ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2022.0373

AN AUTOMATIC DETECTION OF BREAST CANCER USING EFFICIENT FEATURE EXTRACTION AND OPTIMIZED CLASSIFIER MODEL

S. Vani Kumari[1] and K. Usha Rani[2]
[1]*Department of Computer Applications, Government College for Women, Srikakulam, India*
[2]*Department of Computer Science, Sri Padmavati Mahila Visvavidyalayam, India*

Abstract

Image Processing Techniques (IPTs) are widely used in healthcare. IPTs work on images to extract useful information from them. IPTs are effective in the early diagnosis of cancers which helps in their proper treatment. The number of human beings affected by cancer has drastically increased recently and more in women in the form of Breast Cancers (BCs). This work is on detecting BCs from image datasets using IPTs. This work proposes the use of Weiner Filters (WFs) in Preprocessing BC image data. WFs are applied to images as they can improve the quality of images. Image segmentations are executed using the Watershed Algorithm which detects Regions of Interest (ROIs) in images. Feature Extractions are done using the statistical method Gray Level Co-occurrence Matrix (GLCM). These extracted features are then selected using Black Widow Optimization (BWO) Algorithm. The selected features are then trained by Fuzzy Neural Networks (FNNs) which classify BC cases. The results of the classifier when evaluated were found to have a better performance on medical images in terms of Accuracy.

Keywords:

Breast Cancer Detection, Weiner Filter, Feature Extraction, GLCM, Feature Selection, Black Widow Optimization (BWO) Algorithm, Classification, Fuzzy Neural Network (FNN)

1. INTRODUCTION

BCs are the worst type of cancer affecting women mainly due to their higher fatality. One important aspect of reducing these cancer deaths is by early diagnosis of BCs. More than 90% of BCs are curable when detected early and can save the lives of BC affected women [1]. Appropriate categorization of BCs like benign (localized/non-invasive) or malignant (invasive and dangerous) can help in diagnosis. Multitudes of medical modalities exist in healthcare that image body parts. Mammograms, Computerized Tomography (CT)/ Magnetic Resonance Imaging (MRI) scans and biopsy images (histo-pathological) are commonly used for identifying BCs. When BC is identified from a mammogram image, biopsy imaging follows in diagnosis [2]. Analysis of images using systems generally consists of software and hardware and can be viewed as two subsystems encompassing tissue preparation/image production and image analysis. Thus, mortality rates due to BCs in women can be reduced by proper screening and effective diagnosis.

BC medical screening and detections include ultrasonic imaging, MRI/CT scans, and molybdenum target X imaging [3] where Molybdenum Mammography Images (MMIs) have reduced costs, and the convenience of operations, and low levels of harm in patients. Automating BC detections or using CAD (Computer Aided Diagnosis) can help radiologists hasten their diagnostics while improving the accuracy of BC detections. The potential of relevant features found in MMIs [4] makes it usable in MLTs (Machine learning Techniques) though controversies exist in identifying benign or malignant BCs using IPTs. BC diagnostics are based on visual aspects of breast tissues and step by step evaluation including preparing specimens, selecting cell nuclei of tumours, and IPT followed by classification.

Segmentation algorithms working on images lookout for two basic characteristics namely discontinuity and similarity in the image's luminance [5]. In [6] a new classification system based on enhanced neural networks and correlated-contours fuzzy rules (EC-FR) is used. By using an optimization approach known as a bat, more significant aspects from the input samples are extracted. An effective framework is built-in according to fuzzy rules extraction, with similarity-based directional components of data partitioning and cloud data creation used in the second layer. Wavelet functions are used to compute the weight and bias values of neurons are used in the system. These algorithms have been used widely in IPTs processing various types of imaging areas including healthcare, remote sensing and traffic congestions. The history of segmentation algorithmic usage dates back to the 1920s. Traditional image segmentation techniques select regions based on thresholds, areas and edges [7]. Traditional algorithms face issues in feature extractions due to uncertainties and fuzziness in images. To overcome such hurdles and execute proper segmentation of images this work uses a watershed algorithm in image segmentation.

Feature extraction is the simplification of resources that accurately depict large datasets which may include many variables thus increasing the complexity of processing [9]. To eliminate the irrelevant neurons from the neuro-fuzzy network model, density-based regularization approaches and activation functions are presented. The density of the input data determines the formation of efficient fuzzy logical neurons for classification. The technique for generating random weights, combined with more efficient methods for generating the final stage weights at the network's output layer, allows for faster network training. To reduce the complexity and search space, the fuzzy rules' parameters are fine-tuned using Hybrid Ant Colony Particle Swarm Optimization (HASO) [10].

Moreover, as the number of features/variables increases, the memory requirements also increase in parallel leading to overfitting of training samples and poor generalization of samples. Feature extraction methods; construct combinations of variables to determine the accuracy of describing data. Images have textures. Texture analysis finds unique representations of underlying image textures for simpler representations which can be used in better and more accurate classifications or use of images. Despite its dimensionality of pattern recognitions, very few architectures texture-based feature extractions [11]. This work extra the acts image's textural features using a statistical method of GLCM. Though many features can be extracted using GLCM this works uses second-order features: correlation,

entropy, angular second moment and inverse difference moment as these features provide high levels of di and specifically in images with motions. With the aid of parallel and inseparably fuzzy rules, a suitable hybrid optimization problem is posed. To evaluate fuzzy sets, a novel shape able membership feature with adaptive shape is used to outline contours with different shapes. Following that, the derived fuzzy rules parameters are fine-tuned using the hybrid optimization approach [14].

Various MLTs have been used in healthcare CAD applications for their high-performances in detecting objects of clinical significance from medical images in detecting tumours [12]. Hence, CAD systems and mammograms can identify suspicious regions in images with lesser false positives for classifying detected masses. GAs (Genetic Algorithms), and Rule-based fuzzy classifiers have been generally used in Image classifications. These techniques have two major problems [13] in mammographic image feature extractions namely inadequacy of information and design of very few networks in terms of suitability. This work attempts to overcome these issues while using these images for BC detections.

This introductory section is followed by literature on BC detections. The proposed methodology is detailed in section 3 followed by implementation results in the next section. This work concludes in section 5 with future scope.

2. LITERATURE REVIEW

This sect details several studies that have proposed schemes in the area of mammographic image classifications and BC detections along with their advantages or limitations.

Bayes theorem was used by Tyrer et al. [15] in their study where low penetrance and BRCA genes along with personal risks were considered. A woman's family history was exploited iteratively to assess the presence of predisposing genes that can lead to the development of BC. The study refined the risks of BC further using women's personal history and thus provided personalized risk estimates.

ANNs (Artificial Neural Networks were used by Baker et al. [16] for identifying benign/malignant lesions in the breast. They constructed their networks based on standard lexicons from BI-RADS (Breast Imaging Recording and Data System) of the American College of Radiology. The networks were fed with 18 inputs including8 from the patient's personal history and 10 BI-RADS lesion descriptors. Data from 133 benign and 73 malignant cases were used for training and testing ANNs. The scheme's specified output threshold showed the proposed networks improved biopsy PPVs (Positive Predictive Values) from 35 to 61 per cent with a relative sensitivity of 100%. When the sensitivity was fixed to 95%, the results showed 62% which was greater than radiologists' specificity of 30% at $P < .01$. Thus, their use of BI-RADS lexicons provided a standard language for mammographers while ANNs improved PPV breast biopsy values.

SVMs (Support Vector Machines) and their ensembles were assessed by Huang et al. [17] in their study where voluminous as well as small BC datasets were considered. They aimed at proving that single classifiers can outperform single SVMs in BC detections. The study evaluated accuracies, ROCs, F-measures of SVMs and SVM ensemble combinations. LK-SVMs (Linear kernel-based SVMs) using bagging and SVM ensemble combinations using RBF kernel and boosting were suitable for predicting BC from miniature datasets. The study also implied that feature selections were the base for SVM based classifiers to improve their accuracy of predictions.

Wisconsin breast cancer dataset (WBCD) was used by Kumari et al. [18] to formulate a system for predicting BC incidences early. The study was analyzed by selecting a minimum set of attributes from clinical datasets. Their experimental results showed that their proposed scheme achieved classification accuracy comparable to pathologist predicted values. The schema obtained a maximum classification accuracy of 99.28%.

Gene expression profiling was used in the study of Reis-Filho et al. [19] who showed that ER (Estrogen-Receptor) positive and negative BCs are distinct diseases at the transcriptomic level and that other molecular sub-types might exist within these groups. They also proved that ER-positive conditions were framed largely due to the proliferation of related genes. Their molecular classification system and prognostic multi-gene classifiers based on microarrays and their derivatives have been tested in clinical trials and practices.

Multiple LRs (Logistic regressions) were used by Fabian et al. [20] in their study. Their schema used LRs with Cox hazards analysis for predicting cancers from evid of hyperplasia with atypia and a 10-year Gail projected probability of developing BC. Even though epidermal growths, ERs, p53, and HER2/NEU are statistically important in determining hyperplasia with atypia, their scheme multivariable analysis could not predict the development of Bence. Gail's risk model could be used to identify legions using Cytomorphology's randomized periareolar aspirates of the breast and project a high short-term risk of evolving into BC. The study also recommended that cytomorphology could be used as a BC prevention procedure.

MLTs were used by Asri et al. [21] where their scheme proposed multiple MLTs including SVMs, DT (Decision Tree) C4.5, NB (Naive Bayes) and kNNs (k-Nearest Neighbors) with experiments the on the WBCD dataset. The study aimed at assessing the effectiveness/efficiency of data classifications of SVMs. The study used WEKA where SVM's accuracy scores were 97.13% with reduced errors.

The study by Amrane et al. [22] compared the performances of NB and KNN in BC detections. Their implemented evaluation of the classifiers in terms of accuracy using cross-validation showed that the use of KNN could result in high accuracies (97.51%) and low error rates when compared to NB classifier accuracy of 96.19%.

The study by Turgut et al. [23] evaluated MLTs with and without the use of feature selections and compared the final outputs. The study used MTs including SVMs, KNNs, MLPs (Multi-Layer Perceptrons) DTs, RF (Random Forest), LRs, Adaboost and GB (Gradient Boost) in their evaluations. The scheme applied two different feature selection methods for retrieving the 50 best features. Their evaluations showed that SVMs produced the best results. MLPs examined multiple layers and neurons for classification accuracy and found that increasing layers either decreased accuracy or the accuracy did not change. Classification is a technique related to categorization and the method of identifying, distinguishing, and interpreting ideas and artefacts, and it will be useful in several emerging fields. A variety

of machine learning techniques and as well as the creation of suitable algorithms are explained [24].

Sharma et al. [25] also compared MLTs in BC detections. The MLTs considered in the study included RF, KNNs and NB which were then trained and tested on the WDBC data set. Their comparison of MLT performances was in terms of key parameters of accuracy, and precision. Their study showed that these classifiers could be effectively used for BC detections and hence BC treatments.

A novel method for accurate BC detections was proposed by Alarabeyyat et al. [26]. Their method worked in two phases. IPTs formed the first part which prepared mammography images in pattern extraction/ feature selection processes. In the second phase, the extracted features were fed into BPNNs (Back Propagation Neural Networks) and LRs. The resultant outcome or predictions of the study proved that LRs needed more features than BPNNs for predictions of BC.

However, the afore discussed methods have gaps in extracting specific features from mammographic images and their accuracy levels are limited by lesser samples in training. Hence, this work focuses on feature extraction/selection methods which can ultimately improve classifier performances.

3. PROPOSED METHODOLOGY

This work introduces a new scheme for detecting BC cells by its proposed feature extraction/selection technique from mammographic images which can subsequently help classifiers increase their classification accuracy.

1) Initially the breast cancer data is considered as an input dataset, in which Preprocessing was done using WF applied for BC data as it improved image quality.
2) The Segmentations in this work are based on the Watershed Algorithm which segments regions of interest in images.
3) GLCM is used in Feature Extraction from segmented parts of interest.
4) Feature selections are executed using the bio-inspired BWO Algorithm.
5) This work's selected features are evaluated using FNNs on BC classifications.

The overall process of the proposed detection of breast cancer is illustrated in Fig.1.

3.1 IMAGE PRE-PROCESSING

Image Pre-processing is a significant step in IPT as it removes noises or impurities present in images. A well-known image pre-processing technique is the WF (Weiner Filter) which was applied to BC data. The use of WF enhanced the quality of images and is explained below:

- **WF:** Images are prone to noise and degrade IPT performances and hence their quality can be restored using filtering which can be expressed mathematically as follows:

$$g(x,y)=f(x,y)*u(x,y)+n(x,y) \quad (1)$$

$$h(x,y)=R[g(x,y)] \quad (2)$$

where, $f(x,y)$ - an acquired image, $u\ x,y$ - degradation function, * - convolution, $n(x,y)$ - noises like Gaussian noise, $g(x,y)$ - output degraded image, and $h(x,y)$ - final output image on applications of the technique R. The study obtained de-noised gamma images by feeding them to a conventional noise reduction filter like WF which obtains denoised images in non-linear spatial domains [27]. This work improved image quality as detailed below:

Initially, a mask of the $n\times m$ matrix is set in this work for filtering spatial noises. This matrix replaces pixel values with the median value of the neighbourhood in degraded parts of images. Moreover, WF eliminates sudden changes in pixel values and retains the sharpness information of images. A pixel's variance and median used by the mask matrix can be defined as equations given below:

$$\mu = \frac{1}{MN}\sum_{n,m\in\eta} a(n,m) \quad (3)$$

$$\sigma = \frac{1}{MN}\sum_{n,m\in\eta} a^2(n,m) - \mu^2 \quad (4)$$

where, μ - mean, σ^2 - Gaussian noise variance in the image, $n\times m$ - neighbourhood area size η in the mask, and an (n,m) - pixels in η. WF outputs new pixels represented by $b_w(n,m)$ using estimated values.

$$b_w(n,m) = \mu + \frac{\sigma^2 + v^2}{\sigma^2}\cdot\left(a(n,m) - \mu\right) \quad (5)$$

where v^2 - noise variance setting of the mask matrix for application of WF.

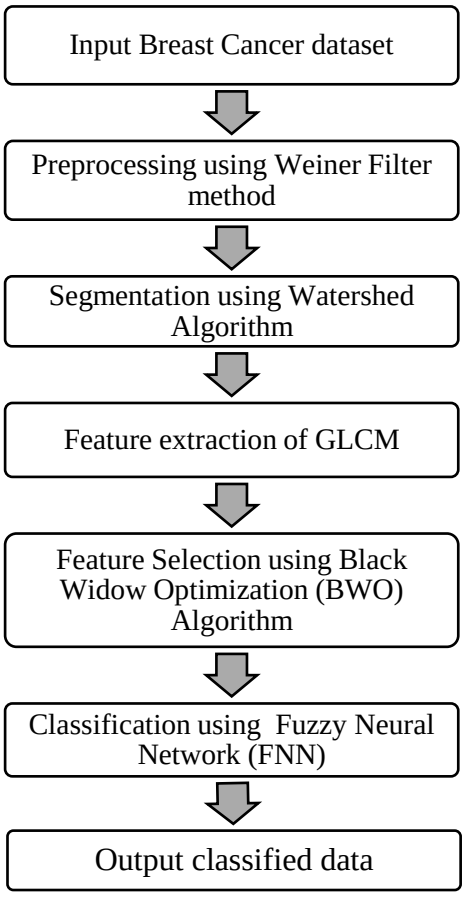

Fig.1. Overall Flow of the Proposed Methodology

3.2 WATERSHED ALGORITHM-BASED SEGMENTATION

Segmentation of images is dividing digital images into multiple segments (sets of pixels). The process aims to simplify images representations as different meaningful segments which can be analyzed [28]. Segmentations highlight objects and boundaries in images while labelling image pixels based on similarities of certain visual characteristics. Thus, seg results in fragments of disjoint homogenous regions which collectively represent the image where the segmented regions carry objects of interest. Pixel homogeneity in segmented image parts is measured by their intensity.

3.2.1 Watershed Method:

Watershed (WS) algorithm is a region-based technique. It can be viewed as landscapes under floods where watersheds divide regions of rain i.e. land submerged in water with pores (local minima). The pores/holes are the starting points for water filling and converge at a point and create a flow where dams are built. This process ends when water levels reach their peaks, dividing land into regions(basins) distinguished by watersheds.

WS is an image transformation/segmentation method based mathematical morphology of pixel's greyscale values. In the case of colour image pixels, it is a multi-component image. Though multiple strategies have been presented, this work pays special focus on the bit mixing approach where multiple dimensions are transformed into a single dimension in a feature space. When an image is viewed from a geological angle, WS lines determine image region boundaries. Topographically, in image *i*, the grey tone values of pixels can be evolved. WS transforms identify basins (Image regions) and ridgelines (Boundaries). When the WS algorithm is used to simulate image segmentation, the basins are first found and then taken as a fixed complement, the image is then partitioned into basins, and finally, the boundaries are detected. The flow of WS is depicted in Fig.2.

3.2.2 Seed Region Growing (SRG) Algorithm

SRG operations are based on collecting pixels with similar properties in forming regions and are based on a few steps: It starts with a seed pixel (starting point) for segmentation of regions; Pixels with similar properties determined by a formula are merged (growing) or pixels around the seed pixel is merged into its domain.

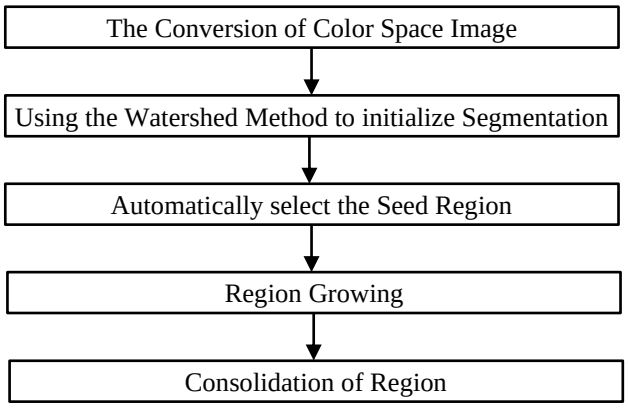

Fig.2. WS algorithm Flowchart

This process is continued with new seed pixels until no more pixels satisfy the inclusion condition. SRG method needs to answer three important questions:

- Determining seed pixel for representing required region;
- Using a formula to accommodate adjacent pixels in region growths;
- Define conditions for terminations of the seed growth process.

The dependency of pixels in SRG can be First order dependence or second-order dependence when pixels have similar difference ratios in their vicinity.

3.3 FEATURE EXTRACTION USING GLCM

Feature extractions and feature selections are techniques for reducing dimensionality in Image Processing and computer vision where extractions create new features by transforming original features irreversibly [28] but with useful information required for the targeted aim. Feature selection on the other hand eliminates unnecessary or unimportant features for further processing.

GLCM examines image textural features statistically where spatial relationships found between pixels are assessed. GLCM computations are based on this relationship which is then used for extractions. In GLCM, the co-occurrence matrix of miniature image regions is calculated for obtaining statistical values including Correlations, Contrasts, Uniformity, Homogeneities, Probabilities, Inverses and Entropies. GLCM converts distances and angles between pixels in multiple angular views. Fourteen image textural features worthy of extractions were defined by Haralick. This work considers four features from Haralick's features namely energies or ASMs (Angular Second Moments), Inertia, Correlations, Entropies, and Inverse Difference Moments for implementations.

3.3.1 Angular Second Moment (ASM):

The Energy feature is the total of the entity's squares also called Uniformity. The picture homogeneity is measured by ASM values. When an image has strong homogeneity or when pixels are very close, the ASM values are high.

$$ASM = \sum_{i=0}^{N_{g-1}} \sum_{j=0}^{N_{g-1}} P_{ij}^2 \tag{6}$$

3.3.2 Inverse Difference Moment (IDM):

This feature indicates Local homogeneity i.e., an image local grey levels are uniform while the inverse values in GLCM are high. It is also the opposite of Contrast weights where weights are defined as follows:

$$IDM = \frac{\sum_{i=0}^{N_{g-1}} \sum_{j=0}^{N_{g-1}} P_{ij}}{1+(i-j)^2} \tag{7}$$

3.3.3 Entropy:

The amount of compression applied to an image is given by its entropy value. This metric indicates loss of information in transmissions given by:

$$Entropy = \sum_{i=0}^{N_{g-1}} \sum_{j=0}^{N_{g-1}} -P_{ij} \log P_{ij} \tag{8}$$

3.3.4 Correlation:

Correlations can measure linear dependency between pixels. Digital Image Correlations (DICs) are optical methods which use tracking and image registrations to identify changes in 2D/3D images. DICs are applied generally in science and engineering to computing deformations or displacements or pressure or optical flows where optical mouse movements are a popular example

$$Correlation = \frac{\sum_{i=0}^{N_{g-1}} \sum_{j=0}^{N_{g-1}} (i,j)\, p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (9)$$

The above-described image features are extracted using Mat lab where GLCM is computed from the image input.

3.4 FEATURE SELECTION USING BWO

The feature selection process involves selecting a subset of attributes from a larger collection of attributes to recognize important features and eliminate features that don't apply to the learning mission, and for creating a good learning model. The advantages of feature selection are twofold: it reduces the induction algorithm's computation time and improved the accuracy of the resulting model. This study introduces BWO, a new metaheuristic optimization algorithm that mimics the unusual mating behaviour of black widow spiders [31]. When compared to other approaches, BWO has key differences in terms of its quick convergences and circumventing local optima issues while exploring or exploiting feature spaces. BWO also balances its exploitations and discoveries i.e. it can inspect wider areas for attaining the best global solutions and thus making it an apt method for optimizations amidst multiple local optimums.

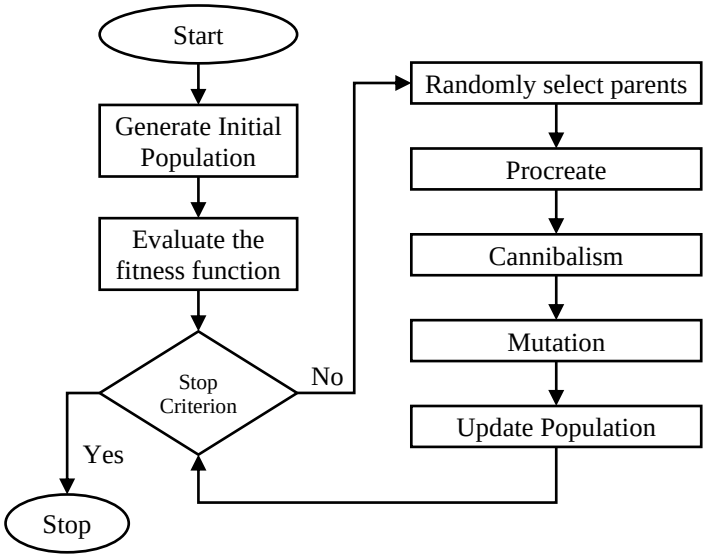

Fig.3. BWO Flowchart

BWO is based on spider population where each spider is a solution. Spiders generate new generations in pairs after mating where the male is eaten by the female spider. The female spider then transports the sperm to an egg sac from where they are released as Spiderlings within 11 days. These Spiderlings' cannibalism can be observed when they cohabitate with their maternal webs before being swept away by a storm.

3.4.1 Initial Population:

A GA chromosome is an acceptable framework for solving an optimization problem which is referred to as a ''widow" in BWO. BWO's solutions are Black widow spiders which contain the problem variables. BWO considers these problem variables as an array for solutions.

In a N_{var} - dimensional optimization problem, a widow is an array of $1 \times N_{var}$ representing the solution of the problem. This array is defined as follows:

$$Widow = \left[x_1, x_2, ..., x_{N_{var}} \right] \quad (10)$$

Each of the variable values (x_1,x_2,…x_{Nvar}) is a floating-point number. The fitness of a widow is obtained by evaluation of fitness function f at a widow of (x_1,x_2,…x_{Nvar}). So

$$Fitness = f\left(widow\right) = f\left(x_1, x_2, ..., x_{N_{var}}\right) \quad (11)$$

A candidate widow matrix of size $N_{pop} \times N_{var}$ is generated from an initial population of spiders for optimizations. Pairs of parents are randomly chosen for mating after which the male spider ceases to exist.

3.4.2 Procreate:

The self-contained pairs begin mating for new offspring generations where pairs mate in their networks. Real spider mating results in 1000 eggs approximately where stronger babies survive. The algorithmic replication is executed by generating an array called alpha. The array has random numbers where offsprings are produced using Eq.(12), in which x_1 and x_2 represent parents while y_1 and y_2 represent children.

$$y_1 = \alpha \times x_1 + (1-\alpha) \times x_2;\ y_2 = \alpha \times x_2 + (1-\alpha) \times x_1 \quad (12)$$

This process is repeated $0.5N_{var}$ times where random numbers that are chosen do not repeat. The parents and their children are put into an array which is sorted on fitness values. Individuals with good cannibalism ranking are added to the newly created population.

3.4.3 Cannibalism:

Cannibalism can be represented in three different ways. Sexual when female spiders eat their husbands during mating or later. BWO differentiates between male and female fitness. The second form is sibling cannibalism where stronger children eat weaker ones. BWO uses Cannibalism Ratings (CRs) to assess survivors. Cannibalism's third form occurs when children eat their mothers which are assessed by the strong fitness values of children. The Fig.3 depicts the proposed algorithm's flowchart.

3.4.4 Mutation:

Select a Mutepop number of individuals from the population at random in this step. As shown in Fig.4, each of the chosen solutions swaps two elements in the array at random. The mutation rate is used to measure mutepop.

3.4.5 Convergence:

Three stop conditions can be considered similar to other evolutionary algorithms: (a) the count of iterations is predetermined. (b) Observation on no changes in the best widow's fitness value over many iterations. (c) Achieving the necessary degree of precision.

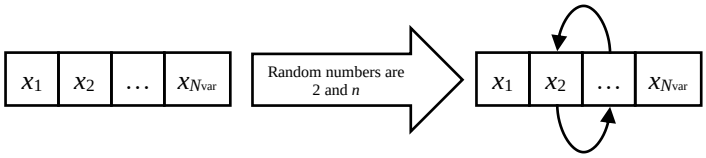

Fig.4. Process of Mutation

3.5 CLASSIFICATION USING FNNS

Computational intelligence systems use methods such as Neural Networks (NNs), Fuzzy Logics (FLs), evolutionary algorithms, multi-agent approaches, and rule-based systems, among others. With a greater understanding of the relationships between various formulation and process parameters, artificial intelligence is increasingly being used in pharmaceutical technology. NNs and FLs are rapidly evolving technologies that could be used in pharmaceutical product formulation and processing [32]. For predicting and optimizing formulation conditions, evolutionary algorithms are used in conjunction with ANNs.

The FLs link NNs to first-order logic. The FNNs are three-layered FFNs where the input layer is a fuzzification layer, a hidden layer with fuzzy rules, and the output layer is a defuzzification layer. The layers are connected by fuzzy sets. Input layers are fuzzy rules-based membership functions which trigger hidden layer rules. Weights between layers represent fuzzy sets where memberships in sets are determined by their relative weights which can be modified in training. Continuous transfer functions move real values in the network to the output layer, where these values construe membership degrees in fuzzy sets based on the firing of fuzzy rules of the hidden layer.

The layered Feed Forward Neural Network (FFNN) with a Back Propagation (BP) least mean-square learning algorithm is commonly used in NNs. The processing units known as neurons are linked by the network edges. Each neuron input has a weight associated with it, which represents its relative significance in the neuron's input collection. A weighted linear combination sum of neuron inputs is the final output. Parameter hierarchy evaluates the pattern as an assessment in multi-criteria analysis. A hierarchical neural network can be used to encode the hierarchy, with each neuron representing a criterion. The network's input neurons adhere to a single criterion. Complex parameters are represented by the secret and output neurons. When the parameters are considered to be distinct, they can be combined linearly.

In reality, standards are linked to failing linear evaluation functions which do not determine the relationships. This work overcomes this issue by its use of Standard Back Propagations (SBPs).

This work uses Fuzzy Back Propagations (FBPs) as fuzzy extensions where net values are computed using LR style fuzzy numbers and thus does fix or assume parameter's independence. FBP are also advantageous in overcoming local minima. Thus, for single and multiple-training patterns, required FBP convergence conditions for single outputs are defined.

3.5.1 FBP Algorithm:

Many Neuro-fuzzy models have recently been introduced for generating the net value net *i* by aggregating the input values of the i^{th} neuron. Sugeno's fuzzy integral, which is based on psychological context, is used to mathematically explain the mapping:

Step 1: Randomly generate the initial weight sets w for the input hidden layer where each $w_{ji}=(w_{mji},w_{\alpha ji},w_{\beta ji})$ is an LR type fuzzy number. Also, generate the weight set for the hidden output layer, where $w'_{kj}=\left(w'_{mkj},w'_{\alpha kj},w'_{\beta kj}\right)$;

$w'_{ij}=\left(w'_{mji},w'_{\alpha ji},w'_{\beta ji}\right)$; $w'_{kj}=\left(w'_{mkj},w'_{\alpha kj},w'_{\beta kj}\right)$.

Step 2: Let (I_p,D_p) p=1,2,…,N input-output pattern set, that fuzzy backpropagation needs to be trained with. Here $I_p=(I_{p0},I_{p1},I_{p1})$ where each I_{pi} is an LR-type fuzzy number.

Step 3: Assign values for α and η; *Alpha* = 0.1; *Neta* = 0.9

Step 4: Get next pattern set (I_p,D_p) Assign $(O_{pi}=I_{pi}, i=1,2,3\ldots)$

Step 5: Compute the input to hidden neurons

$$O'_{pj}=f\left(NET_{pj}\right), j=1,2\ldots.m; O'_{p0}=1$$

where $NET_{pj}=CE(\sum W_{ji}O_{pi})$

Step 6: Compute the hidden to output neurons

$$O''_{pj}=f\left(NET'_{pj}\right), k=1,2\ldots.n;$$

where $NET_{pj}=CE(\sum W_{ji}O'_{pi})$

Step 7: Compute change of weights $\Delta w'(t)$ for the hidden output layer as follows Compute

$$\Delta E_p(t)=(\partial E_p/\partial \mathrm{w}'_{mkj}\ ,\partial E_p/\partial w_{\alpha kj},\partial E_p/\partial \mathrm{w}'_{\beta kj})$$

Compute $\Delta w'(t)=-\eta\Delta E_p(t)+\alpha\Delta w'(t\text{-}1)$

The update weight *i* of hidden to output neuron is

$$W'(t)=W'(t\text{-}1)+\Delta W'(t)$$

Step 8: Compute change of the weights $\Delta w'(t)$ for the input hidden layer as follows

$$\text{Let, } \delta_{pmk}=-\left(D_{pk}-O''_{pk}\right)O''_{pk}\left(1-O''_{pk}\right)\cdot 1$$

$$\delta_{pmk}=-\left(D_{pk}-O''_{pk}\right)O''_{pk}\left(1-O''_{pk}\right)\cdot\left(-\frac{1}{3}\right)$$

$$\delta_{pmk}=-\left(D_{pk}-O''_{pk}\right)O''_{pk}\left(1-O''_{pk}\right)\cdot\left(\frac{1}{3}\right)$$

Compute $\Delta E_p(t)=(\partial E_p/\partial \mathrm{w}'_{mkj}\ ,\partial E_p/\partial w_{\alpha kj},\partial E_p/\partial \mathrm{w}'_{\beta kj})$

Compute $\Delta w'(t)=-\eta\Delta E_p(t)+\alpha\Delta w'(t\text{-}1)$

Step 9: Update weight for the input-hidden-output layer as

$$W'(t)=W(t\text{-}1)+\Delta W(t)$$

$$W'(t)=W'(t\text{-}1)+\Delta W'(t)$$

Step 10: $p=p+1$; if $(p\le N)$ go to step 5

Step 11: COUNT_of_ITRNS=COUNT_OF_ITRNS+1;

if COUNT_of_ITRNS<ITRNS

{

Reset pointer of the first pattern in the training set;

P=1;

Go to step 5;

}

Step 12: Output w' and w'', the final weight sets.

4. RESULTS AND DISCUSSION

The results of experimental studies combining the FNN Classification algorithm with feature extraction algorithms are presented in this section. The data sets for defining BC from the WDBC dataset were derived from the images used for the experimental study. In the experiments, several parameters used in binary classification were used to evaluate the utility of the various methods in predicting BC outcomes.

To determine different performance metrics, first test the True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) rates. The first performance metric was precision, which was specified as the proportion of relevant retrieved instances. The recall metric, defined as the proportion of relevant instances that were retrieved, was the second output metric.

The measures of precision and recall are both critical in evaluating the success of a prediction strategy, despite their sometimes-contradictory existence. As a result, these two metrics can be combined with equal weights to produce the F-measure, a single metric. The accuracy metric, defined as the proportion of correctly predicted instances relative to all predicted instances, was the final output metric.

Precision is defined as the ratio of correctly found positive observations to all of the expected positive observations.

$$\text{Precision} = TP/(TP+FP) \quad (13)$$

Sensitivity or Recall has defined the ratio of correctly identified positive observations to the overall observations.

$$\text{Recall} = TP/(TP+FN) \quad (14)$$

F - measure is defined as the weighted average of precision as well as Recall. As a result, it takes false positives and false negatives.

$$\text{F1 Score} = 2*(Recall * Precision) / (Recall+Precision) \quad (15)$$

Accuracy is calculated in terms of positives and negatives as follows:

$$\text{Accuracy} = (TP+FP)/(TP+TN+FP+FN) \quad (16)$$

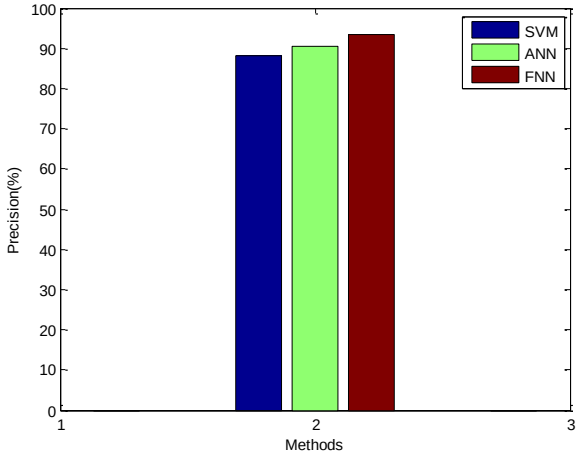

Fig.5. Precision comparison results between the proposed and existing method for classifying the heart disease data

The Fig.5 illustrates the precision comparison results between the proposed and existing method for classifying the breast cancer data. From the results, it concludes that the proposed FNN technique has high precision results compared to the existing classification techniques.

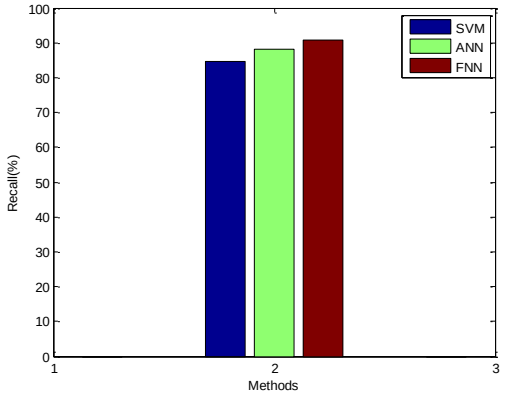

Fig.6. Recall comparison results between the proposed and existing method for classifying the heart disease data

The Fig.6 illustrates the recall comparison results between the proposed and existing method for classifying the breast cancer data. From the results, it concludes that the proposed FNN technique has high recall results compared to the existing classification techniques.

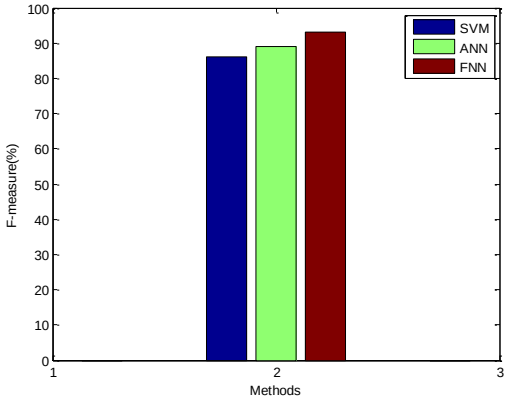

Fig.7. F-measure comparison results between the proposed and existing method for classifying the heart disease data

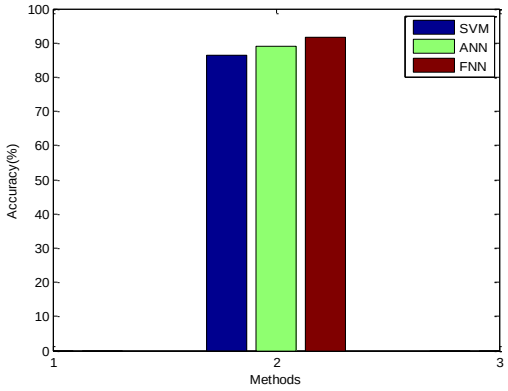

Fig.8. Accuracy comparison results between the proposed and existing method for classifying the heart disease data

The Fig.7 illustrates that the F-measure comparison results between the proposed and existing method for classifying the breast cancer data. From the results, it concludes that the proposed

FNN technique has high F-measure results compared to the existing classification techniques.

The Fig.8 illustrates that the accuracy comparison results between the proposed and existing method for classifying the breast cancer data. From the results, it concludes that the proposed FNN technique has high accuracy results compared to the existing classification techniques.

5. CONCLUSION

The proposed FNN model of this research work was used to establish an automated detection of BC. Initially, the BC data is treated as an input dataset, and WF, a well-known preprocessing technique, is used to treat noises in the data. The well-known WS Algorithm is used to find and segment part of the interest from the image during the Segmentation process. The well-known method GLCM is used to extract features from the segmented part of the interest in the Feature Extraction process. The proposed BWO-based feature selection approach performs exceptionally well in terms of locating real global optima with high accuracy and speed. This study uses FNNs to identify BC data while evaluating its feature extractions. The proposed work's accuracy achieves a classification accuracy of 91.66%.

REFERENCES

[1] M.M. Islam, M.R. Haque and M.N. Kabir, "Breast Cancer Prediction: A Comparative Study using Machine Learning Techniques", *SN Computer Science*, Vol. 1, No. 5, pp. 1-14, 2020.

[2] P.H. Abreu, M.S. Santos, M.H. Abreu and D.C. Silva, "Predicting Breast Cancer Recurrence using Machine Learning Techniques: A Systematic Review", *ACM Computing Surveys*, Vol. 49, No. 3, pp. 1-40, 2016.

[3] J.A. Cruz and D.S. Wishart, "Applications of Machine Learning in Cancer Prediction and Prognosis", *Cancer Informatics*, Vol. 2, pp. 1-13, 2006.

[4] B.M. Gayathri, C.P. Sumathi and T. Santhanam, "Breast Cancer Diagnosis using Machine Learning Algorithms-A Survey", *International Journal of Distributed and Parallel Systems*, Vol. 4, No. 3, pp. 105-114, 2013.

[5] S. Sharma and T. Choudhury, "Breast Cancer Detection using Machine Learning Algorithms", *Proceedings of International Conference on Computational Techniques, Electronics and Mechanical Systems*, pp. 114-118, 2018.

[6] P. Patro, K. Kumar and G. Suresh Kumar, Similarity and Wavelet Transform based Data Partitioning and Parameter Learning for Fuzzy Neural Network", *Journal of King Saud University-Computer and Information Sciences*, Vol. 78, No. 3, pp. 1-17, 2020.

[7] K. Anastraj, T. Chakravarthy and T. Poondi, "Breast Cancer Detection Either Benign or Malignant Tumor using Deep Convolutional Neural Network with Machine Learning Techniques", *Proceedings of International Conference on Computational Techniques, Electronics and Mechanical Systems*, pp. 566-573, 2018.

[8] H.S. Hota, "Diagnosis of Breast Cancer using Intelligent Techniques", *International Journal of Emerging Science and Engineering*, Vol. 1, No. 3, pp. 45-53, 2013.

[9] S. Singh, S. Saini and M. Singh, "Cancer Detection using Adaptive Neural Network", *International Journal of Advancements in Research and Technology*, Vol. 1, No. 4, pp. 1-5, 2012.

[10] P. Patro, K. Kumar and G. Suresh Kumar, "Neuro Fuzzy System with Hybrid Ant Colony Particle Swarm Optimization (HASO) and Robust Activation", *Journal of Advanced Research in Dynamical and Control Systems*, Vol. 12, No. 3, pp. 741-750, 2020.

[11] T. Eklund, M. Collan, P. Jalava and Y. Collan, "On Visual Exploration of Breast Cancer Data using the Self-Organizing Map", *Proceedings of International Workshop on Self-Organizing Maps*, pp. 139-146, 2005.

[12] M.M. Beg and M. Jain, "An Analysis of the Methods Employed for Breast Cancer Diagnosis", *Proceedings of International Conference on Machine Learning*, pp. 1-13, 2012.

[13] A. Raad, A. Kalakech and M. Ayache, "Breast Cancer Classification using Neural Network Approach: MLP and RBF", *Wireless Networks*, Vol. 7, No. 8, pp. 1-9, 2012.

[14] P. Patro, K. Kumar and G. Suresh Kumar, "Optimized Hybridization of Ant Colony Optimization and Genetic Algorithm (HACOGA) Based IC-FNN Classifier for Abalone", *Journal of Computational and Theoretical Nanoscience*, Vol. 17, pp.2756-2764, 2020.

[15] J. Tyrer, S.W. Duffy and J. Cuzick, "A Breast Cancer Prediction Model Incorporating Familial and Personal Risk Factors", *Statistics in Medicine*, Vol. 23, No. 7, pp. 1111-1130, 2004.

[16] J.A. Baker and C.E. Floyd, "Breast Cancer: Prediction with Artificial Neural Network based on BI-RADS Standardized Lexicon", *Radiology*, Vol. 196, No. 3, pp. 817-822, 1995.

[17] M.W. Huang, C.W. Chen and C.F. Tsai, "SVM and SVM Ensembles in Breast Cancer Prediction", *PloS One*, Vol. 12, No. 1, pp. 1-13, 2017.

[18] M. Kumari and V. Singh, "Breast Cancer Prediction System", *Procedia Computer Science*, Vol. 132, pp. 371-376, 2018.

[19] J.S. Reis Filho and L. Pusztai, "Gene Expression Profiling in Breast Cancer: Classification, Prognostication, and Prediction", *The Lancet*, Vol. 378, pp. 1812-1823, 2011.

[20] C.J. Fabian, S. Zeiger and M.S. Mayo, "Short-Term Breast Cancer Prediction by Random Periareolar Fine-Needle Aspiration Cytology and the Gail Risk Model", *Journal of the National Cancer Institute*, Vol. 92, No. 15, pp. 1217-1227, 2000.

[21] H. Asri, H. Mousannif, H. Al Moatassime and T. Noel, "Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis", *Procedia Computer Science*, Vol. 83, pp. 1064-1069, 2016.

[22] M. Amrane, S. Oukid, I. Gagaoua and T. Ensari, "Breast Cancer Classification using Machine Learning", *Proceedings of International Conference on Electric Electronics*, pp. 1-4, 2018.

[23] S. Turgut and T. Ensari, "Microarray Breast Cancer Data Classification using Machine Learning Methods", *Proceedings of International Conference on Electric Electronics, Computer Science, Biomedical Engineering*, p. 1-3, 2018.

[24] P. Patro, K. Kumar and G. Suresh Kumar, "Various Classifiers Performance Based Machine Learning Methods", *International Journal of Recent Technology and Engineering*, Vol. 8, No. 3, pp. 305-310, 2019.

[25] S. Sharma, A. Aggarwal and T. Choudhury, "Breast Cancer Detection using Machine Learning Algorithms", *Proceedings of International Conference on Computational Techniques, Electronics and Mechanical Systems*, pp. 114-118, 2018.

[26] A. Alarabeyyat and M. Alhanahnah, "Breast Cancer Detection using K-Nearest Neighbor Machine Learning Algorithm", *Proceedings of International Conference on Developments in Systems Engineering*, pp. 35-39. 2016.

[27] R.J. Hemalatha and R. Chandrasekaran, "Active Contour based Segmentation Techniques for Medical Image Analysis", *Medical and Biological Image Analysis*, Vol. 4, pp. 1-17, 2018.

[28] C.R. Park, S.H. Kang and Y. Lee, "Median Modified Wiener Filter for Improving the Image Quality of Gamma Camera Images", *Nuclear Engineering and Technology*, Vol. 52, No. 10, pp. 2328-2333, 2020.

[29] T. Srikanth, P.P. Kumar and A. Kumar, "Color Image Segmentation using Watershed Algorithm", *International Journal of Computer Science and Information Technologies*, Vol. 2, No. 5, pp. 2332-2334, 2011.

[30] P. Mohanaiah, P. Sathyanarayana and L. GuruKumar, "Image Texture Feature Extraction using GLCM Approach", *International Journal of Scientific and Research Publications*, Vol. 3, No. 5, pp. 1-5, 2013.

[31] V. Hayyolalam and A.A.P. Kazem, "Black Widow Optimization Algorithm: A Novel Meta-Heuristic Approach for Solving Engineering Optimization Problems", *Engineering Applications of Artificial Intelligence*, Vol. 87, pp. 1-15, 2020.

[32] M. Velmurugan and K. Thangavel, "Mammogram Classification using Fuzzy Neural Network", *International Journal of Computational Intelligence System*, Vol. 3, No. 3, pp. 1-12, 2013.