerous other existing strategies.

Porous nano-rods of ceria (PN-Ceria), a novel ceria nanostructure with an extensive surface range and a high surface Ce_3þ part, showed solid inherent peroxidase action toward a traditional peroxidase substrate within the sight of H_2O_2. Peroxidase-like movement of ceria began from surface Ce3þ species as the synergist focus, accordingly clarifying the superior of PN-Ceria as a simulated compound copying peroxidase. Contrasted and the common protein horseradish peroxidase (HRP), PN-Ceria demonstrated a few preferences, for example, minimal effort, simple stockpiling, high affectability, and, unmistakably, compound and reactant security under cruel criteria. Critically, the enzymatic action of PN-Ceria remained about steady and stable over an extensive variety of temperature and pH values, guaranteeing the precision and dependability of estimations of its peroxidase-like action. Tian, et al. (2015) created a PN-Ceria based novel indicative framework for breast cancer detection with a higher sensitivity than the standard HRP discovery framework. Their work had established a strong framework for the advancement of PN-Ceria as a novel demonstrative instrument for clinical use.

3 Problem definition

There are certain drawbacks associated with the breast cancer detection in existing methods which are as follows:

1 In Li et al. (2015) there is an absence of an optimised classification algorithm which could make more efficient utilisation of a patient's own past scans to improve the performance.

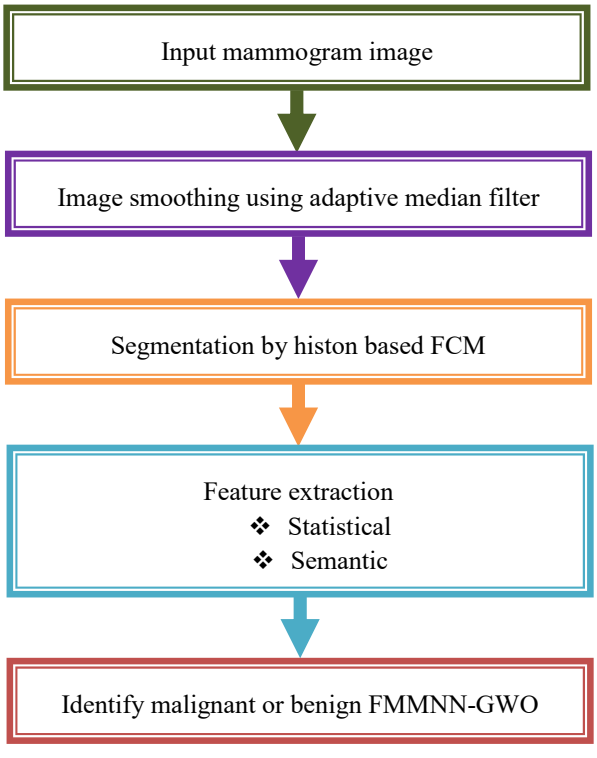
2 There is no combination of aptasensor with electrode chip in (Cai et al., 2016) and hence lacks possibility of promoting portable and affordable tools for cancer cell detection.

3 The usage of data from single clinical centre lead to small amount of prior negative MRI scan with visible lesions in retrospect. Possibility of reading faults. These serve as the major drawback of (Gubern-Mérida et al., 2016).

4 Grid search mechanism utilised in (Karabatak, 2015) for identifying the suitable weight values is more costly and further the initialisation of the weight vector is difficult and application dependent.

All these drawbacks could be easily rectified by using the proposed method.

Figure 1 Proposed tumour segmentation for heart disease prediction (see online version for colours)



4 Proposed methodology

Breast cancer is the most frequently diagnosed cancer in women. It is the second leading cause of cancer death in women after lung cancer. A means to cure breast cancer is not yet identified but treatment before it spreads to other parts of the body is possible on early detection. Moreover, detection of early stage tumour reduces the breast cancer mortality rate. Due to its widespread use, mammography plays an important role in the detection of breast cancer at an early stage. Several researchers worked in the area of breast cancer detection and proposed segmentation methods. Still, no clarification given by researchers is best hopeful and has limitations and it is still a difficult problem to solve. We introduce a simple and easy approach for detection of cancerous tissues in mammogram. In our proposed methodology we have intended to develop breast cancer detection with the aid of segmentation. Initially the input mammogram images are smoothing with the aid of adaptive median filtering technique. The smoothened images are given as the input for Segmentation process. The tissues are segmented by Histon based improved fuzzy C means clustering algorithm. Features such as Statistical, Semantic, etc are extracted from the segmented regions. This extensive feature set will be used to identify whether the segmented regions is malignant or benign using an Optimal fuzzy min-max neural network (OFMMNN). Here the optimisation can be done by grey wolf optimisation (GWO) algorithm. With the aid of the proposed approach, we can easily and effectively detect breast cancer at an early stage. The performance of the proposed system will be analyzed by TP, TN, FP, FN and accuracy and it will be compared with existing techniques.

4.1 Image smoothing using adaptive median filter

4.1.1 Adaptive median filter

In our proposed research we have utilise an adaptive median filter to smoothing an image in this algorithm is a betterment filtering to the medium one. Comparing with medium filter, it can deal with pulse noise in higher density. What's more, it can preserve more image details when dealing with non-pulse noises. We define the symbols as follows: Sx ,y is the template window of the centre pixel(x, y), $Zmin$ is the minimum grey in window Sx, y, $Zmax$ is the maximum grey in Sx, y, $Zmed$ is the medium grey in Sx, y, Zx, y is the grey value in pixel(x, y), $Smax$ is the tolerable.

Maximum scale of Sx, y. The adaptive medium filtering works in two steps:

Step A

$A1 = Zmed - Zmin \quad A2 = Zmax - Zmed$

If $A1 > 0$ and $A2 > 0$, go to step B

Else, enlarge the window.

If window $\leq Smax$, repeat step A

Else, output Zx, y.

Step B

$B1 = Zx, y - Zmin \quad B2 = Zmax - Zx, y$

If $B1 > 0$ and $B2 > 0$, Zx, y is output.

Else, $Zmed$ is output.

Step A is used to judge whether *Zmed* is a pulse. Step B is used to decide whether *Zx*, *y* is a pulse. If *Zmed* and *Zx*, *y* are both not pulse, then a constant value *Zx*, *y* is output to replace the medium value to avoid unnecessary losing of other detail information. In order to resolve the matter of the standard medium filtering on dealing with higher density pulse, the strategy of enlarging windows is adopted in the adaptive medium filtering to reduce the space density of pulse noises. As an improvement of the strategy, it can adopt the idea of the selecting templates, that is, reducing the space density of the pulse noises by change the shape and orientation of the window. An advantage of this method is that it can preserve more details by changing window's shape and orientation in two parts of neighbourhood and by avoiding the situation of one window including pixels in two different parts caused by enlarging window's size. If the image smoothing is completed then these images are given in to segmentation process here we utilise a histogram based fuzzy c-means clustering method to segmented the image.

4.2 Segmentation of modified Histon-based HADEWNNN algorithm

Histon-based ADEWNNN method was used in the paper for segment the multiple sclerosis, it can considerably simplify the segmentation computation.

In this Histon based segmentation was introduced by Mohabey and Ray. In this method introduced the new mean for visualisation of intensity information for the evaluation of similar colour regions in an image. Also in this method provide the segmentation of the elements in the boundaries, which can be applied in the process of image segmentation.

So it was greatly simplified the fuzzy segmentation. By using this segmentation method we get a good segmented result.

$$B(i, j) = \sum_{i=1}^{N} \sum_{j=1}^{M} A^*[C(i, j)]^x \tag{1}$$

In equation (1), where $X = \{x_1, x_2, \ldots\ldots\ldots, x_n\} \subseteq R_p$ it is a dataset of *n* represents the no of data, *M* represent the no of rows, *N* represent the no of column, *A* represent the constant variable as 2, *C* is a input images.

In the proposed modified Histon based fast fuzzy C-means are given below. Let *I* denotes the brain MR image, y_k is the observed log-transformed intensities at the *k-th* pixel and v_i is the log-trans- formed prototype of the *i-th* cluster. The objective function of the modified Histon based fast ADEWNNN (HFADEWNNN) proposed in our paper can be written as,

$$\begin{aligned} J_{FFCM} = & \sum_{i=1}^{c} \sum_{K=1}^{n} \left[(1-\alpha) u_{ik}^{m} \left(y_k - b_k - v_i \right)^2 + \alpha \sum_{r \in n} \mu_{ik}^{m} w_{kr} \left(y_r - b_r - v_i \right)^2 \right] \\ & + \sum_{k=1}^{n} a_k \sum_{i=1}^{c} u_k \left(1 - u_{ik}^{m-1} \right) \end{aligned} \tag{2}$$

where $U = u_{ik}^{m}$ is the fuzzy membership function, $V=\{v_1, \ldots\ldots, v_c\}$, is the intensity centre. n_k is a $q \times q$ neighbourhood centred at the *k-th* pixel, and $q = 2r + 1$ is the width of the window, r is the radius of the neighbourhood window. b_k is the bias field at the *k-th* pixel, ω_{kr} is the weight of the pixels in the neighbourhood centred as the *k-th* pixel. a_k is the control parameter of the term which rewards the crispness membership degrees. The algorithm is an iterative optimisation that minimises the objective function with the following constrains.

$$u_{ik} = \left[\sum_{i=1}^{c} \left(\frac{D_{ik} + \frac{\alpha}{N_k}\gamma_i}{D_{ik} + \frac{\alpha}{N_k}\gamma_j} \right)^{1/(m-1)} \right]^{-1} \tag{3}$$

$$v_i = \frac{\sum_{k=1}^{n} \left[\left(y_k - b_k + \frac{\alpha}{N_k} \sum_{y_r \in N_k} (y_r - b_r) \right) \right]}{(1+\alpha) \sum_{k=1}^{n} u_{ik}^{m}} \tag{4}$$

$$b_k = y_k - \frac{\sum_{i=1}^{c} u_{ik}^{m} v_i}{\sum_{i=1}^{c} u_{ik}^{m}} \tag{5}$$

By utilising this modified HFADEWNNN segmentation algorithm to effectively detect the multiple sclerosis from the brain MRI image.

4.3 Feature extraction

The extraction of image features constitutes one of the basic functions in the domain of cancer detection. In this regard, various categories of features are employed for the purpose of cancer detection which includes the Statistical features of pixel and the Semantic features. The features are mined from the segmented regions for locating the doubtful regions. In accordance with the texture, statistical and semantic features, the regions are classified.

4.3.1 Statistical features

The statistical features are evaluated by means of certain statistical function on the image pattern which include features such as the zoning, projection, profiling, histogram and distance, moments and so on. The structural and statistical features are complementary to each other and several other features can be obtained from the fundamentals of these features.

- *Area:* The simple form descriptor employed in the planned methodology is that the space the world of a selected image is calculated exploitation the expression,

$$A = \frac{I_h}{I_w} \tag{6}$$

I_h = image height, I_w = width

- *Mean:* is the mean of pixel in the image. The nth moment of mean is

$$\mu_0 = \sum_{i=0}^{L-1} (Z_i - m)^n P(Z_i) \tag{7}$$

where,

z is the grey level value

m is the mean value of z.

- *Covariance:* Covariance is a measure of how much two variables change together. The covariance between two real valued random variables X and Y with finite second moments is given by:

$$Cov(X, Y) = E[(X - E(X))\,(Y - E(Y))] \tag{8}$$

where E(X) is the expected value of X. The sample covariance of the K sets of N observations on the variables is the $K \times K$ matrix with $Q = [qjk]$ the entries given by

$$q_{jk} = \frac{1}{N-1} \sum_{j=1}^{n} (x_{ij} - x_j)(x_{ik} - x_k) \tag{9}$$

- *Correlation:* Correlation refers to any of a broad class of statistical relationships involving dependence. Here, correlation of the image is given as the covariance divided by the standard deviation.

$$\rho X, Y = q_{ik} / \sigma X \sigma Y \tag{11}$$

where $\rho X, Y$ is the correlation and σ is the standard deviation. After extracting the features of the each of the regions, the details are fed into the fuzzy min max neural network for training.

4.3.2 Semantic features

The semantic data symbolises a very significant factor in the co-reference resolution. The amalgamation of huge corpora and 'deep' analysis processes has facilitated the acquisition of a wide range of semantic data for the purpose of applying it to this task. These statistical and semantic features are given in to the next phase for optimisation here we utilise a fuzzy min max neural network is utilised to optimise a feature.

4.4 Fuzzy min-max neural network

In this heart disease prediction study we have utilised this FMMNN technique to classify the results. Fuzzy min-max learning is an expansion or contraction process. The training set consists of set of ordered pairs $\{A, I\}$, where $A = \{A_1, A_2 \ldots A_n\}$ is the input data and $I \in (1, 2, m)$ is the index of one of the class. The learning process begins by selecting an

ordered pair and finding a hyper box for the same class that can expand (if necessary) to include the input. If a hyper box cannot be found that meets the expansion criteria, a new hyper box is formed and added to the neural network. The membership function is defined with respect to the minimum and maximum points of a hyper box. It describes the degree to which a pattern fits in the hyper box. The hyper boxes have a range from 0 to 1 along each dimension. A pattern which is contained in the hyper box has a unity membership function. Mathematically, the definition of each hyper box fuzzy set H_j (Houssami and Skaane, 2013) is defined by,

$$H_j = \left\{ A, V_{\min j}, W_{\max j}, F\left(A, V_{\min j}, W_{\max j}\right) \right\} \tag{12}$$

where

A input data

$V_{minj}(V_{min1}, V_{min2}, \ldots. V_{minN})$ is the minimum points of H_j

$W_{maxj}(W_{max1}, W_{max2}, \ldots. W_{maxN})$ is the maximum points of H_j

$F(A, V_{mini}, W_{maxi})$ is the membership function

The membership function for the *j-th* hyper box (H_j) is given below,

$$H_j = \frac{1}{2n} \sum_{i=1}^{n} \left[\begin{array}{l} \max\left(0, 1 - \max\left(0, \gamma \min\left(1, A_j - w_{ij}\right)\right)\right) \\ + \max\left(0, 1 - \max\left(0, \gamma \min\left(1, v_{ji} - A_j\right)\right)\right) \end{array} \right] \tag{13}$$

where,

γ is a sensitivity parameter that regulates how fast the membership value decreases as the distance between A and H_j increases.

The architecture of fuzzy min max neural network consists of three layers of node. First layer represent the input layer that contains input data. Last layer represent the output layer that contain the number of classes. The middle or hidden layer is called hyper box layer. Here we include a binary cuckoo search algorithm in to a fuzzy min max neural network technique. These fuzzy min max neural network is used for identify a malignant in the tumour detected region.

4.5 Grey wolf optimisation

In this grey wolf optimisation technique is used for identify a malignant region. The grey wolves effectively enclose a Canidae's section ancestors and are valued as the head predators viewing their arrangement at the sustenance's food sequence. They normally illustrate a partiality to formulate suitable as an assembly The leader represent a male and a female, marked as alpha, which are for the majority division in accusation of enchanting proper variety screening diverse features, for exemplar, the hunting, sleeping location, time to wake, and so forth. The selections prepared by the alpha are accepted on to the assembly. The Beta addresses to the second grade in the pecking array of the grey wolves. They are, fundamentally, secondary wolves which effectively propose a few backing to the alpha in the variety constructing or equivalent assembly utility.

The omega is the smallest division of the grey wolf pack and huge task as a replacement present into the further foremost wolves very almost on every occasion and is allowed to include just the diminutive leftovers enchanting following a grand feast by the leader wolves. A wolf is indicated as secondary or as delta each so frequently in the occasion that it doesn't fit in among the congregation of an alpha, beta, or omega. Due to the truth that these delta wolves require to admiration the alphas and betas, they have a flourishing high proportion over the omegas. In our method, the alpha (α) is valued as the most appropriate collection by an outlook to replicating rationally the community pecking categorise of wolves whereas visualising the GWO. Therefore, the second and the third most excellent provision are beta (β) and delta (δ) independently. The left over hopeful provision are observed to be the omega (ω). In the GWO technique the hunting (optimisation) is conducted by the α, β, δ and ω.

- Initialisation process

  At this point we initialise the preprocessed output data as well a, A and C as coefficient vectors.

- Fitness evaluation

  Evaluate the fitness utility depend on the equation (14) and moreover select the finest result.

  $$Fit_i = \max accuracy \tag{14}$$

- Separate the solution based on the fitness

  Currently, we get the split result depend on the fitness value. Consider the initial finest fitness results be d_α, the second finest fitness results d_β and the third finest fitness results d_δ.

- Encircling prey

  The tracking is directed by α, β, δ and ω tag along these three contenders. In order for the collection to track a victim is foremost surrounding it.

  $$d(t+1) = d(t) + \vec{A}.\vec{K} \tag{15}$$

  $$\vec{K} = \left|\vec{C}.d(t+1) - d(t)\right| \tag{16}$$

  $$\vec{A} = 2\vec{a}r_1 - \vec{a} \text{ and } \vec{C} = 2r_2 \tag{17}$$

  Where, t represents the iteration n number.

  $d(t)$ represents the prey position.

  A and C represents the coefficient vector.

  $\vec{a}$ is linearly decreased from 2 to 0.

  r_1 and r_2 represents the random vector [0, 1].

- Hunting

  We suppose that the alpha (best candidate solution), beta and delta include the enhanced information about the probable position of the victim in order to imitate

accurately the tracking activities of the grey wolves. Since an outcome, we accumulate the earliest three finest result accomplished up to now and necessitate the further explore mediator (including the omegas) to alter their pose based on the arrangement of the finest explore mediator. For replication, the novel result is predictable by the formulae stated beneath.

$$\vec{K}^{\alpha} = \left|\vec{C}_1.d_{\alpha} - d\right|, \vec{K}^{\beta} = \left|\vec{C}_2.d_{\beta} - d\right|, \vec{K}^{\delta} = \left|\vec{C}_3.d_{\delta} - d\right| \tag{18}$$

$$d_1 = d_{\alpha} - \vec{A}_1.\left(\vec{K}^{\alpha}\right), d_2 = d_{\beta} - \vec{A}_2.\left(\vec{K}^{\beta}\right), d_3 = d_{\delta} - \vec{A}_3.\left(\vec{K}^{\delta}\right) \tag{19}$$

$$d(t+1) = \frac{d_1 + d_2 + d_3}{3} \tag{20}$$

It can be experimental that the concluding location would be in a casual position surrounded by a circle which is distinct by the location of alpha, beta, and delta in the explore gap. By means alpha, beta, and delta guess the location of the victim, and further wolves' modernise their location erratically in the region of the victim.

- Attacking prey (exploitation) and search for prey (exploration)

  Investigation and utilisation are assured by the adaptive values of *a* and *A*. The adaptive values of limitation *a* and *A* permit GWO to efficiently conversion among investigation and utilisation. By declining *A*, half of the iterations are constant to investigation ($|A| \geq 1$) and the further half are committed to utilisation ($|A| < 1$). The GWO contains only two foremost limitations to be accustomed (*a* and *C*). Though, we encompass reserved the GWO algorithm as effortless as achievable through the smallest number of operative to be accustomed The procedure will be persistent awaiting the greatest exactness is acquired. At last the finest attribute are elected and supply to the additional procedure.

5 Result and discussion

Our proposed FMMHH-GWO for the effective segmentation of multiple sclerosis is implemented using the MATLAB platform on the brain MRI images from the dataset. The segmentation figures of benign and malignant cancers are as shown in Figures 2–3.

Figure 2 Images of benign cancer

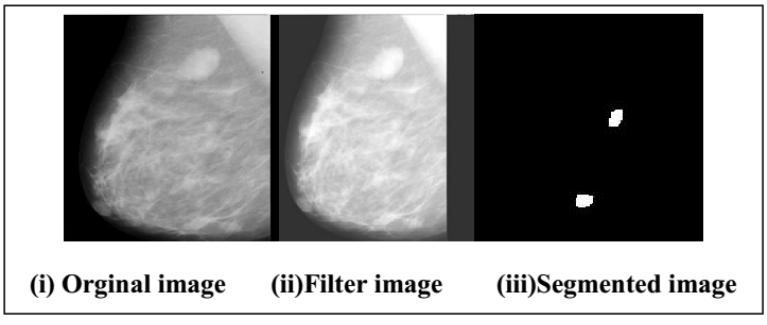
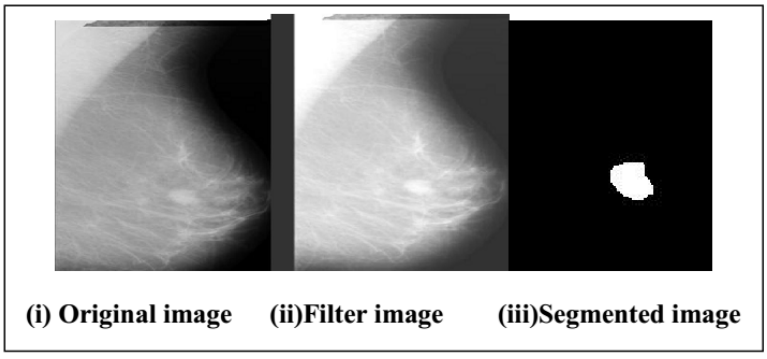


Figure 3 Images of malignant cancer



5.1 Evaluation metrics

By using the evaluation metrics sensitivity, specificity and accuracy, the performance of the system is evaluated. The normal count values such as true positive (TP), true negative (TN), false positive (FP) and false negative (FN) are utilised here.

- Sensitivity

  The proportion of actual positives which are correctly identified is the measure of the sensitivity. It relates to the ability of test to identify positive results.

$$Sensitivity = \frac{Number\ of\ true\ positives}{Number\ of\ true\ positives + Number\ of\ false\ negatives} \times 100$$

- Specificity

  The proportion of negatives which are correctly identified is the measure of the specificity. It relates to the ability of test to identify negative results.

$$Specificity = \frac{Number\ of\ true\ negatives}{Number\ of\ true\ negatives + Number\ of\ false\ positives} \times 100$$

- Accuracy

  We can compute the measure of accuracy from the measures of sensitivity and specificity as specified below.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100$$

5.2 Performance evaluation

The performance of proposed method is evaluated by using accuracy, sensitivity and specificity values. Table 1 indicates the parameter values for different images.

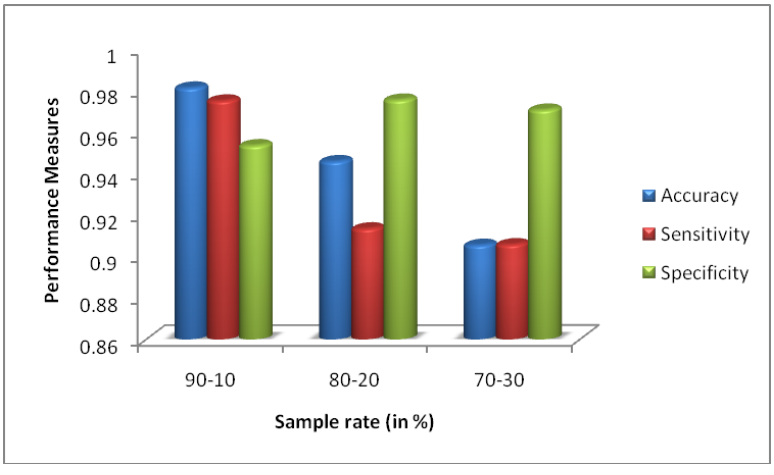
Utilising the proposed method in the training samples in ratio 90–10, the average accuracy value is 0.956521739% and the sensitivity value is 0.995% and the specificity value is 0.893442623%. When we decrease the training sample to 80% and testing

sample to 20% means, the accuracy value is 0.905791937%, sensitivity value is 0.913375856% and the specificity value is 0.97540405%. The accuracy value is 0.875468815, the sensitivity value is 0.905735525% and the specificity value is 0.970592782% for 70-30 samples ratio.

Table 1 Performance measures of the proposed technique

Training samples ratio (in %)	*Accuracy*	*Sensitivity*	*Specificity*
90–10	0.956521739	0.995	0.893442623
80–20	0.905791937	0.913375856	0.97540405
70–30	0.875468815	0.905735525	0.970592782

Figure 4 Graph for sensitivity, specificity and accuracy results of segmentation (see online version for colours)



From Figure 4, it is clear that for all the images the accuracy value is automatically decreased when the number of training samples decreased. Here the highest accuracy obtained in 90-–10 ratio.

5.3 Comparative analysis

For the classification of normal and abnormal brain images our proposed work made use of FMMNN-GWO classifier. From the below table we could say that our proposed method has better accuracy on comparing to the existing method adaptive differential evolution wavelet neural network (ADEWNN) classifier (Singh et al., 2016) and MLP based CNN segmentation method (Rouhi et al., 2015). The comparison outcomes could be better understood using the following Table 2.

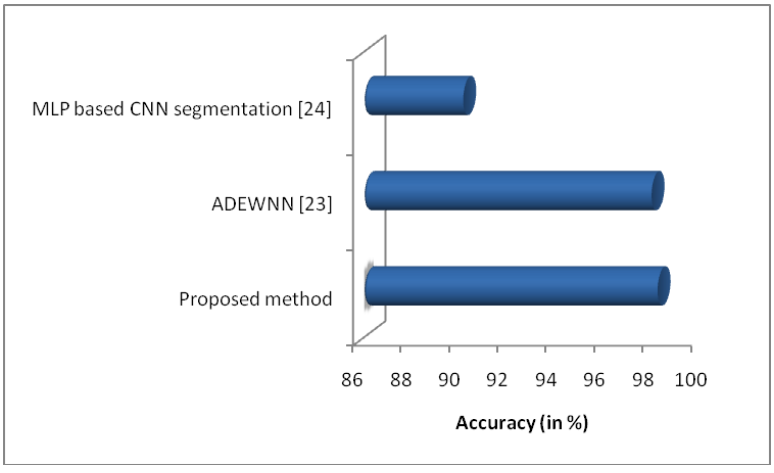
From Table 2 it is clear that the accuracy value is 98.21% for the proposed method while utilising the existing classifier ADEWNNN (Singh et al., 2016) the accuracy value is 97.96, because in our proposed method, we are combining fuzzy and neural network which is optimised by GWO. So it will classify the images based on optimal features. Our proposed method is further compared with MLP based CNN segmentation method,

which has 92% accuracy. Figure 5 represents the accuracy comparison of proposed methods with existing study.

Table 2 Comparative analysis

Methods	*Classifier used*	*Sensitivity*	*Specificity*	*Accuracy*
Proposed method	Fuzzy min-max neural network	98.45	97. 32	98.21
ADEWNN (Singh et al., 2016)	Wavelet NN	98.19	97.19	97.96
MLP based CNN segmentation (Rouhi et al., 2015)	MLP	92.70	90.54	90.16

Figure 5 Accuracy comparison with existing methods (see online version for colours)



From the results and discussion, we observed that the proposed method outperforms the existing research in terms of accuracy.

6 Conclusions

The Histon based fuzzy c-means clustering based breast cancer detection four phases such as image smoothing, segmentation, feature extraction, optimisation. Initially the input mammogram images are smoothing with the aid of adaptive median filtering technique. The smoothened images are given as the input for segmentation process. The tissues are segmented by Histon based improved fuzzy C means clustering algorithm. Features such as statistical, semantic, etc., are extracted from the segmented regions. This extensive feature set will be used to identify whether the segmented regions is malignant or benign using an optimal fuzzy min-max neural network (OFMMNN). Here the optimisation can be done by GWO algorithm. With the aid of the proposed approach, we can easily and effectively detect breast cancer at an early stage. The efficiency of the optimisation of images is very high by presenting very good accuracy outcomes and also the segmentation of input mammogram image very accurate outcomes. From the outcomes, we have showed that the fuzzy min-max neural network-GWO utilised in our

proposed work outperforms the other classifiers ADEWNNN by facilitated very good accuracy of 98.21% in categorising the images into normal and abnormal.

References

Arif, S., Qudsia, S., Urooj, S., Chaudry, N., Arshad, A. and Andleeb, S. (2015) 'Blueprint of quartz crystal microbalance biosensor for early detection of breast cancer through salivary autoantibodies against ATP6AP1', *Biosensors and Bioelectronics*, Vol. 65, pp.62–70, Elsevier.

Azimzadeh, M., Rahaie, M., Nasirizadeh, N., Ashtari, K., and Naderi-Manesh, H. (2016) 'An electrochemical nanobiosensor for plasma miRNA-155, based on graphene oxide and gold nanorod, for early detection of breast cancer', Biosensors and Bioelectronics, Vol. 77, pp.99–106, Elsevier.

Cai, S., Chen, M., Liu, M., He, W., Liu, Z., Wu, D. and Chen, J. (2016) 'A signal amplification electrochemical aptasensor for the detection of breast cancer cell via free-running DNA walker', Elsevier on Biosensors and Bioelectronics, Vol. 85, pp.184–189.

Çalışkan, R., Gültekin, S.S., Uzer, D. and Dündar, Ö. (2015) 'A microstrip patch antenna design for breast cancer detection', *Procedia-Social and Behavioral Sciences*, Vol. 195, pp.2905–2911, Elsevier.

Chen, S., Zhao, Q., Zhang, L., Wang, L., Zeng, Y. and Huang, H. (2015) 'Combined detection of breast cancer biomarkers based on plasmonic sensor of gold nanorods', *Sensors and Actuators B: Chemical*, Vol. 221, pp.1391–1397, Elsevier.

Ciatto, S., Bernardi, D., Calabrese, M., Durando, M., Gentilini, M.A., Mariscotti, G. and Stevanin, C. (2012) 'A first evaluation of breast radiological density assessment by QUANTRA software as compared to visual classification', *The Breast*, Vol. 21, No. 4, pp.503–506, Elsevier.

El Saghir, N.S., Adebamowo, C.A., Anderson, B.O., Carlson, R.W., Bird, P.A., Corbex, M., Harness, J.K. et al. (2011) 'Breast cancer management in low resource countries (LRCs): consensus statement from the Breast Health Global Initiative', *The Breast*, Vol. 20, pp.S3–S11, Elsevier.

Gentilini, O. and Veronesi, U. (2012) 'Abandoning sentinel lymph node biopsy in early breast cancer? A new trial in progress at the European Institute of Oncology of Milan (SOUND: sentinel node vs. observation after auxiliary ultra sound)', *The Breast*, Vol. 21, No. 5, pp.678–681, Elsevier.

González-Salido, N., Medina, L. and Camacho, J. (2016) 'Full angle spatial compound of ARFI images for breast cancer detection', *Full Angle Spatial Compound of ARFI Images for Breast Cancer Detection*, Vol. 71, pp.1–32, Elsevier.

Gubern-Mérida, A., Vreemann, S., Martí, R., Melendez, J., Lardenoije, S., Mann, R.M., Platel, B. et al. (2016) 'Automated detection of breast cancer in false-negative screening MRI studies from women at increased risk', *European Journal of Radiology*, Vol. 85, No. 2, pp.472–479, Elsevier.

Hong, W., Lee, S., Chang, H.J., Lee, E.S. and Cho, Y. (2016) 'Multifunctional magnetic nanowires: a novel breakthrough for ultrasensitive detection and isolation of rare cancer cells from non-metastatic early breast cancer patients using small volumes of blood', *Biomaterials*, pp.78–86, Elsevier.

Houssami, N. and Skaane, P. (2013) 'Overview of the evidence on digital breast tomosynthesis in breast cancer detection', *The Breast*, Vol. 22, No. 2, pp.101–108, Elsevier.

Huober, J., Fasching, P.A., Barsoum, M., Petruzelka, L., Wallwiener, D., Thomssen, C., Kubista, E. et al. (2012) 'Higher efficacy of letrozole in combination with trastuzumab compared to letrozole monotherapy as first-line treatment in patients with HER2-positive, hormone-receptor-positive metastatic breast cancer–results of the eLEcTRA trial', *The Breast*, Vol. 21, No. 1 pp.27–33, Elsevier.

Karabatak, M. (2015) 'A new classifier for breast cancer detection based on naïve Bayesian', *Measurement*, Vol. 72, pp.32–36, Elsevier.

Kohler, R.E., Gopal, S., Miller, A.R., Lee, C.N., Reeve, B.B., Weiner, B.J. and Wheeler, S.B. (2017) 'A framework for improving early detection of breast cancer in Sub-Saharan Africa: a qualitative study of help-seeking behaviors among Malawian women', *Patient Education and Counselling*, Vol. 100, No. 1, pp.1–7, Elsevier.

Li, Y., Santorelli, A., Laforest, O. and Coates, M. (2015) 'Cost-sensitive ensemble classifiers for microwave breast cancer detection', in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP*, pp.952–956.

Magna, G., Casti, P., Jayaraman, S.V., Salmeri, M., Mencattini, A., Martinelli, E. and Di Natale, C. (2016) 'Identification of mammography anomalies for breast cancer detection by an ensemble of classification models based on artificial immune system', *Knowledge-Based Systems*, Vol. 101, pp.60–70, Elsevier.

Nielsen, D.L., Kümler, I., Palshof, J.A. and Andersson, M. (2013) 'Efficacy of HER2-targeted therapy in metastatic breast cancer. Monoclonal antibodies and tyrosine kinase inhibitors', *The Breast*, Vol. 22, No. 1, pp.1–12, Elsevier.

Rastghalam, R. and Pourghassem, H. (2016) 'Breast cancer detection using MRF-based probable texture feature and decision-level fusion-based classification using HMM on thermography images', *Pattern Recognition*, Vol. 51, pp.176–186, Elsevier.

Rouhi, R., Jafari, M., Kasaei, S. and Keshavarzian, P. (2015) 'Benign and malignant breast tumors classification based on region growing and CNN segmentation', *Expert Systems with Applications*, Vol. 42, No. 3, pp.990–1002.

Singh, A.K. and Gupta, B. (2015) 'A novel approach for breast cancer detection and segmentation in a mammogram', *Procedia Computer Science*, Vol. 54, pp.676–682, Elsevier.

Singh, S.P., Urooj, S. and Lay-Ekuakille, A. (2016) 'Breast cancer detection using PCPCET and ADEWNN: a geometric invariant approach to medical X-ray image sensors', *IEEE SENSORS JOURNAL*, Vol. 16, No. 12, pp.4847–4855.

Telegrafo, M., Rella, L., Ianora, A.S., Angelelli, G. and Moschetta, M. (2016) 'Effect of background parenchymal enhancement on breast cancer detection with magnetic resonance imaging', *Diagnostic and Interventional Imaging*, Vol. 97, No. 3, pp.315–320, Elsevier.

Tian, Z., Li, J., Zhang, Z., Gao, W., Zhou, X. and Qu, Y. (2015) 'Highly sensitive and robust peroxidase-like activity of porous nanorods of ceria and their application for breast cancer detection', *Biomaterials*, Vol. 59, pp.116–124, Elsevier.