



## **Multi-Objective Feature Selection Method by Using ACO with PSO Algorithm for Breast Cancer Detection**

**Rajesh Saturi<sup>1\*</sup>      Parvataneni Premchand<sup>1</sup>**

*<sup>1</sup>Department of Computer Science and Engineering, University College of Engineering,  
Osmania University, Hyderabad, India*

\* Corresponding author's Email: [rajeshsaturi.oucse@gmail.com](mailto:rajeshsaturi.oucse@gmail.com)

---

**Abstract:** Breast cancer is the second most cause of cancer deaths in women after lung cancer, and the treatment and diagnosis of breast cancer at an earlier stage is an important requirement for the reduction of the mortality rate. The detection of cell nuclei is the core operation involved in Computer-Aided Diagnosis (CAD) that is used in breast cancer detection. The existing deep learning methods used contour information for accurate detection. But, the existing cancer detection method extracted only the semantic information from the former layer and was not provided the details about the semantic shallow layers in detecting cancer. To solve such an issue, a multi-objective feature selection method was proposed by using Ant Colony Optimization (ACO) with Particle Swarm Optimization (PSO) for the detection of breast cancer. The proposed method utilized the BreakHis dataset and data augmentation was applied to analyze the internal morphological features of the images. The features were extracted by using Convolutional Neural Network (CNN), Histogram of Gradients (HoG) and Local Ternary Pattern (LTP). Then, the features are selected by the proposed multi-objective feature selection method by using ACO with PSO. The selected features are forwarded to the Long Short Term Memory (LSTM) network for the process of classification. The proposed multi-objective feature selection method has the advantage of minimizing the features and reduces the rate of error classification, through a selection of optimal features to increase the rate of breast cancer detection. The proposed multi-objective feature selection method achieved a higher accuracy rate of 95.72% in the detection of breast cancer. Whereas, the existing method showed an accuracy of 93.68% in the detection of breast cancer.

**Keywords:** Ant colony optimization, Breast cancer, Computer-aided-diagnosis, Convolutional neural network, Histogram of gradients.

---

### **1. Introduction**

Breast cancer is the second most cause of cancer deaths in women after lung cancer. The death rate of women due to breast cancer is around 1 in 37 individuals that amount to 2.7%. The treatment and diagnosis of breast cancer in an early stage is necessary for the reduction of death rates and to prevent the progression of the disease [1]. The pathological images serve as effective standards for doctors and research scholars for the diagnosis of disease. The variations in morphological features of nuclei images are the source for tumor detection and diagnosis [2]. So, the effective detection of nuclei in images is basic requirement for efficient diagnosis of

breast cancer processes [3]. The biopsy samples of tissue or cells are kept on the glass slide for the process of staining and frequent microscopic examination. The pathologists view the glass slides under the microscope for analyzing various components of tissue that is dyed with more than one stain to examine the cellular level of the tumor [4]. Hematoxylin-Eosin (H-E) staining is the most utilized staining method by pathologists. The Nuclei are stained by purple and blue color with Hematoxylin stain. In the background structure, the stroma and cytoplasm are stained by color pink by using Eosin stain. The size of the nucleus, shape, ratio of cytoplasm, and chromatin patterns are considered in the tissue affected by breast cancer [5].

The detection of cell nuclei is the core operation in CAD for cancer detection. The CAD serves as the basis for counting the cells and for the study of subcellular morphology such as the investigation of shape, texture, and size of the cellular components. The detection of cell nuclei is difficult during the analysis of histopathological images such as target cells that are in various states [6]. The smaller cells will be surrounded by the background clutters made up of histopathological structures such as collagen, capillaries, etc., along with the irrelevant visual aspects and artefacts which occur during the acquisition of images [7]. The challenges faced during the automatic analysis of breast cancer from the histopathological images were mainly the complex and diverse structures. The difference in the appearance of H-E stained images was due to the various concentrations of the applied stains and inconsistent scanning of the images which significantly degrades the performance of subsequent processes of nuclei detection [8]. The normalization of stain is carried out to enhance the quality and consistency of appearance in the images. Another issue faced was less availability of training data, a larger size of biopsy images, and resizing the images, all of which resulted in the loss of detailed information [9]. The existing deep learning methods used the contour information for accurate detection of breast cancer. But, the existing method showed lesser performance by extracting only the semantic information from the former layer and not extracted the details from the shallow layer which is required for effective breast cancer detection [10]. To solve this issue, a multi-objective feature selection method that uses ACO with PSO for breast cancer detection is proposed. The BreakHis dataset is used in the proposed method and the images undergo pre-processing by using color variation and data augmentation. Then, the features are extracted by using CNN, HoG, and LTP to identify the salient region from images. After extracting the features, the features are selected by the proposed multi-objective feature selection method by using ACO with PSO. The selected features are forwarded to the LSTM network for the process of classification. The proposed optimization method has the advantage of reducing the features and minimizes the error rate of classification by selecting optimal features to increase the rate of breast cancer detection.

The paper is organized as follows, the review of existing methods is described in section 2, the proposed methodology for nuclei segmentation is explained in section 3, the experimental results and discussion are present in Section 4, the conclusion of this research is explained in section 5.

## 2. Literature review

The existing techniques that were used for breast cancer detection, are reviewed and the advantages and limitations of these reviewed techniques are also explained in this section.

Budak et al. [11] Developed Fully Convolutional Network (FCN) and Bi-directional Long Short Term Memory (Bi-LSTM) for breast cancer detection. The developed method exploited FCN as an encoder which facilitated the efficient processing of larger scale of images and input of variable image sizes. The developed FCN method consumes more time in extracting the features due to the limitation of space. Wang et al. [12] Established deep feature fusion and enhanced routing by utilizing Convolutional Neural Network (CNN) and CapsNet for classification of breast cancer images. The developed method effectively obtained discriminative features by extracting convolutional features and integrating the semantic as well as spatial features. It was difficult to use the developed feature fusion method to classify the dataset as it is based on convolutional and capsule features which are extracted simultaneously and were time-consuming.

Alom et al. [13] Presented Inceptional Recurrent Residual Convolutional Neural Network (IRRCNN) for classification of breast cancer images. The developed IRCNN method improved the performance of invariant binary and multiple classifications for breast cancer detection by magnifying the features of the image. In the IRRCNN method, when single-center patches were present it showed the lowest accuracy for binary and multiple classifications.

Beltran-Perez et al. [14] Developed generalized Multi-Scale Radial Basis Function (MSRBF) and Discrete Cosine Transform (DCT) for the detection of breast cancer. The MSRBF combined with the DCT method extracted features of higher quality from black and white images as well as color images. The performance of the developed method was lesser in terms of lesion distinction, sensitivity and detection rate because of higher similitude between healthy tissue and abnormal tumor.

Sudharshan et al. [15] Established deep transfer network with the deep convolutional generative adversarial network for the detection of breast cancer. The developed method solved the problem of class imbalance and worked well with a different types of microscopic images for breast cancer detection. The developed method during training was failed to generate higher quality data due to the lesser number of minority samples.

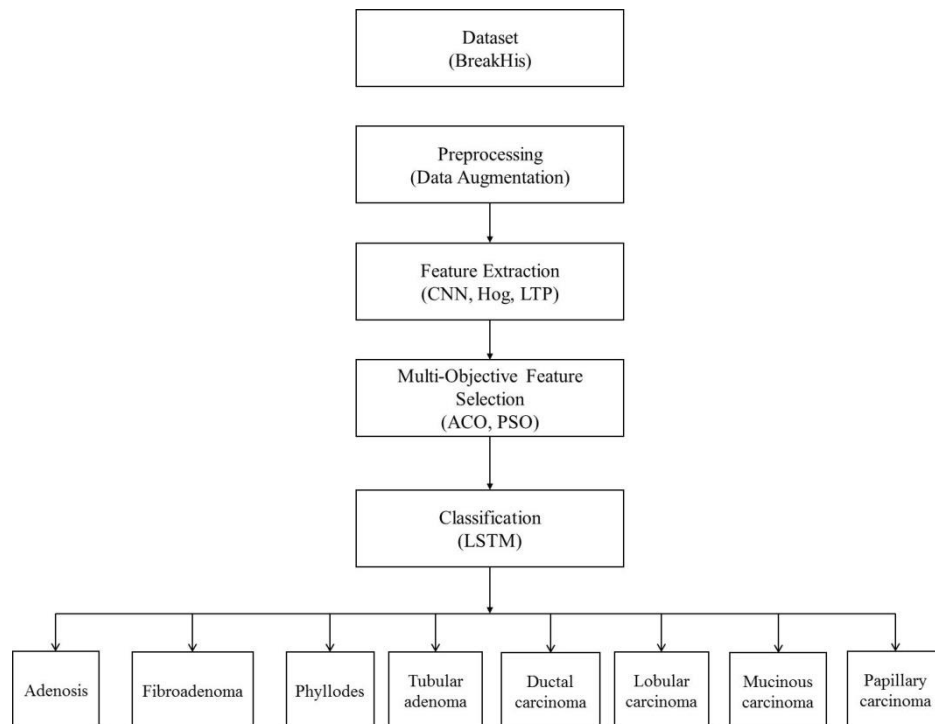


Figure. 1 Block diagram of proposed Multi-Objective feature-selection method used for breast cancer detection

### 3. Proposed methodology

Breast cancer patients suffer from irreversible conditions which lead to a higher mortality rate because of late treatment and diagnosis, therefore early detection and diagnosis are necessary to reduce the rate of mortality

In this research, a multi-objective feature selection method that uses ACO with PSO was proposed, for the identification of breast cancer. The BreakHis dataset is used in the proposed method which includes microscopic biopsy images of malignant tumors and benign tumors. The BreakHis images are made to undergo pre-processing for analysis of the internal morphological features of the image by color variation and data augmentation. Then, the CNN, HoG, and LTP are utilized to remove noises and extracts CNN, histogram and texture features. After extracting the features, the features are selected by the proposed feature selection method by using ACO with PSO to minimize the number of features and to reduce the error rate of classification. The selected features are forwarded to the LSTM network for the process of classification. The block diagram of the proposed breast cancer detection method is shown in Fig. 1.

#### 3.1 Dataset

The Breast Cancer Histopathological Image classification (BreakHis) dataset is used in the

proposed method, which includes microscopic images of malignant tumors and benign tumors. In a malignant tumor, it includes 4 classes as Ductal Carcinoma, Lobular Carcinoma, Mucinous Carcinoma and Papillary Carcinoma. In benign tumor, it includes Adenosis, Fibroadenoma, Phyllodes-tumor, and Tubular-adenoma and the benign tumors are the slower-growing tissue of cancer and remain as localized. The BreakHis consists of 9109 image of breast tumor collected from 82 patients by utilizing various magnifying factors such as 40x, 100x, 200x and 400x. It includes 5429 samples of malignant images and 2480 samples of benign images of 700×460 pixels with three channels of RGB and 8-bit of depth in every channel. Histological benign is a term that refers to a lesion which will not match with the conditions of malignant such as cellular-atypia, disruptions in the membrane of basement, mitosis, metastasize, etc.

#### 3.2 Preprocessing

The image data of the BreakHis dataset undergoes pre-processing, as the microscopic breast images are stained in H-E which helps the doctors to analyze the internal morphology of tissues. But, the variation of color occurs because of differences in staining procedures, which affects the training process of the model. So, the images are pre-processed in the proposed multi-objective method by imposing the color characteristics of one image with another to achieve a correlation of color by selecting

a particular source of the image. To overcome the problem of overfitting, data augmentation is performed such as flipping transformation, random rotation and shearing transformation. The microscopy images are repeated randomly without any fixed angles. The images are flipped horizontally and vertically. The vertical flipping is similar to rotating the image by 180 degrees and then the horizontal flipping is performed. The shearing transforms the part of the images like a parallelogram or slants the shape of the image to create stretches. The pre-processed images of the BreakHis dataset are shown in Fig. 3.

### 3.3 Feature extraction

After pre-processing the microscopy image data, the process of feature extraction takes place. The features are extracted by using CNN, Histogram of Gradients (HoG) and Local Ternary Pattern (LTP) to detect the salient regions from the images. The feature extraction reduces the number of features in the BreakHis dataset by providing new features that improve the accuracy and increase the speed of training.

#### 3.3.1. Convolutional neural network (CNN)

A CNN model includes 4 types of layers that are implemented as follows. The input preprocessed images are passed into convolutional layers and estimated the number of neurons connected with the regions of input images. Every neuron is estimated by dot product among smaller weights and regions connected to the volume of input images. Then, the activation function is performed to verify the neurons are correct or not by ReLU layers it will not change the dimension of the input images. Further, the pooling layer is utilized decrease the effect of noise among the extracted features. Finally, the higher-level features are obtained by using Fully-Connected (FC) layer. In the proposed method, the pre-trained deep CNN models like VGG19 are utilized for feature extraction. The VGG19 network includes 16 convolutional layers, 19 learnable weight layer and 3 FC layer with the soft-max function. The VGG19 model includes convolutional, pooling, ReLU, normalization, and FC layer. The convolutional layer extracts the local features from the images as shown in Eq. (1).

$$g_i^L = b_i^L + \sum_{j=1}^{m_1(L-1)} \psi_{i,j}^L \times h_j^{L-1} \quad (1)$$

where,  $g_i^L$  is the output layer of  $L$ ,  $b_i^L$  is the base value,  $\psi_{i,j}^L$  is the filter connection with

$i^{th}$ ,  $j^{th}$  feature map and  $h_j$  is the output layer of  $L - 1$ . The pooling layer extracts maximum responses from the lesser convolutional layer to reduce unwanted features and it solves the problem of overfitting which is explained through Eqs. (2-4).

$$m_1^L = m_1^{L-1} \quad (2)$$

$$m_2^L = \frac{m_2^{L-1} - F(L)}{S^L} + 1 \quad (3)$$

$$m_3^L = \frac{m_3^{L-1} - F(L)}{S^L} + 1 \quad (4)$$

where,  $S^L$  are the strides which are parameters of neural network that changes the movement of images,  $m_1^L$ ,  $m_2^L$ , and  $m_3^L$  are the filter of feature maps, the other layers like ReLU and FC are explained in Eqs. (5, 6).

$$Re_i^L = \max(h, h_i^{L-1}) \quad (5)$$

$$FC_i^L = f(z_i^L) \text{ with } z_i^L = \sum_{j=1}^{m_1(L-1)} \sum_{r=1}^{m_2^{L-1}} \sum_{s=1}^{m_3^{L-1}} w_{i,j,r,s}^L (FC_i^{L-1})_{r,s} \quad (6)$$

where,  $Re_i^L$  is the ReLU layer,  $h$  is the output layer,  $FC_i^L$  is the FC layer that follows pooling and convolutional layer that performs activation FC layer for deeper feature extractions.

#### 3.3.2. Histogram of gradients (HoG)

The process of feature extraction takes place by using a Histogram of Oriented Gradient (HOG) to detect salient regions from the image. The histogram counts the number of occurrences of gradients orientation in the region of local spatial of the image which is called a cell. To extract the histogram features from images the gradients are calculated, followed by creating the histogram of orientation at every cell. The histograms obtained from every cell are normalized, which provides the HOG descriptor of particular blocks. The process involved in the HOG feature extraction is described below.

Initially, the RGB frames are converted to grayscale, followed by computation of gradients by convoluting the images into horizontal and vertical mask such as  $[-1 \ 0 \ 1]$  and  $[-1 \ 0 \ 1]^T$ . The gradients of horizontal and vertical masks are represented as shown in Eqs. (7, 8).

$$G_x(x, y) = [-1 \ 0 \ 1] * I(x, y) \quad (7)$$

$$G_y(x, y) = [-1 \ 0 \ 1]^T * I(x, y) \quad (8)$$

where,  $G(x, y)$  is the gradients of a horizontal and vertical mask,  $I(x, y)$  is the preprocessed image \* is the convolution, the  $\theta$  is the orientation of every pixel computed by using a ratio of gradients in horizontal and vertical directions as shown in Eq. (9).

$$\theta(x, y) = \arctan \frac{G_x(x, y)}{G_y(x, y)} \quad (9)$$

Further, every block is divided into  $M \times M$  cells, where  $M \leq N$ . For every cell, pixels will be calculated by a weighted vote which is the gradient of magnitude at every pixel and the votes are accumulated in the bins of orientation. The bins  $L$  are spaced between  $0^0$  to  $180^0$  for the unsigned gradients or  $0^0$  to  $360^0$  for signed gradients and  $l_{th}$  value of bins as shown in Eq. (10).

$$\Omega_l(x, y) = \begin{cases} G(x, y) & \text{if } \theta(x, y) \in \text{bin}_l \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

where,  $G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}$  is the magnitude of gradients at the pixels  $(x, y)$ . The HOG features are extracted from every cell and the extracted features are equal to a number of bins and features are normalized as shown in Eq. (11).

$$f_{cell} = \frac{\sum_{(x, y) \in \text{Block}} G(x, y) + \epsilon}{(\sum_{(x, y) \in \text{Block}} G(x, y) + \epsilon)} \quad (11)$$

where,  $\epsilon$  is an immeasurably small quantity of normalized feature of each cell in blocks from  $d$  dimensions of vector which are similar to product numbers of the oriented bin and overall cell in blocks that is last descriptors of the block.

### 3.3.3. Local ternary pattern (LTP)

The LTP is a three-value texture operator which is an efficient and simple descriptor for describing the features. The LTP labels the pixels of the image with a threshold value by using the centered value of every neighborhood pixel multiplied by 2 and adding to generate the new label. The threshold  $t$  will be defined and pixel values within the range of  $-t$  to  $+t$  are considered that assigns the value of zero to pixels. The value of 1 will be assigned to the pixels if it is higher than the threshold value and value of -1 if it is lesser than the centered pixels value. The evaluation of the LTP operator is explained in Eq. (12).

$$LTP(i) = \begin{cases} 1 & \text{if } p_i - p_c \geq t \\ 0 & \text{if } |p_i - p_c| < t \\ -1 & \text{if } p_i - p_c \leq -t \end{cases} \quad (12)$$

where,  $t$  is the user-specified threshold,  $p_i$  is the pixel values in neighbor and  $p_c$  is the central pixel value. The LTP obtains the texture operators which are less sensitive to noise because the texture features are not be based on the value of centered pixels and not strictly invariant to transformations of gray level. The negative values of LTP are changed to 1. The information of features from D-CNN and LTP are fused into one matrix to get a feature vector that is fed to Region-based Convolutional Neural Network (R-CNN) for the classification. The feature extraction process removes noise and extracts robust features such as layers feature from CNN, histogram features from Hog and texture features from LTP.

## 3.4 Multi-objective feature selection by using ACO with PSO

After extracting the features, the proposed multi-objective feature selection takes place by using ACO with PSO. The multi-objective feature selection method minimizes the number of features and reduces the error rate of classification by selecting optimal features.

### 3.4.1. Ant colony optimization (ACO)

The ACO iteratively finds optimum solutions for the targeted problem with the help of guided search by the movements of a number of ants in the solution space by developing the pheromone details. The ACO algorithm develops the directed paths by  $n$  features, at initial moments the  $m$  ants are placed randomly in  $n$  feature nodes. The tabu list  $tabu_k$  from tabu search algorithm records the nodes, which the ant has visited which is set for every ants. The pheromone concentration of  $\tau_{ij}(0)$  on each of the sides of ants are initialized with 0 and the ants choose the next node based on pheromone concentration. The ants move with the probability of  $P_{ij}^k(t)$  from feature,  $i$  to feature  $j$  during  $t$  iteration as explained in Eq. (13).

$$P_{ij}^k(t) = \begin{cases} \frac{\tau_{ij}^\alpha(t) \eta_{ij}^\beta(t)}{\sum_{s \notin tabu_k} \tau_{is}^\alpha(t) \eta_{is}^\beta(t)}, & j \notin tabu_k \\ 0 & \text{else} \end{cases} \quad (13)$$

where,  $P_{ij}^k(t)$  is the probability of ant movements, The  $\tau_{ij}$  is the concentration of pheromone in the path from  $i$  features to  $j$  features during the iteration, The

$\alpha$  and  $\beta$  are the heuristic factors information and expected heuristic factors that are utilized to distribute the weights of pheromone concentration and heuristic information.  $\eta_{ij}$  is the heuristic details which is  $\frac{1}{d_{ij}}$  and  $d_{ij}$  is the Euclidean distance among two features... When the traversal of ant is completed, the concentration of information on every path is updated by using Eq. (14).

$$\Delta\tau_{ij}^k = \begin{cases} \frac{Q}{L_k}, & (i, j) \in \text{path of } k \\ 0, & \text{else} \end{cases} \quad (14)$$

where,  $\Delta\tau_{ij}^k$  is the pheromone concentration at every path,  $Q$  is the constant and  $L_k$  is the path length of  $k$  ant during traversal. The ACO algorithm selects the features of classifiers and effectively improves the performance of classifiers.

### 3.4.2. Particle swarm optimization (PSO)

PSO is a population based evolutionary computation algorithm that is developed by the simulation of simplified models of behavior of swarms. The PSO algorithm is based on the behavior of swarms such as bird flocking and fish schooling. The PSO includes fitness evaluation function which considers the position of particles and assigns the fitness value. The position having higher value of fitness in the entire run is considered as global best ( $P_g$ ). Every particle will keep track of the highest fitness value and the best location at particular value is considered as personal best ( $P_i$ ). The algorithm includes casting the population of particles over the search space that finds the best solutions which is solved. The PSO performs search operation through the swarm of particles which updates the position during every iteration. To get the optimal solution, every particle moves in the direction of  $P_g$  and  $P_i$  positions in the swarm as shown in Eqs. (15, 16).

$$pbest(i, t) = \arg_{k=1, \dots, t} \min[f(P_i(k))], \quad i \in \{1, 2, \dots, N\} \quad (15)$$

$$gbest(t) = \arg_{i=1, \dots, N, k=1, \dots, t} \min[f(P_i(k))] \quad (16)$$

where,  $pbest(i, t)$  is the personal best value of particles,  $gbest(t)$  is the global best value of  $i$  is the index of particles,  $N_p$  is the total number of particles,  $t$  is the present number of iterations,  $f$  is the fitness function and  $P$  is the position,  $arg$  is the arguments of fitness function and position.. The velocity  $v$  and position of particle  $p$  are updated by using Eqs. (17) and (18).

$$V_i(t+1) = \omega V_i(t) + c_1 r_1 (pbest(i, t) - p_i(t)) + c_2 r_2 (gbest(t) - p_i(t)) \quad (17)$$

$$p_i(t+1) = p_i(t) + V_i(t+1) \quad (18)$$

where,  $V$  is the velocity,  $\omega$  is the inertia weights utilized for balancing the local exploration and global exploration, the  $r_1$  and  $r_2$  are the random variables that are randomly generated in the range of  $[0, 1]$ . The  $c_1$  and  $c_2$  are the positive component parameters called as acceleration co-efficient. The proposed multi-objective feature selection by using ACO with PSO method selects the optimum patch of features in images by minimizing the error rates in classification. In existing hybrid ACO with PSO algorithm for feature selection the performance of the best subset selected was not significantly improved for several generations. Further, when the fitness value was not changed by hybrid ACO-PSO [17] algorithm the process of searching was reduced. Whereas, in case of proposed multi-objective feature selection based ACO-PSO method when the fitness value will not change, the PSO starts the process of searching and identifies the best subsets found by the population of ants which improves the performance of selection.

### 3.5 Classification

The selected features are forwarded to the LSTM network for the process of classification. The proposed multi-objective method classifies the data set into 8 classes such as Ductal-Carcinoma, Lobular-Carcinoma, Mucinou-Carcinoma, Papillary-fibroadenoma, carcinoma, adenosis, phyllodes-Tumor and tubular-adenoma. The LSTM is architecture of Recurrent Neural Network (RNN) and a type of neural network which models the temporal dependencies. The direct cyclic connection among the units of RNN stores the hidden states of neural network in internal region and helps to model the temporal dynamic behavior of images. The LSTM is the upgraded version of RNN with three gates such as output gate, input gate and forget gate. LSTM is the hidden layer neural network consisting of 100 nodes in every layer. The LSTM determines the dependencies of long terms temporal direction of neural network with the three gates. The LSTM network is easy to optimize because the gates enable the input features to be propagated via the hidden layers without affecting the output. The LSTM network also effectively deals with the gradient of vanishing problem because it frees the memory location in the temporal dimensions which are not helpful in detecting the labels of final classification.

The hidden state  $h_t$  for  $t^{th}$  word in the sub-path is the function of previous state  $h_{t-1}$  and present word  $x_t$ . The input is linearly transformed by the weight matrix and squashed non-linearly by the activation function as shown in Eq. (19).

$$h_t = f(W_{in}x_t + W_{rec}h_{t-1} + b_h) \quad (19)$$

Where,  $W_{in}$  and  $W_{rec}$  are the weight matrix for the input and recurrent connections, respectively. The  $b_h$  is the bias term for the vector in hidden state and  $f_h$  is the activation function of non-linear layer.

The LSTM based RNN includes four components that are input gate  $i_t$ , output gate  $o_t$ , forget gate  $f_t$  and memory cell  $c_t$ . The three adaptive gates such as  $i_t$ ,  $f_t$  and  $o_t$  are dependent on the previous  $h_{t-1}$  layer. The current input  $x_t$  and the extracted vectors of features  $g_t$  are calculated by using Eqs. (20-24).

$$i_t = \sigma(W_i \cdot x_t + U_i \cdot h_{t-1} + b_i) \quad (20)$$

$$f_t = \sigma(W_f \cdot x_t + U_f \cdot h_{t-1} + b_f) \quad (21)$$

$$o_t = \sigma(W_o \cdot x_t + U_o \cdot h_{t-1} + b_o) \quad (23)$$

$$g_t = \tanh(W_g \cdot x_t + U_g \cdot h_{t-1} + b_g) \quad (24)$$

Where,  $i_t$  is the input gate,  $f_t$  is the forget gate,  $o_t$  is the output gate and  $g_t$  is the extracted feature gate. The  $W$  is the weights of particular gate and the current memory of cells  $c_t$  is the combination of previous cell content  $c_{t-1}$  and the content of candidate  $g_t$  that is weighted by input gate  $i_t$  and forget gate  $f_t$ . The current memory of cell  $c_t$  is calculated by using Eq. (25).

$$c_t = i_t \otimes g_t + f_t \otimes c_{t-1} \quad (25)$$

The output of LSTM units is the hidden states recurrent network that is calculated by using Eq. (26).

$$h_t = o_t \otimes \tanh(c_t) \quad (26)$$

Where, the  $\sigma$  is the sigmoid function,  $\otimes$  is the element wise multiplication.

#### 4. Experimental results and discussion

The detection of cell nuclei is difficult in the analysis of Histopathological images such as target cells which are in various states. The smaller cells will be surrounded by the background clutters made-up of histopathological structures such as collagen, capillaries etc., along with irrelevant visual aspects

and artefacts which occurs during the acquisition of images. The present research performs effective analysis on BreakHis dataset for the nuclei identification. The proposed method is evaluated by Python 3 interpreter programming language in windows 10, i7 core processor, 16 GB RAM, and 6 GB 2080Ti NVIDIA GTX edition GPU environments. The proposed method considered 70% of images as training data and 30% of images for testing process. The performance metrics that were considered to evaluate the proposed method, and the quantitative and comparative analysis of the proposed multi-objective feature selection method are explained in this section.

##### 4.1 Performance metrics

The parameters considered to evaluate the proposed multi-objective feature selection by using ACO with PSO for the detection of breast cancer, is explained in the following:

- Accuracy: Accuracy is defined as the measure utilized to determine the exactness of model. The accuracy Eq. is explained in Eq. (27).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (27)$$

- Precision: Precision is defined as the ratio of truly predicted observation for positives to the overall predicted observation for positives. The precision equation is explained in Eq. (28).

$$Precision = \frac{TP}{TP+FP} \quad (28)$$

- Sensitivity: Sensitivity is defined as the ratio of truly predicted as the fault-modules which is explained in Eq. (29).

$$Recall = \frac{TP}{TP+FN} \quad (29)$$

- F-Measure: F-Measure is the calculation of accuracy test and is defined as mean of precision and recall. The F-Measure is explained in Eq. (30).

$$F - Measure = \frac{2PR}{P+R} \quad (30)$$

- Specificity: Specificity is defined as the ratio of negatives that are identified correctly which is explained in Eq. (31).

$$Specificity = \frac{TN}{TN+FP} \quad (31)$$

## 4.2 Quantitative Analysis

The result obtained by the proposed ACO with PSO method is shown in Table 1. Table 1 includes evaluation results for breast cancer detection in terms of accuracy, precision, recall and f-measure also shows the performance estimation of proposed ACO-PSO method used for the detection of breast cancer in terms of accuracy, precision, F1-score, recall and specificity. Accuracy is the measure utilized to determine the exactness of the proposed model used in the detection of breast cancer. The proposed multi-objective feature selection by using ACO with PSO method achieved an accuracy of 95.72% for the detection of breast cancer. The precision achieved by the proposed method is 94.52%, achieved F1-Score of 91.57%, sensitivity of 97.35% and specificity of 94.54%. The time taken to complete the process of detection by proposed ACO-PSO algorithm is 50 seconds. The graphical representation of quantitative analysis of the proposed multi-objective feature selection by using ACO with PSO method is shown in Fig. 2.

## 4.3 Comparative analysis

The comparative analysis of the proposed ACO with PSO method, used for breast cancer identification, is carried out and the values are tabulated as described in Table 2.

The existing techniques such as FCN-BiLSTM [11], CNN [12] and MSRBF-DCT [14] are compared with the proposed approach. The performance metrics or parameters considered in the existing FCN-BiLSTM [11] and CNN [12] is accuracy, sensitivity and specificity for the analysis of results. The experimental result result of FCN-BiLSTM[11] was conducted in Intel (R) Xeon (R) Gold 6132 @2.60 GHz CPU with 14 core and NVIDIA Quadro P6000 GPU. The CNN [12] is evaluated by using Python 3.6 in PyTorch 1.1.0 and GeForce GTX 1080Ti GPU to develop the model. Table 2 shows the comparison results of the proposed method with

existing methods. Table 2 shows the comparative data of the proposed method by using ACO with PSO for breast cancer detection with existing methods such as FCN-BiLSTM [11] and CNN [12].

The existing FCN-BiLSTM method showed an accuracy of 93.68%, Sensitivity of 95.62%, and Specificity of 89.44% for breast cancer detection. Similarly, the existing CNN method showed an accuracy of 92.71%, sensitivity of 92.15% and specificity of 93.27% for the detection of breast cancer. The Existing MSR-BF [14] showed accuracy of 93.99%, sensitivity of 92.65% and specificity of 94.35%. The performance of the existing method was lesser because of higher similitude among healthy tissue and abnormal tumor. The existing FCN-BiLSTM [11] method consumed more time in extracting the features due to the limitation of space. The existing CNN [12] method was very time consuming in order to classify based on the capsule features and convolution features extracted in two parallel channels. But, the time taken to complete the process was not described in values in both te existing research. Whereas, the proposed feature selection method achieved higher performance than existing methods in detecting breast cancer and consumed lesser amount of time. The proposed ACO with PSO method requires only primitive and simple mathematical operators, and is computationally

Table 1. Performance evaluation of the proposed method by using ACO with PSO for breast cancer detection in terms of accuracy, precision, F1-score, Sensitivity and specificity

Metrics	Proposed ACO-PSO
Accuracy	95.72
Precision	94.52
F1-Score	91.57
Sensitivity	97.35
Specificity	94.54
Time (sec)	50

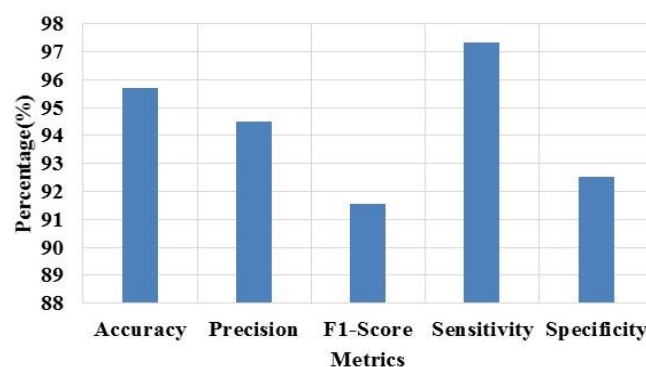


Figure. 2 The graph of quantitative analysis of the proposed ACO with PSO for breast cancer detection



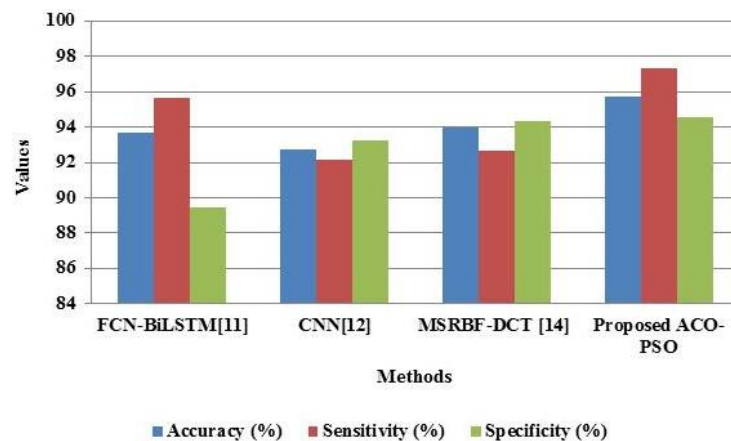


Figure. 3 The graph of comparative analysis of proposed multi-objective method with existing methods for breast cancer detection

Table 2. Comparative table of the proposed method with the existing methods in terms of Accuracy, Sensitivity, and Specificity.

Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)
FCN-BiLSTM[11]	93.68	95.62	89.44
CNN[12]	92.71	92.15	93.27
MSRBF-DCT [14]	93.99	92.65	94.35
Proposed ACO-PSO	95.72	97.35	94.54

inexpensive in memory and time. The graphical representation of comparative analysis of proposed multi-objective feature selection method by using ACO with PSO for breast cancer detection is shown in Fig. 3.

## 5. Conclusion

The existing deep learning methods for breast cancer detection utilized the outline information for effective performance. But, the existing methods showed lesser performance because they utilized an insufficient amount of semantic details. In order to solve such an issue, a multi-objective feature selection method, consisting of ACO with PSO, was proposed for the detection of breast cancer. The proposed method utilized the BreakHis dataset and data augmentation are applied to analyze the internal morphology features of the image. To overcome the problem of overfitting, data augmentation is performed such as flipping transformation, random rotation and shearing transformation. Then, the features are extracted by using Convolutional Neural Network CNN, HoG and LTP. The feature extraction reduces the number of features in the BreakHis dataset by providing the new features which improve

the accuracy. The features are selected by the proposed multi-objective feature selection method by using ACO with PSO. The selected features are forwarded to the LSTM network for the process of classification. The proposed multi-objective feature selection method by using ACO with PSO method achieved higher performance in detecting breast cancer than an existing method by achieving an accuracy rate of 95.72%, whereas the existing method showed an accuracy rate of 93.68%. In future work, image segmentation is carried out to effectively identify the malignant region position.

## Conflicts of interest

The authors declare no conflict of interest.

## Author contributions

The paper background work, conceptualization, methodology, dataset collection, implementation, result analysis and comparison, preparing and editing draft, visualization have been done by first author. The supervision, review of work and project administration, have been done by second author.

## References

- [1] R. Ruskone, S. Airault, and O. Jamet, "Vehicle Detection on Aerial Images", *International Journal of Intelligent Engineering and Systems*, Vol. 1, No. 1, pp. 123-456, 2009.
- [2] Y. Li, J. Wu, and Q. Wu, "Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning", *IEEE Access*, Vol. 7, pp. 21400-21408, 2019.
- [3] Z. Zeng, W. Xie, Y. Zhang, and Y. Lu, "RIC-Unet: An improved neural network based on Unet for nuclei segmentation in histology

- images”, *IEEE Access*, Vol. 7, pp. 21420-21428, 2019.
- [4] S. Mojriari, G. Pinter, J. H. Joloudari, I. Felde, A. S. Gali, L. Nadai, and A. Mosavi, “Hybrid machine learning model of extreme learning machine radial basis function for breast cancer detection and diagnosis; a multilayer fuzzy expert system”, In: *Proc. of RIVF International Conference on Computing and Communication Technologies (RIVF)*, pp. 1-7, 2019.
- [5] A. Ghasemzadeh, S. S. Azad, and E. Esmaeili, “Breast cancer detection based on Gabor-wavelet transform and machine learning methods”, *International Journal of Machine Learning and Cybernetics*, Vol. 10, No. 7, pp. 1603-1612, 2019.
- [6] K. George and P. Sankaran, “Computer assisted recognition of breast cancer in biopsy images via fusion of nucleus-guided deep convolutional features”, *Computer Methods and Programs in Biomedicine*, Vol. 194, pp. 105531, 2020.
- [7] E. B. Kablan, H. Dogan, M. E. Ercin, S. Ersoz, and M. Ekinici, “An ensemble of fine-tuned fully convolutional neural networks for pleural effusion cell nuclei segmentation”, *Computers & Electrical Engineering*, Vol. 81, pp. 106533, 2020.
- [8] T. Iesmantas, A. P. Taraseviciene, and K. Sutiene, “Enhancing multi-tissue and multi-scale cell nuclei segmentation with deep metric learning”, *Applied Sciences*, Vol. 10, No. 2, pp. 615, 2020.
- [9] Y. Kong, G. Z. Genchev, X. Wang, H. Zhao, and H. Lu, “Nuclear Segmentation in Histopathological Images Using Two-Stage Stacked U-Nets with Attention Mechanism”, *Frontiers in Bioengineering and Biotechnology*, Vol. 8, pp. 1246, 2020.
- [10] K. George, S. Faziludeen, and P. Sankaran, “Breast cancer detection from biopsy images using nucleus guided transfer learning and belief based fusion”, *Computers in Biology and Medicine*, Vol. 124, pp. 103954, 2020.
- [11] N. Wahab, A. Khan, and Y. S. Lee, “Transfer learning based deep CNN for segmentation and detection of mitoses in breast cancer histopathological images”, *Microscopy*, Vol. 68, No. 3, pp. 216-233, 2019.
- [12] U. Budak, Z. Cömert, Z. N. Rashid, A. Şengür, and M. Çıbuk, “Computer-aided diagnosis system combining FCN and Bi-LSTM model for efficient breast cancer detection from histopathological images”, *Applied Soft Computing*, Vol. 85, pp. 105765, 2019.
- [13] P. Wang, J. Wang, Y. Li, P. Li, L. Li, and M. Jiang, “Automatic classification of breast cancer histopathological images based on deep feature fusion and enhanced routing”, *Biomedical Signal Processing and Control*, Vol. 65, pp. 102341, 2021.
- [14] M. Z. Alom, C. Yakopcic, M. S. Nasrin, T. M. Taha, and V. K. Asari, “Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network”, *Journal of digital imaging*, Vol. 32, No. 4, pp. 605-617, 2020.
- [15] C. B. Perez, H. L. Wei, and A. R. Solis, “Generalized multiscale RBF networks and the DCT for breast cancer detection”, *International Journal of Automation and Computing*, Vol. 17, No. 1, pp. 55-70, 2019.
- [16] P. J. Sudharshan, C. Petitjean, F. Spanhol, L. E. Oliveira, L. Heutte, and P. Honeine, “Multiple instance learning for histopathological breast cancer image classification”, *Expert Systems with Applications*, Vol. 117, pp. 103-111, 2020.
- [17] K. Menghour and L. S. Meslati, “Hybrid ACO-PSO based approaches for feature selection”, *International Journal of Intelligent Engineering and Systems*, Vol. 9, No. 3, pp. 65-79, 2016.