

# ALOA: Ant Lion Optimization Algorithm-based Deep Learning for Breast Cancer Classification

**Jack Russel Fernandis**

*Leitek Informations, Schwerin, Germany*  
*russelfernandis33@gmail.com*

**Abstract:** Normally, breast cancer is the most familiar disease, and it is the second leading type of cancer death among women. Breast cancer disease begins in malignant stage and it slowly develops from breast cell. Periodic clinical checks and self-tests assist for early prediction of breast cancer, thus the survival rate can increase significantly. In this paper, Ant Lion Optimization (ALO)-based Deep Neural Network (DNN) technique is developed for breast cancer classification. Here, the Type 2 Fuzzy and Cuckoo Search (T2FCS) method is utilized for pre-processing the input image. The color-based thresholding scheme is employed for segmenting the blood cells from the input breast cancer image. Furthermore, the statistical features are extracted from an input image for predicting breast cancer. Besides, the cancer classification is performed using a DNN classifier. The ALO algorithm is developed in this method to train the classifier for better performance. In addition, the developed breast cancer classification approach achieves a better classification performance based on several parameters, such as specificity, accuracy, and sensitivity. Thus, the developed ALO-based DNN method obtained enhanced results with high accuracy of 95.34%, sensitivity of 97.54%, and specificity of 95.34%.

**Keywords:** Breast Cancer, Ant Lion Optimization Algorithm, Deep Neural Network, Color-Based Thresholding, CNN Features.

## 1. Introduction

Cancer is a more dangerous disease and it directs to a huge amount of deaths worldwide. There are more than a hundred types of cancers are available, among them breast cancer is the most familiar cancer in women especially for aged women, which means above 40 years [1,5]. Normally, Breast cancer is a type of malevolent development beginning with breast tissue typically, in the inner layer of milk ducts or breast lobules and metastasizing to other body divisions [2,6]. In worldwide, the second more extensive disease is breast cancer. Hence, it is essential to verify the number of deaths because of breast cancer in victorious and dropped treatment [2]. Moreover, there are numerous classifications utilized in breast cancer grading system, familiar one is T-Tumor, N-lymph node and M-metastasis (TNM) system. Generally, TNM classifies the breast tumor into five levels, such as stage 0, stage 1, stage 2, and stage 4. Here, stage 0 is a pre-cancerous situation, like Ductal Carcinoma in situ (DCIS) and Lobular Carcinoma in situ (LCIS). Similarly, stage 1 and 2 is persistent tumors in which tumors are limited with breast and sentinel lymph nodes. Stage 3 indicates that cancer extended outside to the instant tumor area and affected closed lymph muscles and nodes, while stage 4 is metastatic cancer where cancer tissue is increased away from the breast and nearby lymph nodes [3]. Commonly, breast cancer is divided into two types, namely malignant and benign. Benign is non-cancerous, whereas malignant is a cancerous form of a tumor. In addition, both tumors have more sub-kinds, which require analyzing independently since every kind may need various types of predictions and treatment strategies.

Meanwhile, early prediction leads to several advantages, like the survival capacity of women with breast cancers. Breast cancer can be avoided or alleviated through efficient defensive treatments based on long-term clinical experience [1,7]. Thus, introducing the efficient prediction approach is most significant for breast cancer. Besides, several classifiers are developed in breast cancer prediction system, like Artificial Neural Network (ANN) [8] [1] [28], Rotation Forest (RF) [9] [1], Support Vector Machine (SVM) [1,10] and Particle Swarm Optimization technique-based Classifier (PS-classifier) [1]. Additionally, the Deep Active Learning Framework (DALF) was developed for breast cancer detection. Moreover, the Human Epidermal growth factor Receptor-2 (Her2Net) classifier was also devised for the detection and segmentation process of cell nuclei and membranes in the breast cancer classification [11]

[2], along with this, Computer Assisted Diagnosis (CAD) system was also introduced for cancer disease prediction and classification. The CAD system was developed for balancing the detection of the pathologist and avoids inters pathologist variation for enhancing the cancer classification [12,13]. Moreover, GoogLeNet system-based classification techniques [14,15] were developed for detecting cell samples based on their microscopic high magnification multi-views. Additionally, a regularization approach, termed as transition module was introduced, and it is helpful to design generalization capability and a gradual decrease in size of the network [15,16]. Along with this, transfer learning on Convolutional Neural Networks (CNN) was developed in [15,17] for detecting breast cancer using histopathological images. Furthermore, CNN was integrated with Fish vector (FV) and developed a new layer for improving the accuracy and discriminative power [3].

The major purpose of this research work is to design breast cancer classification using ALO-based DNN. This breast cancer classification method mainly includes four stages, namely pre-processing, blood cell segmentation, feature extraction, and breast cancer classification. At first, the input image is given to pre-processing stage where the T2FCS filter is utilized for removing the noises in the input image. After that, pre-processed image is passed to blood cell segmentation for segmenting blood cells using color-based thresh holding. Consequently, the significant statistical features, like variance, mean, standard deviation, kurtosis, skewness, and CNN features are extracted for effective classification outcome. Finally, the generated feature vector is fed to DNN for a breast cancer classification where the classifier is trained using the developed ALO technique.

The major contribution of the paper is:

- ❖ Proposed Ant Lion Optimizer-based Deep Neural Network for breast cancer classification: The breast cancer detection model is performed using developed ALO-based DNN. The characteristic features of ALO attain the best classification outcome by optimal fitness value.

The rest of the paper is arranged as follows: section 2 illustrates a literature survey of present breast cancer classification approaches. Section 3 describes the developed ALO-based DNN for detecting breast cancer. Section 4 displays results and discussion of developed breast cancer detection technique, and the conclusion of the paper is presented in section 5.

## 2. Motivation

This section elaborates on the motivation of breast cancer classification. The downsides of existing breast cancer detection approaches are considered to devise a novel breast cancer classification technique using the ALO-based DNN method.

### 2.1. Literature Survey

This section explains about literature survey of existing breast cancer classification techniques. A generic driven wrapper embedded with feature type aware hybrid Bayesian classifier was devised by Qiqige Wuniri et al. [1]. Here, the Bayesian classifier and combined genetic-driven structure were included for the feature extraction process using kernel probability density estimation. The kernel-based Bayesian classifier and the closed interpretable solution were employed for handling discrete features and continuous features. After that, every classifier was subjected to feature subsets and it was estimated by their fitness index, termed as Area Under the Curve (AUC). Moreover, a Genetic Algorithm (GA) was employed to get an near-optimal feature subset, and it produces better AUC. Finally, one class F-score was utilized for obtaining fast convergence. This method was more reliable and achieved high accuracy than other GA techniques, but this method failed to check feasibility by considering other datasets. Jing Zheng et al. [2] developed Deep Learning assisted Efficient Adaboost Algorithm (DLA-EABA) for breast cancer classification. In this method, an ensemble classifier was employed with boosting technique for identifying various metastases in breast tumors. The CNN-based transfer learning model was utilized for early and accurate breast cancer classification using various modalities. This classification approach achieved high accuracy for breast cancer detection. Although, the computational time of the developed breast cancer classification approach was high. Yu-Dong Zhang et al. [3] introduced the combination of Graph Convolutional Network (GCN) model and Convolutional Neural Network (CNN) scheme for breast cancer classification. At first, standard 8-layer CNN was utilized for the breast cancer classification process. Moreover, the improved approach, like Drop Out (DO) and Batch Normalization (BN) was integrated for enhancing the accuracy of breast cancer classification. Consequently, Rank-based Stochastic Pooling (RSP) was applied for reserving conventional max pooling. Moreover, this developed approach was hybridized with two-layer GCN for analyzing breast mammogram images. This method achieved better classification results but still failed to include huge databases for better performance in huge clinical studies. Pin Wanga et al. [4] modeled cross-task Extreme Learning Machine (ELM) scheme with deep convolutional features for breast cancer classification. Initially, high-level features were

extracted by double-step deep transfer learning and a deep transfer learning method. Subsequently, extracted high-level features were mutually utilized as regularization for improving the classification accuracy. In addition, interactive cross-talk ELM was also introduced for enhancing the accuracy of detection with deep convolutional features and loss functions. This developed technique achieved better classification accuracy, but still, this method does not decrease the computational complexity.

## 2.2 Challenges

The challenges faced by existing breast cancer classification approaches are explained as follows:

- ❖ The generic driven wrapper model embedded with feature type aware hybrid Bayesian classifier [1] was developed for the breast cancer identification process. However, this method not utilized a better dataset for validating the effectiveness and accuracy.
- ❖ The combination of the CNN and GCN model [3] was introduced for breast cancer classification. Although, this technique failed to include deeper GCN with more than two layers for improving classification outcomes.
- ❖ Cross task extreme learning machine with deep convolutional features technique [4] was devised for breast cancer classification, but still, this method failed to decrease the training time and computational time.
- ❖ The new transfer learning-based approach was developed in [27] for automatic breast cancer classification. However, this approach not included a stain color normalization scheme for handling color difference management.
- ❖ Normally, automatic multi-class breast cancer classification faces various problems because of low inter-class connection and high intra-class similarity concerns.

## 3. Proposed Ant Lion Optimization Algorithm for Breast Cancer Classification

This section explains the developed ALO-based DNN technique for breast cancer classification. This breast cancer classification method mainly includes four stages, namely pre-processing, blood cell segmentation, feature extraction, and breast cancer classification. Originally, the input breast image is given to pre-processing stage for removing noises from the input image. In the pre-processing phase, the T2FCS filter [18] is applied to improve the quality of the input image. After that, blood cell segmentation is carried out using a color-based thresholding model [19] in which blood cells is extracted for the classification process. Once the blood cells are extracted, the feature extraction process is executed. The significant features, namely statistical features, namely skewness, standard deviation, mean variance, kurtosis, and CNN features are extracted for breast cancer classification. Finally, the extracted features are taken as input to the classification phase where DNN classifier [20] [22] is employed for detecting cancerous regions. Moreover, the DNN classifier is trained by a developed ALO algorithm [21] for obtaining effective classification outcomes. Fig. 1 displays the block diagram of the developed ALO-based DNN model for breast cancer classification.

### 3.1 Input Image Acquisition

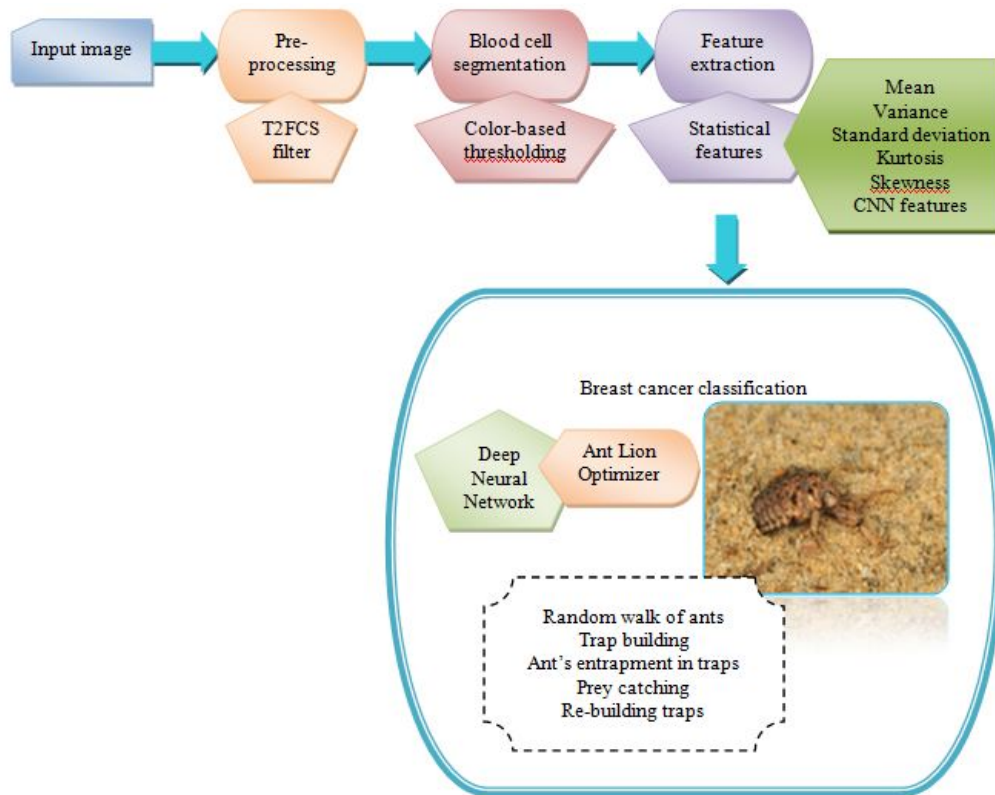
The input breast image is taken from the database and it is utilized for the cancer classification process. Let us consider the dataset  $G$  with  $m$  quantity of breast cancer images, which is expressed as

$$G = \{S_1, S_2, \dots, S_r, \dots, S_q\}; 1 \leq r \leq q \quad (1)$$

where,  $q$  is the total amount of images,  $S_r$  signifies the image placed at  $r^{\text{th}}$  index, and  $G$  is the dataset.

### 3.2 Pre-Processing using T2FCS Filter

The input image  $S_r$  is obtained from the database and it is given to pre-processing phase in which the noises are removed; thereby the image quality is improved. The T2FCS filter [18] is generally developed by combining the cuckoo search optimization technique and type II fuzzy system for reducing the noises in images. Finally, the pre-processed image is signified as  $T_r$  and it is utilized for further processing.



*Fig.1. Block diagram of developed Ant Lion-based Deep Neural Network for breast cancer detection*

### 3.3 Blood Cell Segmentation Using Color-Based Thresholding

The pre-processed image  $T_r$  is taken as input for the segmentation stage where blood cells are segmented for the classification process. Segmentation is a process of subdividing the image into its element objects or regions. Here, the color-based thresholding scheme [19] is applied using the color information of the bacterium for extracting the blood cells. This approach indicates the range of Red Green Blue (RGB) intensities for the thresholding process. The objects, which present the exterior from the selected range are eliminated. Thus, the selection range identification is a significant task for segmenting the accurate blood cells. Moreover, the blood cell segmented output is specified as  $B_r$ .

### 3.4 Feature Extraction Based On Statistical Features

After the segmentation process, the feature extraction stage is executed using the output of blood cell segmentation  $B_r$ . The statistical features, like variance, mean, skewness, standard deviation, and kurtosis are extracted from the segmented output. The feature extraction process is significant for classifying breast cancer, whereas the statistical features ensure the effectiveness of breast cancer classification. The statistical features are extracted from segmented output  $B_r$  is explained as follows.

**a) Mean:** Mean value of segmented blood cell is evaluated by estimating average pixels, which are enclosed in an image, and it is expressed as,

$$s_1 = \sum_{t=0}^{n-1} t \cdot H(t) \quad (2)$$

**b) Variance:** The variance is defined as a measure of signal deviation from the mean value of the segmented image. It is formulated by,

$$s_2 = \sum_{t=0}^{n-1} (t - s_1)^2 \cdot H(t) \quad (3)$$

**c) Standard deviation:** Standard deviation is estimated using the mean value of the segmented image, and it is computed by,

$$s_3 = \sqrt{\sum_{t=0}^{n-1} (t - s_1)^2 \cdot H(t)} \quad (4)$$

**d) Skewness:** Skewness is extracted by a symmetry measure, and it is estimated by,

$$s_4 = \frac{\sum_{t=0}^{n-1} (t - s_1)^3 \cdot H(t)}{(s_3)^3} \tag{5}$$

**e) Kurtosis:** Kurtosis finds whether the data is peak or flat relative with regards to normal distribution.

$$s_5 = \frac{\sum_{t=0}^{n-1} (t - s_1)^4 \cdot H(t)}{(s_3)^4} \tag{6}$$

where,  $H(t)$  is the probability of incidence for the individual grey level  $t$  of the segmented image  $B_r$ .

**CNN features:** CNN feature is the most important feature extracted from segmented blood cells for improving classification accuracy. The CNN feature mainly includes three layers, a convolution layer, a pooling layer, and a fully connected layer. The initial layer is the convolution layer in which the segmented blood cell is subjected as the input. Furthermore, the output produced by the convolution layer is considered as CNN features and it is indicated as  $s_6$ .

Moreover, the feature vector is generated by extracted CNN feature and statistical features and it is formulated by,  $s_w = \{s_1, s_2, s_3, s_4, s_5, s_6\}$  with  $[6 \times 1]$  dimension.

### 3.5 Breast Cancer Classification Using Developed ALO-Based DNN

Once the significant features are extracted, breast cancer identification is carried out by a DNN classifier. The feature vector  $s_w$  is taken as the input to the DNN classifier for breast cancer classification. Furthermore, the DNN classifier is trained by a developed ALO algorithm for effective classification.

#### 3.5.1 Structure of DNN

Here, DNN classifier [20] [22] is employed for the classification process of breast cancer using the generated feature vector from a segmented blood cell. The DNN classifier includes a number of hidden layers linked among the input layer and the output layer. DNN estimates accurate mathematical manipulation to tune the input values to obtain the output. The design of the DNN classifier is portrayed in fig. 2. The main benefit of utilizing the DNN classifier is that it has decreased computational time. Let us assume  $d$  is the network layer,  $M$  is the input subjected to  $d^{th}$  network and  $N^d$  is the output value of  $d^{th}$  layer. Moreover, the weight of  $d^{th}$  layer is denoted as  $W^d$ , and it linearly changes the input value to output value,  $a^d$  signifies the activation vector function, and  $q$  indicates the bias. The input vector is characterized by,

$$U^d = M \cdot c^d \tag{7}$$

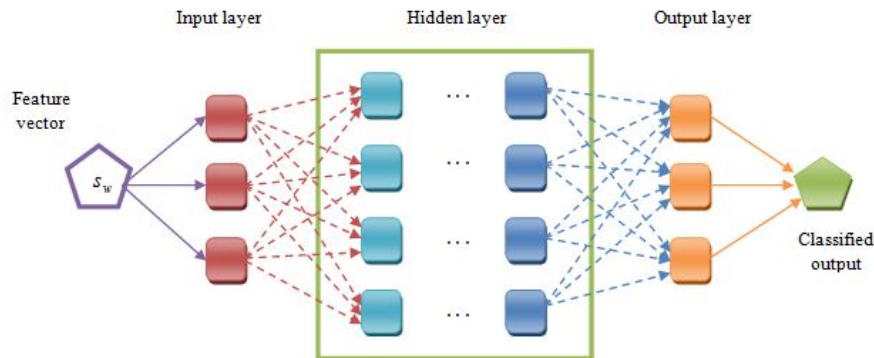
The hidden layer is estimated as,

$$V^d = W^d a^d + q^d \tag{8}$$

Additionally, the output calculated using DNN classifier is signified as,

$$N^d = a(V^d) \tag{9}$$

where,  $c$  denotes independent samples and  $\cdot$  specifies the element-wise vector product. The value of  $c$  is re-sampled at every update process of training.



**Fig.2.** Architecture diagram of DNN classifier

### 3.5.2 Algorithmic procedure of developed ALO algorithm

The training process of the DNN classifier is performed by a developed optimization technique, named the ALO algorithm. Here, the ALO technique [21] is a foraging characteristic of antlions, which imitates the interface among antlions in trap and ants. The ants are permitted to shift in search space, while antlions chase the ant to become fitter with traps. Moreover, a random walk is selected for representing the ant's movement in a stochastic nature environment. The algorithmic process of developed ALO algorithm is explained as follows,

**i) Initialization:** The population of antlions is randomly initialized, denoted as  $I$ , and it is formulated as,

$$I = \{I_1, I_2, \dots, I_k, \dots, I_l\} ; 1 \leq k \leq l \quad (10)$$

**ii) Fitness function estimation:** The fitness function is utilized to calculate the best solution to classify breast cancer and it is expressed as,

$$J = \frac{1}{f} \sum_{\beta=1}^f [A_x - Z^\beta] \quad (11)$$

where,  $Z^\beta$  is the target output,  $A_x$  represents the classified output,  $J$  indicates fitness function, and  $f$  is the total amount of samples.

**iii) Update C and D:** The antlion generates the traps with respect to fitness function such that, ants can randomly shift to search the food. On the other hand, shoot of antlions sands outwards to pit center only if it recognizes that the ant is in the trap. The trapping behavior of the ant slides down the ant to escape. Besides, the below equations are mathematically expressed for representing the trapping behavior of ants.

$$C^g = \frac{C^g}{X} \quad (12)$$

$$D^g = \frac{D^g}{X} \quad (13)$$

where,  $X$  is a ratio,  $D^g$  signifies the vector, which contains the maximum value of every variable at  $g^{\text{th}}$  iteration and  $C^g$  denotes the minimal variables at  $g^{\text{th}}$  iteration.

**iv) Random walks of ants:** The random walk of ant is chosen for modeling the ant movements because ant stochastically shifts for searching food and it is represented by,

$$E(g) = [0, I(2h(g_1) - 1), I(2h(g_1) - 1), \dots, I(2h(g_i) - 1)] \quad (14)$$

where,  $I$  is utilized to calculate the cumulative sum,  $h(g)$  indicates the stochastic function, and  $i$  signifies the maximum amount of iterations. Furthermore, the stochastic function is illustrated as,

$$h(g) = \begin{cases} 1 & \text{if rand} > 0.5 \\ 0 & \text{if rand} \leq 0.5 \end{cases} \quad (15)$$

where,  $\text{rand}$  specifies the random number, which ranges from  $[0,1]$ . Moreover, ants update their position using random walk at each stage of optimization. The ants are normalized for maintaining the random walk in a search space, and the normalization is executed by the below equation.

$$I_u^g = \frac{(I_u^g - o_u) \times (D_u^g - C_u^g)}{D_u^g - o_u} + C_u \quad (16)$$

where,  $o_u$  is the minimum of the random walk at  $u^{\text{th}}$  variable,  $D_u^g$  is the maximum variable of  $u^{\text{th}}$  variable at  $g^{\text{th}}$  iteration, and  $C_u^g$  is the minimum variable of  $u^{\text{th}}$  variable at  $g^{\text{th}}$  iteration.

**v) Elitism:** Elitism is a significant factor of the evolutionary approach, which permits to sustain the optimal solution attained by the optimization process. The best antlion attained at  $g^{\text{th}}$  iteration is considered elite. Normally, the elite will influence the movement of ants at every iteration because the elite is the best antlion. Therefore, ant walks around the roulette wheel is formulated as,

$$K_u^g = \frac{F_u^g + F_v^g}{2} \quad (17)$$

where,  $F_v^g$  is random walkabout elite at  $g^{\text{th}}$  iteration and  $F_u^g$  represents random walk in the order of elite at  $g^{\text{th}}$  iteration.

**vi) Catch prey and re-building pit:** The final segment of the hunt is attained, while the ant arrives at bottom of the pit and is trapped in an antlion's mouth. After that, the antlion pulls the ant in the sand and eats its body. Here, it is considered that the prey catching happens, while an ant becomes fitter than the corresponding antlion to imitate this process. However, antlion is needed for updating the

location of hunted ant, hence it improves the possibility of caching new prey. The below equation is expressed to catch the prey and re-creating a pit.

$$R_w^g = K_u^g \quad \text{if } f(K_u^g) > f(R_w^g) \quad (18)$$

where,  $R_w^g$  is the position of  $w^{\text{th}}$  variable at  $g^{\text{th}}$  iteration and  $K_u^g$  is the position of  $u^{\text{th}}$  variable at  $g^{\text{th}}$  iteration.

**vii) Re-evaluate the fitness:** The fitness value is estimated using the objective function, which is defined in equation (11) in which the fitness measure with optimal value is considered as the best solution.

**viii) Termination:** The above process is continued until an optimal solution is achieved for predicting breast cancer. Algorithm 1 illustrates a pseudo-code of the developed ALO technique.

**Algorithm1.** Pseudo code of developed ALO algorithm

---

```

Input: Ant lion population
Output: Best solution
Start
    Randomly initialize the population of antlions
    Estimate the fitness function
    Identify best antlions
while (end condition is not fulfilled)
for all ant
    Choose antlion
    Update C and D by equation (12) and (13)
    Generate random walk and regulate it based on equation (14) and (15)
    Update location of an ant by equation (17)
end for
    Estimate fitness for every ant
    Replace the antlion with the best one using equation (18)
Update elite
end while
Optimal solution is achieved
End

```

---

## 4. Results and discussion

The results and the discussion of the developed breast cancer detection method are expressed in this section. The experimental setup, performance metrics, dataset description, competing methods, and comparative analysis are given below.

### 4.1 Experimental setup

The implementation of the developed breast cancer classification method is carried out using MATLAB with 4GB RAM, Windows 8 OS with the Intel Core i-3 processor.

### 4.2 Performance Metrics

The performance of the developed ALO-based DNN is estimated using three performance metrics, namely accuracy, sensitivity, and specificity.

**a) Accuracy:** Accuracy is estimated as degree of closeness of approximated rate regarding its original, and it is expressed as,

$$\text{Accuracy} = \frac{F_p + F_N}{F_p + T_N + F_N + T_p} \quad (19)$$

**b) Sensitivity:** The percentage of positives, which are perfectly predicted through a classifier, is termed as sensitivity and it is indicated as,

$$\text{Sensitivity} = \frac{F_p}{F_p + T_N} \quad (20)$$

**c) Specificity:** Specificity is the relative amount of negatives that are appropriately determined through a classifier, and it is denoted as,

$$\text{Specificity} = \frac{F_N}{F_N + T_p} \quad (21)$$

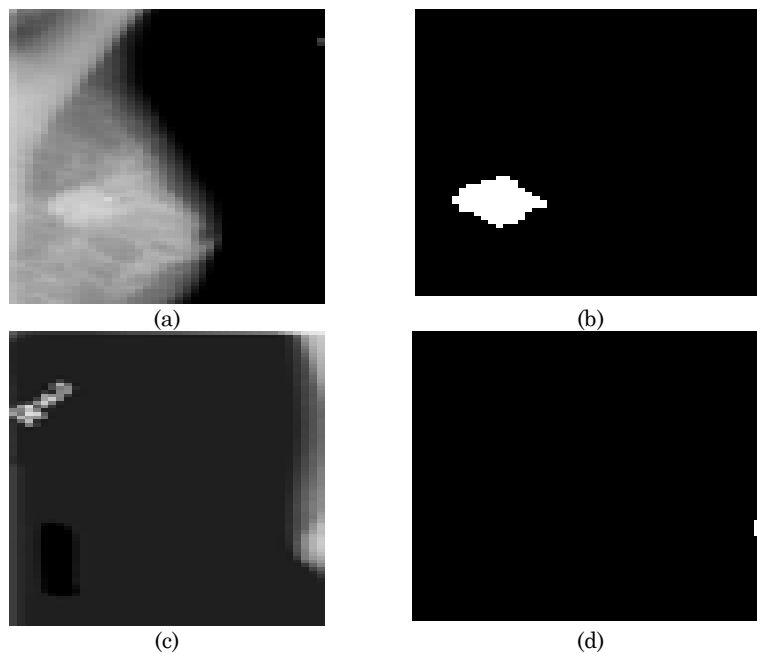
where,  $F_N$ ,  $F_p$ ,  $T_p$  and  $T_N$  represents false negative, false positive, true positive and true negative.

### 4.3 Dataset Description

The database considered for breast cancer classification is Digital Database for Screening Mammography (DDSM) [26]. DDSM includes mammographic images and it is offered by the Breast Cancer Research Program of U.S. Army Medical Research. The data is gathered from Sandia National Laboratories and Massachusetts General Hospital. This DDSM is arranged as 2620 cases and volumes, which are available on 8mm tape. A case is a group of information and images equivalent to one mammography assessment of a single patient. Besides, the volume is a set of cases gathered mutually for the intention of easy allocation.

### 4.4 Experimental Results

The experimental results for breast cancer detection are displayed in this section. The input image of the left side breast is shown in fig. 3 a) and fig. 3 b) indicates the ground truth image of the left side breast image. Similarly, fig. 3 c) denotes the input image of the right-side breast image, and fig. 3 d) signified the ground truth image of the right-side breast input image.



*Fig.3. Experimental results a) input image-1, b) input image-2, c) ground truth-1, d) ground truth image-2*

### 4.5 Competing Techniques

The developed ALO-based DNN model for breast cancer classification is compared using existing techniques, like Deep Belief Network (DBN) [23], Support vector machine (SVM) [24], and DCNN [25].

### 4.6 Comparative Analysis

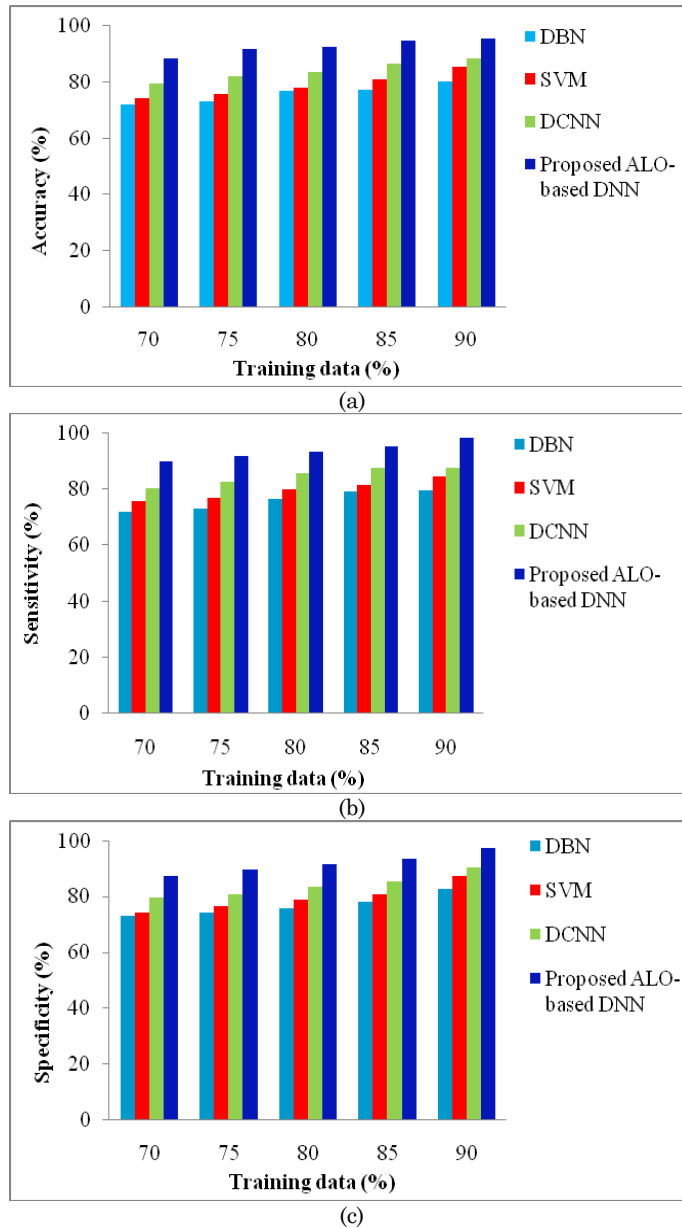
The comparative analysis of the developed ALO-based DNN algorithm with respect to the accuracy, sensitivity, and specificity parameters are estimated by changing training data percentage and K-fold value.

#### 4.6.1 Analysis using training data percentage

The comparative analysis of developed ALO-based DNN by changing training data percentage with respect to accuracy, sensitivity, and specificity is displayed in fig. 4. Fig. 4 a) portrays the analysis of accuracy through changing the training data percentage. The accuracy attained through existing breast cancer classification methods, like DBN, SVM, DCN and developed ALO-based DNN are 76.84%, 77.89%, 83.47%, and 92.56% for 80% of training data. Moreover, the performance improvement achieved by the developed ALO-based DNN technique is 16.98%, 15.84% and 9.82%, when compared with existing DBN, SVM, DCN approaches. The comparative analysis of sensitivity by changing training data percentage is



represented in fig. 4 b). In 80% of training data, the developed breast cancer classification techniques achieved 93.27% of sensitivity, while existing techniques, such as DBN, SVM, DCN has 76.51%, 79.84% and 85.74%. In addition, the performance improvement of developed ALO-based DNN technique with existing DBN, SVM and DCN approaches is 17.96%, 14.39% and 8.07%. Fig. 4 c) presents the comparative analysis of specificity by varying the training data percentage. The developed breast cancer classification method has a specificity value of 91.56% and existing techniques, like DBN is 75.84%, SVM is 78.89% and DCN is 83.47% for 80% of training data. The performance improvement of developed ALO-based DNN method with DBN is 17.16%, SVM is 13.83% and 8.83%.

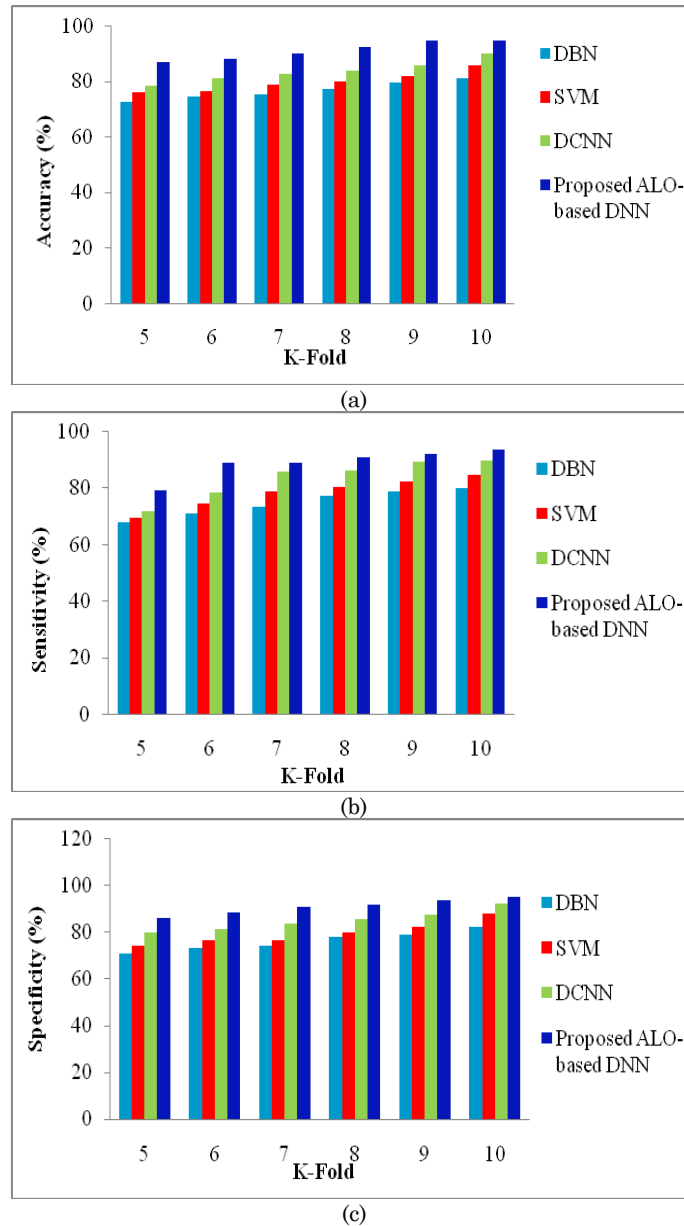


**Fig.4.** Analysis of developed ALO-based DNN a) accuracy, b) sensitivity and c) specificity by changing training data percentage

#### 4.6.2 Analysis using K-fold value

The comparative analysis of proposed ALO-based DNN through varying K-Fold value with respect to accuracy, sensitivity and specificity is represented in fig. 5. Fig. 5 a) displays the comparative analysis of accuracy by changing the K-Fold value. The developed breast cancer classification technique has an accuracy value of 94.51% and existing approaches, like DBN, SVM and DCN methods are 79.38%, 81.94%, and 85.59%, when the K-Fold value is 9. The performance improvement of the developed ALO-based DNN method with DBN is 16%, SVM is 13.30% and 9.43%. Fig. 5 b) presents the comparative analysis of sensitivity by varying the K-Fold value. The sensitivity attained by existing breast cancer

classification methods, like DBN, SVM, DCN and developed ALO-based DNN are 78.67%, 82.17%, 89.2% and 91.96% for K-Fold value 9. Moreover, the performance improvement achieved by the developed ALO-based DNN technique is 14.45%, 10.64% and 3%, when compared with existing DBN, SVM, DCN techniques. The comparative analysis of specificity by changing K-Fold value is illustrated in fig. 5 c). When the K-Fold value is 9, then the developed breast cancer classification techniques achieved 93.74% of specificity, while existing techniques, such as DBN are 79.2%, SVM is 82.53% and DCN has 87.62%. In addition, the performance improvement of developed ALO-based DNN technique with existing DBN, SVM and DCN approaches are 15.51%, 11.95% and 6.52%.



**Fig.5.** Analysis of developed ALO-based DNN with respect to a) accuracy, b) sensitivity and c) specificity by varying K-Fold value

#### 4.7 Comparative Discussion

The comparative discussion of several breast cancer classification techniques in terms of accuracy, sensitivity and specificity parameters is represented in Table 1 by changing the training data percentage and K-Fold value. The developed ALO-based DNN techniques obtain better accuracy of 95.34% for 90% of training data, whereas existing approaches, like DBN, SVM and DCNN have 79.46%, 84.23% and 87.63%. The sensitivity value of the developed breast cancer classification scheme is 97.54% and existing techniques, such as DBN has 80.05%, SVM has 85.47% and DNN has 88.4% in 90% of training data.

Similarly, the developed ALO-based DNN achieved 95.34% of specificity, while existing DBN, SVM and DCNN schemes are 79.46%, 84.23% and 87.63%.

The accuracy value of the developed breast cancer classification method is 95.35% and existing techniques, such as DBN have 81.56%, SVM has 85.95% and DNN has 88.61%, when K-Fold value is 10. The developed ALO-based DNN techniques obtained better sensitivity of 93.62% for 90% of training data, whereas existing approaches, such as DBN, SVM and DCNN have 79.96%, 84.56% and 89.88%. Furthermore, the developed ALO-based DNN achieved 93.62% of specificity, while existing DBN, SVM and DCNN methods have 79.96%, 84.56% and 89.88%. Thus, from the below table, it is well-known, that the developed breast cancer classification model, named ALO-based DNN technique achieved maximum accuracy, sensitivity and specificity.

**Table1.** Comparative discussion

Analysis	Metrics	DBN	SVM	DCNN	Proposed ALO-based DNN
Based on training data	Accuracy (%)	80.05	85.47	88.4	<b>95.34</b>
	Sensitivity (%)	79.46	84.23	87.63	<b>98.05</b>
	Specificity (%)	82.61	87.43	90.57	<b>97.54</b>
Based on K-Fold value	Accuracy (%)	81.2	85.87	90.09	<b>94.74</b>
	Sensitivity (%)	79.96	84.56	89.88	<b>93.62</b>
	Specificity (%)	82.56	87.95	92.61	<b>95.35</b>

## 5. Conclusion

In this research work, an efficient breast cancer classification approach was introduced using a developed ALO-based DNN scheme. At first, the breast image was considered as input and it was subjected to pre-processing for removing noise and improving the image quality. Here, T2FCS filter was utilized for pre-processing of the input image. After that, color-based thresholding was utilized to segment the blood cells from the pre-processed image. Once the blood cells are segmented, a feature extraction process was performed for extracting statistical features, namely skewness, mean, standard deviation, kurtosis, variance, and CNN features. Moreover, extracted features are utilized for further breast cancer classification processes. Finally, breast cancer detection is carried out using DNN classifier, and it was trained by a developed optimization algorithm, named ALO. Here, the implementation of the developed breast cancer classification process was carried out using DDSM dataset. Moreover, the performance of the developed breast cancer classification approach was estimated using three performance metrics, such as sensitivity, accuracy and specificity. Therefore, the developed ALO-based DNN technique achieved maximum accuracy, sensitivity and specificity of 95.34%, 97.54% and 95.34%. In the future, other new classifiers and optimization algorithms are included for better performance of breast cancer classification.

## Compliance with Ethical Standards

**Conflicts of interest:** Authors declared that they have no conflict of interest.

**Human participants:** The conducted research follows the ethical standards and the authors ensured that they have not conducted any studies with human participants or animals.

## References

- [1] Wuniri Q, Huangfu W, Liu Y, Lin X, Liu L, Yu Z, "A Generic-Driven Wrapper Embedded with Feature-Type-Aware Hybrid Bayesian Classifier for Breast Cancer Classification", *IEEE Access*, vol.7, pp.119931-42, August 2019.
- [2] Zheng J, Lin D, Gao Z, Wang S, He M, Fan J, "Deep Learning assisted Efficient AdaBoost Algorithm for Breast Cancer Detection and Early Diagnosis", *IEEE Access*, May 2020.
- [3] Zhang YD, Satapathy SC, Guttery DS, Górriz JM, Wang SH, "Improved Breast Cancer Classification Through Combining Graph Convolutional Network and Convolutional Neural Network", *Information Processing & Management*, vol.58, no.2, pp.102439, November 2020.
- [4] Wang P, Song Q, Li Y, Lv S, Wang J, Li L, Zhang H, "Cross-task extreme learning machine for breast cancer image classification with deep convolutional features", *Biomedical Signal Processing and Control*, vol.57, pp.101789, March 2020.
- [5] Jardines L, Goyal S, Fisher P, Weitzel J, Royce M, Goldfarb SB, "Breast cancer overview: Risk factors, screening, genetic testing, and prevention", *Cancer*, June 2015.

- [6] Fakoor R, Ladhak F, Nazi A, Huber M, "Using deep learning to enhance cancer diagnosis and classification", In Proceedings of the international conference on machine learning, vol.28, June 2013.
- [7] Shah C, Arthur DW, Wazer D, Khan A, Ridner S, Vicini F, "The impact of early detection and intervention of breast cancer-related lymphedema: a systematic review", *Cancer medicine*, vol.5, no.6, pp.1154-62, June 2016.
- [8] Ahmad F, Isa NA, Hussain Z, Osman MK, Sulaiman SN, "A GA-based feature selection and parameter optimization of an ANN in diagnosing breast cancer", *Pattern Analysis and Applications*, vol.18, no.4, pp.861-70, November 2015.
- [9] Aličković E, Subasi A, "Breast cancer diagnosis using GA feature selection and Rotation Forest", *Neural Computing and Applications*, vol.28, no.4, pp.753-63, April 2017.
- [10] Zhao M, Fu C, Ji L, Tang K, Zhou M, "Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes", *Expert Systems with Applications*, vol.38, no.5, pp.5197-204, May 2011.
- [11] Saha M, Chakraborty C, "Her2Net: A deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation", *IEEE Transactions on Image Processing*, vol.27, no.5, pp.2189-200, January 2018.
- [12] Veta M, Pluim JP, Van Diest PJ, Viergever MA, "Breast cancer histopathology image analysis: A review", *IEEE Transactions on Biomedical Engineering*, vol.61, no.5, pp.1400-11, January 2014.
- [13] Li Y, Wu J, Wu Q, "Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning", *IEEE Access*, vol.7, pp.21400-8, February 2019.
- [14] Garud H, Karri SP, Sheet D, Chatterjee J, Mahadevappa M, Ray AK, Ghosh A, Maity AK, "High-magnification multi-views based classification of breast fine needle aspiration cytology cell samples using fusion of decisions from deep convolutional networks", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp.76-81, 2017.
- [15] Li L, Pan X, Yang H, Liu Z, He Y, Li Z, Fan Y, Cao Z, Zhang L, "Multi-task deep learning for fine-grained classification and grading in breast cancer histopathological images", *Multimedia Tools and Applications*, vol.79, no.21, pp.14509-28, June 2020.
- [16] Akbar S, Peikari M, Salama S, Nofech-Mozes S, Martel A, "Transitioning between convolutional and fully connected layers in neural networks", In proceedings of Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp.143-150, September 17.
- [17] Zhi W, Yueng HW, Chen Z, Zandavi SM, Lu Z, Chung YY, "Using transfer learning with convolutional neural networks to diagnose breast cancer from histopathological images", In proceedings of International Conference on Neural Information Processing, pp.669-676, November 2017.
- [18] Kumar SV, Nagaraju C, "T2FCS filter: Type 2 fuzzy and cuckoo search-based filter design for image restoration", *Journal of Visual Communication and Image Representation*, vol.58, pp.619-41, January 2019.
- [19] Kulkarni N, "Color thresholding method for image segmentation of natural images", *International Journal of Image, Graphics and Signal Processing*, vol.4, no.1, pp.28, February 2012.
- [20] Saxe, J. and Berlin, K., "Deep neural network based malware detection using two dimensional binary program features", In IEEE 10th International Conference on Malicious and Unwanted Software (MALWARE), pp. 11-20, October 2015.
- [21] Mirjalili S, "The ant lion optimizer", *Advances in engineering software*, vol.83, pp.80-98, May 2015.
- [22] Duan, G., Hu, W. and Wang, J., "Research on the natural image super-resolution reconstruction algorithm based on compressive perception theory and deep learning model," *Neurocomputing*, vol.208, pp.117-126, 2016.
- [23] Abdel-Zaher AM, Eldeib AM, "Breast cancer classification using deep belief networks", *Expert Systems with Applications*, vol.46, pp.139-44, March 2016.
- [24] Rejani Y, Selvi ST, "Early detection of breast cancer using SVM classifier technique", arXiv preprint arXiv:0912.2314, December 2009.
- [25] Tu F, Yin S, Ouyang P, Tang S, Liu L, Wei S, "Deep convolutional neural network architecture with reconfigurable computation patterns", *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol.25, no.8, 2220-33, April 2017.
- [26] DDSM: Digital Database for Screening Mammography taken from, "<http://www.eng.usf.edu/cvprg/Mammography/Database.html>", accessed on January 2021.
- [27] Boumaraf S, Liu X, Zheng Z, Ma X, Ferkous C, "A new transfer learning based approach to magnification dependent and independent classification of breast cancer in histopathological images", *Biomedical Signal Processing and Control*, vol.63, pp.102192.
- [28] Amol V Dhumane, "Examining User Experience of eLearning Systems using EKhoool Learners", *Journal of Networking and Communication Systems (JNACS)*, Volume 3, Issue 4, October 2020.