

An Improved Feature Selection Algorithm for Medical Image Classification

Mustafa S. Kadhm^{1*}, Israa B. Mohammed², Amani Y. Noori¹

¹ Department of Computer, Collage of Basic Education, Mustansiriyah University, Baghdad 10052, Iraq

² College of Laser and Optoelectronic Engineering, University of Technology, Baghdad 10001, Iraq

Corresponding Author Email: mst.salam@uomustansiriyah.edu.iq

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.310520>

ABSTRACT

Received: 12 February 2026

Revised: 10 April 2026

Accepted: 25 April 2026

Available online: 31 May 2026

Keywords:

breast ultrasound image classification, feature selection, Grey Wolf Optimizer, Cauchy Mutation, chaotic initialization, texture features, machine learning

Accurate classification of breast ultrasound images remains challenging because lesion appearance is heterogeneous, and handcrafted descriptors often contain redundant information. This study develops Grey Wolf Optimizer-Cauchy Mutation (GWO-CM), a feature-selection method that augments the GWO with hybrid Tent-Logistic chaotic initialization and Cauchy Mutation. The method is embedded in a breast ultrasound classification pipeline comprising image preprocessing, gray-level co-occurrence matrix and statistical-moment feature extraction, wrapper-based feature selection, and supervised classification of normal, benign, and malignant images. The hybrid initialization is intended to diversify the search population, whereas CM is applied to reduce premature convergence during feature-subset optimization. The method was evaluated on the BUSI dataset, which contains 780 ultrasound images. Feature selection and classifier fitting were performed exclusively within the training data, and performance was assessed on an independent test set using accuracy, precision, recall, and F1-score. Under the reported experimental setting, the best configuration achieved 99.12% accuracy, 99% precision, 100% recall, and 99% F1-score. Compared with standard GWO, the proposed method improved accuracy from 96.49% to 99.12%. These results suggest that chaotic population initialization and CM can improve wrapper-based feature selection for BUSI image classification. However, the results should be confirmed using a clearly specified stratified split or nested cross-validation, class-wise metrics, and external validation before clinical generalization is claimed.

1. INTRODUCTION

Breast cancer is a disease in which aberrant breast cells proliferate and form tumors [1]. The cancer can be fatal and spread throughout the body if left untreated. The World Health Organization (WHO) estimates that breast cancer will kill 670,000 people worldwide in 2022. Women are the most affected by breast cancer, who have no identifiable risk factors other than their gender and age [2]. Effective treatments for early and prompt diagnosis, as well as thorough treatment, rehabilitation, and palliative care, are essential for lowering the burden of breast cancer and achieving and maintaining optimal functioning and well-being [3].

The breast cancer should be detected in the early stage to increase the survival chance of patients. The late-stage breast cancer poses significant hurdles for treatment, highlighting the vital relevance of early identification [4]. The primary approaches for detecting and treating early breast cancer are ultrasound, thermography, and mammography [5]. It has been shown that radiologists who scan medical imaging quickly during long workdays are more likely to make analytical mistakes and encounter unstable results due to eye tiredness [4].

Machine learning (ML) is successfully applied in various medical fields [6]. Many ML classifiers support researchers

and systems to achieve efficient classification results. Linear Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Artificial Neural Networks (ANN) are the common powerful classifiers that are employed in various medical field systems [7]. On the other hand, metaheuristic algorithms helped to rise the results of the existing systems by selecting the desired features in different medical systems [8-10]. One of the efficient and effective metaheuristic algorithms that is applied for feature selection in various optimization problems is Grey Wolf Optimizer (GWO) [11]. GWO provides competitive results among other metaheuristic algorithm via global exhaustive search ability and the high convergence speed [12]. GWO successfully applied to solve many medical field issues and enhance the results via the feature selection process and using various medical data from UCI warehouse [13, 14].

The proposed work enhances the breast cancer diagnoses by using a new accurate feature selection algorithm based on GWO, chaotic maps, and Cauchy Mutation (CM). The proposed algorithm GWO-CM raise the prediction rate of the breast cancer detection system by address the issue of selecting only the relative and informative features that identify the cancer. A standard dataset BUSI is used to evaluate the system and the effectiveness of the proposed system compared using different ML methods and metaheuristics algorithms. Besides,

several performance measures are used accuracy, f-score, precision, and recall for evaluation.

2. RELATED WORKS

Number of works have been done in the field of breast cancer detection. Some works, employee various deep neural networks models and other employee the ML methods for classification purpose.

Halder et al. [15] presented a robust breast cancer classification system that employs many efficient approaches. The authors employ a number of classifiers (Random Forest (RF), Gradient Boosting, AdaBoost, XGBoost, Naive Bayes, and SVM), a variety of feature selection techniques (methods, Chi-Square, Recursive Feature Elimination with Cross-Validation (RFECV), and Ridge Regression), as well as thorough preprocessing (data label encoding, data imputation, and data normalization). Thirty percent of the Wisconsin Original Breast Cancer (WOBC) and seventy percent of the Wisconsin Prognostic Breast Cancer (WPBC) datasets were used for testing and training, respectively. For WOBC and WPBC, the system's classification accuracy was 92.33% and 89.32%, respectively.

Alotaibi et al. [16] proposed an ultra sound breast cancer classification approach based on CNN and image fusion. The approach uses three preprocessing steps for the input images starting from noise removing by (block matching technique), highlight the ROI, and RGB fusion in order to enhance the performance of the used deep learning models. The proposed approach achieved a high classification accuracy using the proposed preprocessing approach using three breast cancer datasets (BUSI 780 images, B 180 images, and KAIMRC 5693 images) and VGG19 transfer-learning model.

Benghazouani et al. [17] suggested innovative approach for optimizing breast cancer diagnoses to enhance the accuracy. The proposed approach combines the nature-inspired optimization (NIO) algorithms (GWO, Jaya, Genetic Algorithm, Harris Hawks, SALP Swarm, Whale Optimization, Sine Cosine, Flower Pollination, and Cuckoo Search) for feature selection with soft voting ML classifiers (Decision Tree, Gradient Boosting, Logistic Regression, SVM, and Gaussian Naive Bayes). The proposed approach achieves the highest results using the Jaya algorithm within 99.6% accuracy, 99.6% F1-score, 99.21% precision, 100% recall, and 99.6% AUC.

Ismael et al. [18] proposed a new model based on deep neural networks and GWO called DeepWolfNet for enhancing the medical disease diagnosis. The proposed model uses two datasets namely BUSI, and the BreakHis-400X within images that classified into benign, malignant, and normal. Histogram of Oriented Gradients (HOG) is used first to extract the required features from the input images then GWO algorithm is used to select the most relevant features to improve the system's effectiveness and accuracy. After that, the selected features went to the DeepWolfNet model for training. The proposed model achieved 95.23% accuracy using BUSI dataset, and 94.11% using BreakHis-400X dataset.

Mohammed et al. [19] employed ML (for feature extraction and selection) and deep learning (for data analysis) algorithms to detect breast cancer based on clinical data and histological whole slide pictures. In their proposed work, Histopathological Whole Slide Imaging (WSI) is used for feature extraction, Particle Swarm Optimization (PSO) and

Principal Component Analysis (PCA) are used for feature selection process and several classifiers (Naïve Bayes, SVM, RF, Logistic Regression and KNN). The proposed achieved the highest accuracies of 95.7% and 97.2% using PCA and PSO with RF classifier and Wisconsin Diagnostic Breast Cancer (WDBC) dataset.

Yavuz et al. [20] proposed a combined feature selection approach called JASAL using filter and wrapper features selection methods for high dimensional medical classification. First, the relief filter is used for image preprocessing, then a hybrid binary optimization includes the Sine-Cosine Algorithm (SCA) and JAYA algorithms are used for feature selection. In the proposed JASAL approach, the Levy flight strategy is used to address the local optimum problem in JAYA algorithm. In order to evaluate the performance of JASAL approach eight medical datasets (seven high dimensional medical datasets) are used. The JASAL achieved a highest result comparing to eleven metaheuristic algorithms using various metrics (accuracy, number of selected features, and standard deviation).

Mustafa et al. [21] proposed a hybrid optimization approach and a clear deep learning model for breast cancer screening. The suggested work aims to improve model convergence and classification appreciation by combining the MobileNet deep learning model with two optimization algorithms: the Firefly Algorithm (FLA) and the Dingo Optimization Algorithm (DOA). The fusion test had an accuracy of 98.96% when using the DOA and MobileNet models, while the FLA and MobileNet models had 98.06% and 95.44% accuracy on mammographic and ultrasound tests, respectively. Furthermore, Grad-CAM visualizations revealed clinically consistent lesion localization, improving Grad-CAM's interpretability and model diagnostic reliability.

Zafar et al. [22] presented BU-DLNet, a breast ultrasonography (BU) image-based system for women's breast cancer diagnosis that uses deep-learning network selection and feature optimization. Several pre-trained networks are used to extract the deep features from the input BU pictures. On the other hand, SVM calculates the optimal subset of deep features, and ten optimization algorithms are included (generalized normal distribution optimization, Henry gas solubility optimization, rich optimization, equilibrium optimizer (EO), manta-ray foraging optimization, atom search optimization, Harris hawks optimization, path finder algorithm, poor slime mold algorithm, and marine predator algorithm). In addition, the best pre-trained network was chosen using a selection method. The suggested framework achieved the highest accuracy of 96.79% utilizing the EO method and the BU dataset.

Kumar et al. [23] offered EGWO-SVM, an upgraded feature selection algorithm for breast cancer detection that is based on the GWO and SVM. The GWO is enhanced using a weighted position update (Instead of merely average, each wolf's position is updated as a weighted combination of the locations of the three best grey wolves), then combined with SVM for selecting the better tumor features. The suggested approach improves benign and malignant tumor detection accuracy by 98.24%, precision by 97.18%, recall by 97.64%, and F-measure by 97.40% when using the WDBC dataset.

Ayana et al. [24] proposed an accurate ultrasound breast cancer picture categorization using a unique multistage transfer learning (MSTL) algorithm. The proposed MSTL approach employs three optimizers: Adam, Adagrad, and stochastic gradient descent (SGD) in conjunction with three

pre-trained deep learning models: EfficientNetB2, InceptionV3, and ResNet50. The proposed MSTL is evaluated using two standard datasets: Mendeley (20,400 cancer cell pictures and 200 ultrasound images) and MT-Small-Dataset (400 ultrasound images). The best results were obtained using ResNet50 (Adagrad) with 99.061% using the Mendeley dataset and 98.7% using the MT-Small dataset.

Based on a thorough review of the above-related works. This paper focuses on enhancing the performance of the breast cancer system by applying an efficient and accurate feature selection algorithm in order to raise the existing accuracy and speed up the processing time by eliminate the unwanted features.

3. GREY WOLF OPTIMIZER

The Canidae family includes the grey wolf. Being at the top of the food chain, grey wolves are regarded as apex predators. In a pack of wolves, four types of solutions are considered: alpha, beta, omega and delta. Alpha is the wolf's leader, and the less powerful wolves follow him. Delta wolves must be subject to alphas and betas, but they control the omega. Figure 1 depicts the dominance hierarchy of grey wolves [25].

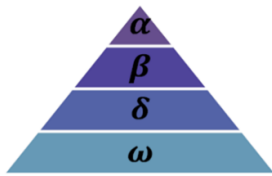


Figure 1. Hierarchy of grey wolf

The hunting mechanism of the wolves begins by dividing the optimization problems into four classes: alpha α , beta β , delta δ , and omega ω . The first three best solutions are considered as α , β , and δ . While ω contains the other solutions (wolves) as illustrated in Figure 2.

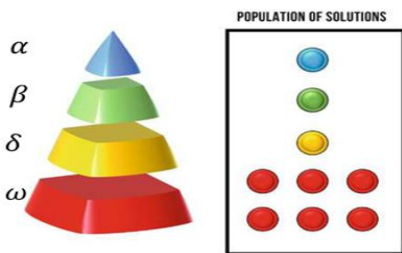


Figure 2. Divide the population into four Grey Wolf Optimizer (GWO) classes

The encircling mechanism is used to hunt pray. The mathematical formula of hunting is shown in Eq. (1):

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \quad (1)$$

where, \vec{X} is the grey wolf position vector, \vec{X}_p is the prey position vector \vec{A} and \vec{D} are coefficient vectors, and t is the current iteration. \vec{A} and \vec{D} vectors calculated using the following formulas:

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (2)$$

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_p(t) - \vec{X}(t) \right| \quad (3)$$

However, the distance of the prey location \vec{C} is calculated using Eq. (4):

$$\vec{C} = 2 \cdot \vec{r}_2 \quad (4)$$

where, r_1 and r_2 are random vectors between [0, 1], \vec{a} component is decreased from 2 to 0 during optimization.

Grey wolves can recognize and encircle their prey. Typically, the alpha leads the hunt. The alpha is responsible for hunting and guiding the other wolves. The beta and delta might also hunt on occasion. However, the optimal (prey) location is known in an abstract search space. Eq. (5) shows the updating formula for each wolf position.

$$\vec{X}(t+1) = \frac{X_1 + X_2 + X_3}{3} \quad (5)$$

where, \vec{X}_1 , \vec{X}_2 and \vec{X}_3 calculated using Eq. (6):

$$\begin{aligned} \vec{X}_1 &= \vec{X}_\alpha(t) - \vec{A}_1 \cdot \vec{D}_\alpha \\ \vec{X}_2 &= \vec{X}_\beta(t) - \vec{A}_2 \cdot \vec{D}_\beta \\ \vec{X}_3 &= \vec{X}_\delta(t) - \vec{A}_3 \cdot \vec{D}_\delta \end{aligned} \quad (6)$$

\vec{D}_α , \vec{D}_β , and \vec{D}_δ calculated using Eq. (7):

$$\begin{aligned} \vec{D}_\alpha &= \left| \vec{C}_1 \cdot \vec{X}_\alpha - \vec{X} \right| \\ \vec{D}_\beta &= \left| \vec{C}_2 \cdot \vec{X}_\beta - \vec{X} \right| \\ \vec{D}_\delta &= \left| \vec{C}_3 \cdot \vec{X}_\delta - \vec{X} \right| \end{aligned} \quad (7)$$

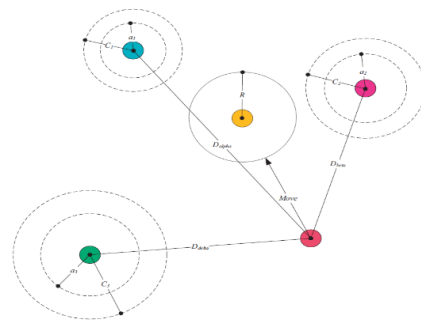


Figure 3. Updating position in Grey Wolf Optimizer (GWO)

```

Initialize the grey wolf population  $X_i (i = 1, 2, \dots, n)$ 
Initialize  $a$ ,  $A$ , and  $C$ 
Calculate the fitness of each search agent
 $X_\alpha$  = the best search agent
 $X_\beta$  = the second best search agent
 $X_\delta$  = the third best search agent
while ( $t < \text{Max number of iterations}$ )
  for each search agent
    Update the position of the current search agent
  end for
  Update  $a$ ,  $A$ , and  $C$ 
  Calculate the fitness of all search agents
  Update  $X_\alpha$ ,  $X_\beta$ , and  $X_\delta$ 
   $t = t + 1$ 
end while
return  $X_\alpha$ 

```

Figure 4. Pseudo code of Grey Wolf Optimizer (GWO) algorithm

The search agent updates its position in a 2D search space according to alpha, beta, and delta as shown in Figure 3. The grey wolves finish the hunt by attacking the prey (exploitation) when it stops moving.

The pseudo code of the GWO algorithm is presented in Figure 4.

4. CAUCHY MUTATION

CM is a random mutation operation based on the Cauchy distribution with a large tail that is used to mutate particles in evolutionary algorithms. This helps the algorithm escape local minima by allowing large jumps more frequently than Gaussian mutation [25].

The probability density function of the Cauchy distribution is given by:

$$f(x; x_o, \gamma) = \frac{1}{\pi \gamma \left[1 + \left(\frac{x - x_o}{\gamma} \right)^2 \right]} \quad (8)$$

where,

x_o : the centre of distribution,

γ : the scale vector.

In CM disturbance, the disturbance vector is typically generated by random sampling from the Cauchy distribution and then added to the current solution to generate a new solution. Therefore, the Cauchy distribution is presented by the following formula:

$$x_{new} = x + \gamma \cdot Cauchy(0,1) \quad (9)$$

where,

x : the original individual (solution),

x_{new} : the mutated individual (solution),

Cauchy(0,1): the Cauchy distribution.

The standard and Cauchy distributions' density function curves are shown in Figure 5.

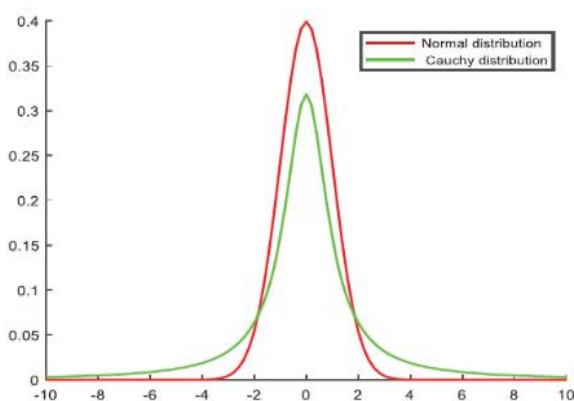


Figure 5. The standard and Cauchy distributions

5. PROPOSED SYSTEM

The proposed system for breast cancer detection consists of five steps that work together to produce the best acceptable results. These stages include preprocessing, feature extraction, feature selection, classification, and evaluation. Figure 6

depicts the main stages of the proposed system.

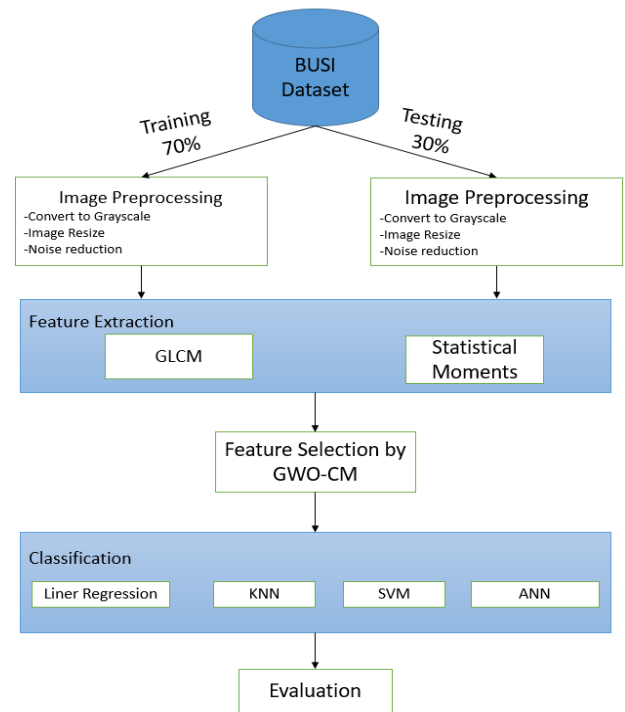


Figure 6. The proposed ultrasound detection system diagram

5.1 Dataset

In the proposed work, a standard breast ultrasound images dataset called BUSI is used since it has been successfully used in various previous works for breast cancer segmentation, detection, classification, and recognition. The dataset has three image classes, which are: normal, benign, and malignant as in Figure 7 [26].

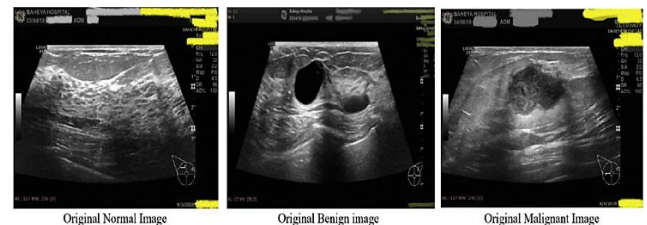


Figure 7. Sample images of BUSI dataset

Table 1. BUSI dataset details

Case	No. of Images
Normal	487
Benign	210
Malignant	133
Total	780

The BUSI dataset collected in 2018 from various women and contain 780 images with resolution of 1280×1024 pixels. All the images are collected and stored as DICOM format from Baheya Hospital. In order to preserve patient privacy and conditionality, all the patient information is removed prior to dataset publication. Besides, to ensure the labeling reliability, medical specialist and radiologist reviewed the images and corresponding labels in the dataset. Details of the BUSI dataset are illustrated in Table 1.

5.2 Pre-processing

In the pre-processing stage, all input images are converted into Grayscale and store it in PNG format then resize to $500 \times$

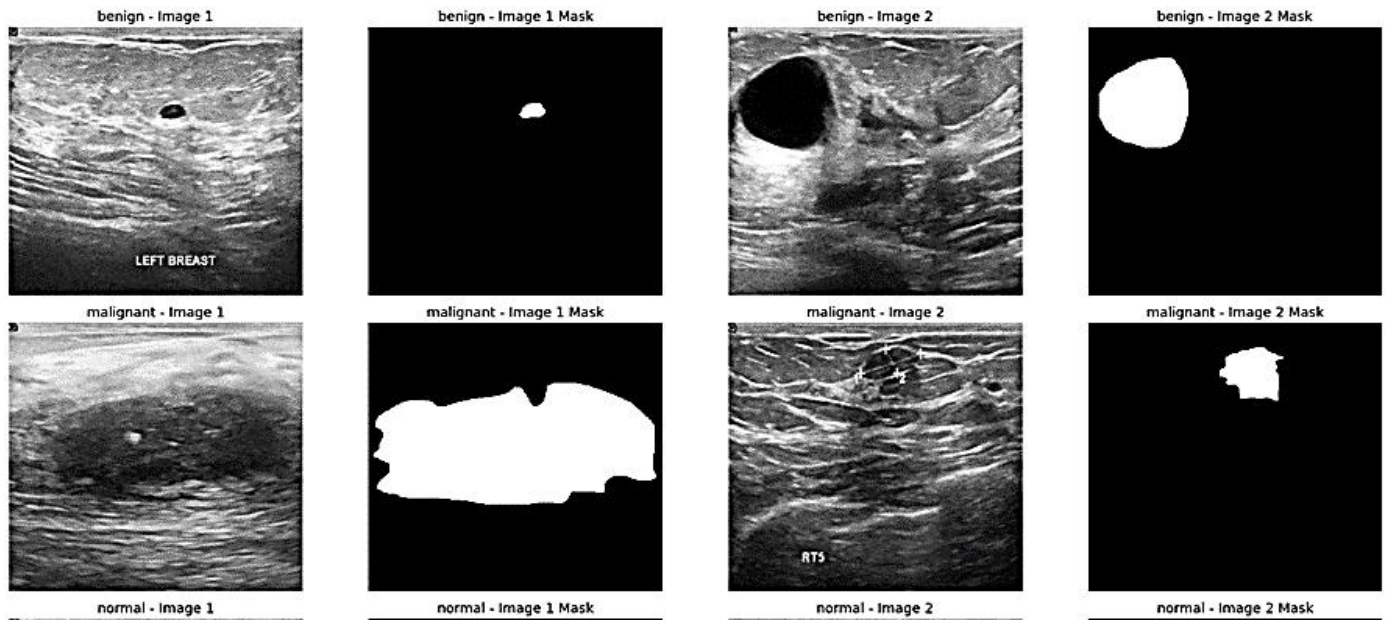


Figure 8. Output of pre-processing stage

5.3 Feature extraction

During the feature extraction stage, two types of features are taken from the input photos to assist the classifiers in determining the class label of the images. The first feature is Gray-Level Co-occurrence Matrix (GLCM) features, while the second is statistical moment's features. Using two types of features, the number of extracted features is huge, thus making the calculation time complex.

5.4 Feature selection

Feature selection is a critical stage in any detection system. In features selection stage, the most relevant features are selected that support the classifier to make the right decision. Selecting the relevant features make the system work fast and accurate by eliminate the irrelevant features. In the proposed system, a new feature selection algorithm based on GWO and CM is proposed.

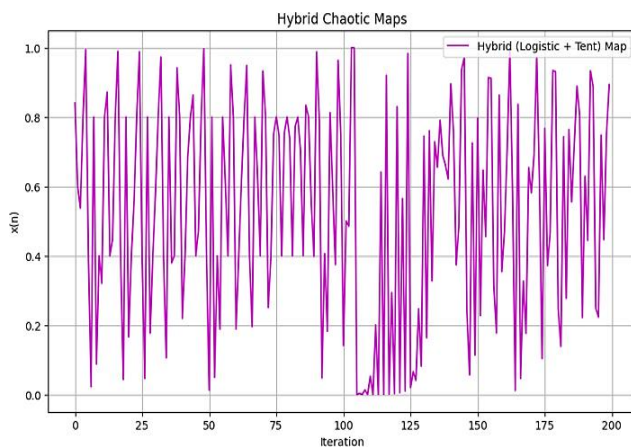


Figure 9. Proposed hybrid chaotic maps (Tent + Logistic)

500 pixels. After that, the noise in input images is removed using 3×3 Median filters in order to enhance the overall system results. A sample images after applying the pre-processing stage are illustrated in Figure 8.

5.4.1 Proposed Grey Wolf Optimizer-Cauchy Mutation

Reducing the number of features make any system work faster. Besides, many irrelevant features could make a false detection, which is a critical issue in the medical field system. As a result, a proposed feature selection algorithm based on GWO and CM named GWO-CM is present.

Population is initialized in GWO-CM algorithm using chaotic maps to provide a strong diversity over the input features. A hybrid approach of initialization is used in the proposed system based on Logistic and Tent chaotic maps as shown in Figure 9. This approach will enhance the population diversity, give better coverage, and avoid periodicity.

The computation formula of the Tent and Logistic maps shown in Eqs. (10) and (11) respectively.

$$x_{i+1} = \begin{cases} \frac{x_i}{\mu}, & x_i < \mu \\ \frac{1-x_i}{1-\mu}, & x_i \geq \mu \end{cases} \quad (10)$$

$$x_{i+1} = rx_i(1 - x_i) \quad (11)$$

After the initialization step, KNN is employed for the fitness function to select the best possible solution. Choosing KNN in the proposed algorithm over the other method due to the simple and fast behavior of the KNN, which makes the algorithm work faster. Eq. (12) shows the proposed fitness function based on KNN.

$$X_{\infty} = \alpha \cdot (1 - F1(KNN)) + \beta \cdot \frac{\text{selected features}}{\text{all features}} \quad (12)$$

where,
 F1: f-score,
 α : 0.9,
 β : 0.1.

The CM is used to improve the original GWO's performance and prevent it from falling into the local optima. CM will generate a massive jump and make the alpha wolves avoid getting stuck in the local optima, thus enhancing the exploration stage.

After applying the Eq. (5) to find the new position of wolves, CM is applied on the alpha wolves (α) the best possible solution in order to avoid the stagnation.

A few iterations are performed then we check the improvement of the wolves. If the alpha wolf is improved then we keep the solution. However, if the alpha wolf does not improve then we apply the CM in Eq. (13) to the alpha wolf to make it do a massive jump and find the best solution.

$$Y_{\alpha, new}(t) = Y_{\alpha, old}(t) + Y_{\alpha, old}(t) \cdot CM(1) \quad (13)$$

where,

$Y_{\alpha, old}(t)$: the t-dimension (features) of alpha,

$CM(1)$: Cauchy distribution number.

By applying this equation, the alpha wolf will teleport to explore new region and lead the pack to this new region. The complete steps of the proposed GWO-CM are shown in Figure 10.

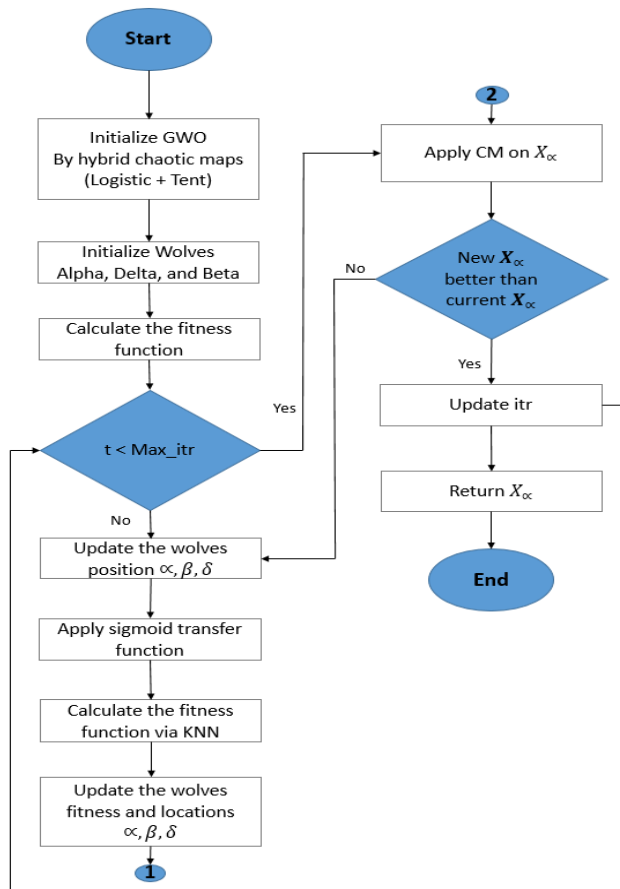


Figure 10. Grey Wolf Optimizer-Cauchy Mutation (GWO-CM) algorithm

CM distribution generate a large number (jump) to guide the alpha wolf to a new unexplored region. However, the Gaussian distribution generate a tight random number that clustered around the mean, thus make the wolf move closely. Therefore, using the CM is the best choice for the proposed algorithm.

The proposed algorithm makes a balance between the exploitation (hunting) than generated from the original GWO

and the exploration (scouting) that generated by CM, which make the algorithm to avoid being stuck with the sub optimal features.

Tent + Logistic used to initialize the population to generate a appropriate diversity and avoid the local optima. After that, KNN is used to evaluate the fitness for each wolf. Besides, CM updates the wolf positions by generating a massive jump to ensure better diversity and avoid the local optimum.

The essential steps of the proposed GWO-CM algorithm are illustrated in Algorithm 1.

Algorithm 1. GWO-CM Algorithm

Input

- BUSI dataset
- Population size = 20 (no. of wolves)
- Max_it = 100
- k neighbor = 9

Output best solution $Y_{\alpha, new}$

Initialize population using Tent + Logistic maps

Determine the wolves α, δ, β

Calculate the fitness function of each wolf using KNN error rate

While $t < Max_it$

For each wolf X_i

Update the wolves' positions

End For

Convert the X_i features into binary using sigmoid

Evaluate fitness function using KNN Eq. (12)

Update the wolves (α, δ, β)

Apply CM on alpha wolf Y_α Eq. (13)

Greedy selection ($Y_{\alpha, new} > Y_{\alpha, old}$) replace

$Y_{\alpha, old}$ with $Y_{\alpha, new}$

$t = t + 1$

End While

Return $Y_{\alpha, new}$

The proposed algorithm addressed the issue of the local optima by improving the population diversity via using the proposed Tent + Logistic approach. In the Tent map the value $\mu = 0.5$ is selected after testing several values 0.3 – 0.7 to ensure the best possible uniform destitution, thus improving the search space convergence. While the r values assign as $r = 4$ in the logistic map to achieve a stronger randomness via the generated chaotic behaviour.

The number of nearest neighbors is determined by the k value. Using small value of k leads to noise and overfitting, while the class discrimination capability is reduced with large number of k value. Therefore, in the proposed system the k value explored in the range [1, 41] in order to get the best solution. The best selected value was 9 which as achieved the minimum classification errors. However, the chosen maximum iteration was 100 in the proposed algorithm, which is improved the algorithm performance via providing a sufficient search capability while maintain the computational cost. In contrast, the scale vector $\gamma = 0.1$ with mutation probability = 0.2 were the better selected values in CM which are led to an appropriate balance between exploration and exploitation.

5.5 Classification

In the classification stage, the system will classify the class label of each input images whether it is benign, malignant, or normal. The input to the classifier will be the selected features

from the previous stage using the proposed GWO-CM and the output will be the classified model in the training phase. However, in the testing phase the input features input directly to the stored model for testing the classified model. The proposed system used four common ML classifiers for results comparison and achieve the best possible results. These classifiers are LR, SVM, KNN, and ANN.

5.6 Evaluation model

The proposed system's performance is evaluated using a variety of measures. In this paper confusion matrix is used by compare the predicted results with the real dataset results. The predicted classes represented in the columns and the real classes presented in the rows. In following the equations of the used metrics.

$$precision = \frac{TP}{TP + FP} \tag{14}$$

$$recall = \frac{TP}{TP + FN} \tag{15}$$

$$F1-score = \frac{2 * recall * precision}{recall + precision} \tag{16}$$

$$accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{17}$$

where,

- TP: the true positive
- TN: the true negative
- FP: the false positive
- FN: the false negative

6. RESULTS AND DISCUSSION

The proposed system was built with the Python programming language and runs in a Windows 11 64-bit environment. The method was tested and evaluated using a standard ultrasound breast cancer dataset known as BUSI. Experimental is conducted by splitting the dataset into 70% for training and 30% for testing by random sampling in order to preserve a satisfied class distribution of the dataset. Six independent runs are performed with fixed number of random seeds (28) to ensure the reproducibility and reducing the bias of the initialization. The proposed system. The system achieved high results across a variety of measurement criteria, as shown in Table 2.

Table 2. Results of Grey Wolf Optimizer-Cauchy Mutation (GWO-CM) algorithm

Dataset	Precision	Recall	F1-Score	Accuracy
BUSI	99%	100%	99%	99.12%

The results of the proposed feature selection algorithm GWO-CM, are compared with the results of the standard GWO. The comparison reveals that the proposed algorithm for GWO-CM outperformed the original GWO in all assessment measures, as shown in Table 3.

Table 3 shows a rise in the obtained results by GWO-CM

compared to GWO. The precision rose to 3%, 1% in recall, 2% F1-score, and 2.63% in accuracy. However, the Receiver Operating Characteristic (ROC) curve of the classification results is illustrated in Figure 11.

Table 3. Results comparison of Grey Wolf Optimizer (GWO) and GWO-Cauchy Mutation (CM) algorithms

Dataset	Alg.	Pre.	Recall	F1-score	Acc.
BUSI	GWO	96%	99%	97%	96.49%
	GWO-CM	99%	100%	99%	99.12%

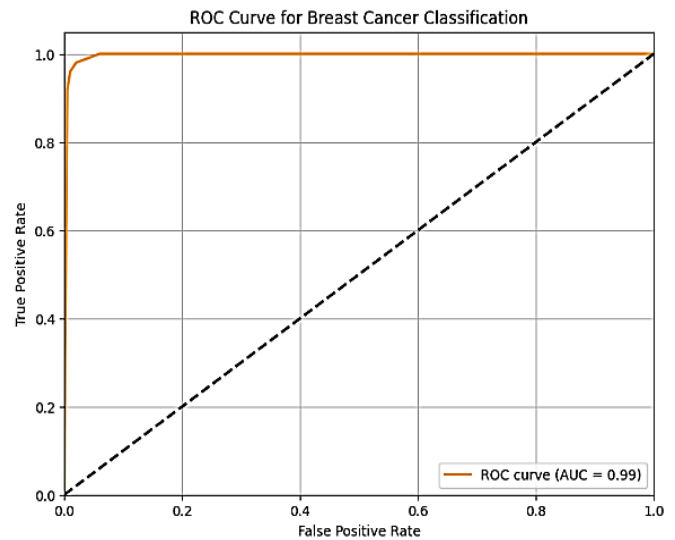


Figure 11. Receiver Operating Characteristic (ROC) curve of the classification

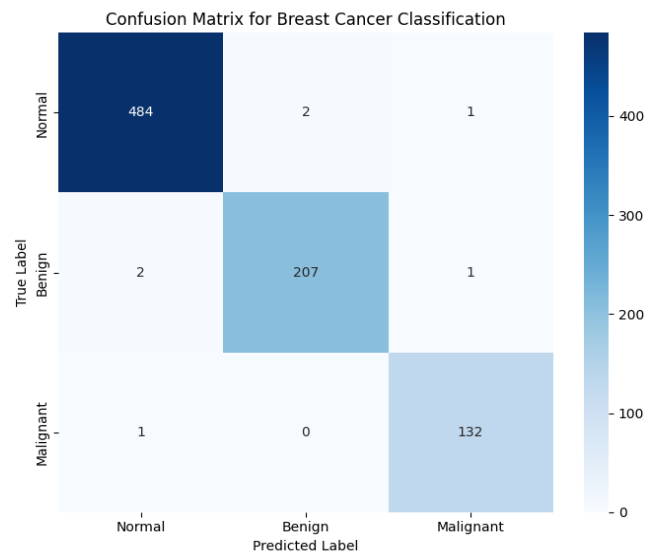


Figure 12. Confusion matrix

The confusion matrix of the obtained results by applying the proposed system is shown in Figure 12.

In the proposed system, several k-fold cross validations are tested to get the appropriate value. After multiple runs 5-fold is selected as shown in Figure 13.

Besides, in the proposed system several machine-learning classifiers are used for classification purpose to get the best classification results. Figure 14 shows the accuracy of using the proposed algorithm GWO-CM with different ML classifiers.

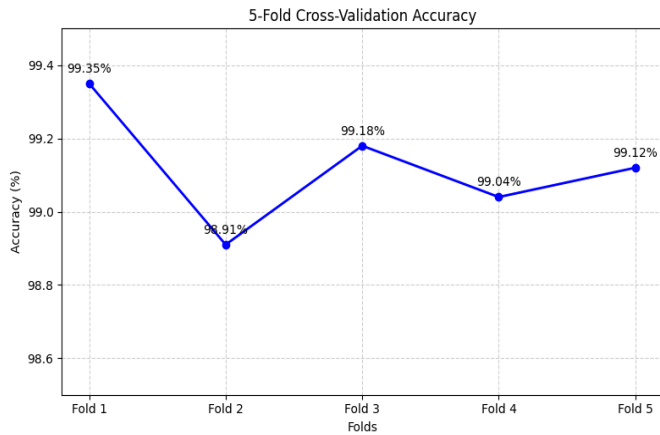


Figure 13. Cross validation

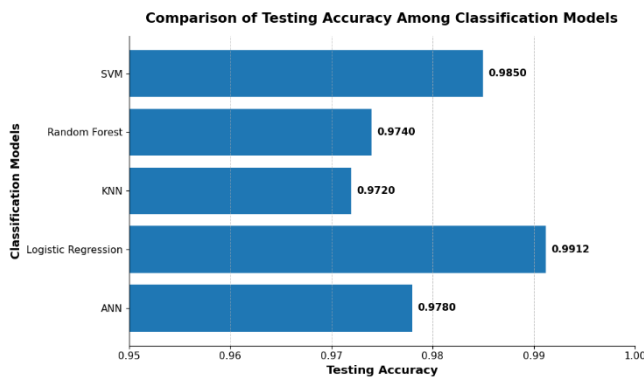


Figure 14. Comparison of model accuracies

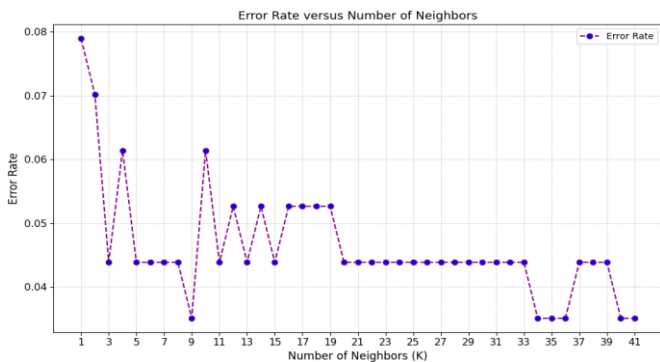


Figure 15. The lowest mean error value by Grey Wolf Optimizer-Cauchy Mutation (GWO-CM)

As observed in Figure 11, the LR achieve the highest classification accuracy than other classifiers. Thus, in our proposed system the LR is considered.

The proposed GWO-CM reduced the number of extracted features by select the best and informative features only. This process leads to make the computing time less and make the system work accurately. The other unselected features may lead the system to fall in the local optima, thus rise the error rate.

By using the chaotic maps (logistic + tent), KNN, with CM in order to enhance the GWO performance in feature selection stage, the process of finding lowest error mean values became more effective and robust. The generated diversity by the hybrid chaotic maps, and the calculated fitness function by KNN, then the long jump by the CM, made the steps of the GWO more stable to find the desired values with lowest error.

Figure 15 shows the process of GWO-CM for finding the lowest mean error values.

In addition, the proposed breast cancer detection system compared with several most existing works. Results of the proposed system are compared with different methods, models with different datasets as shown in Table 4. The comparison showed that the highest accuracy results are obtained by the proposed system compared to other systems' results. The system has a rise in the accuracy results by 2-7% comparing to the other works that used WOBC and WPBC datasets. However, the system outperformed the other system that used WDBC dataset by 0.88%. In another hand, the results obtained by the proposed system increased by 1.28% comparing to other system that use same BUSI dataset.

Table 4. Results comparisons between the proposed Grey Wolf Optimizer-Cauchy Mutation (GWO-CM) algorithm and other recent works

Authors	Dataset and Size	Model	Accuracy
Halde et al. [15]	WOBC + WPBC (569)	FS Based Stacking	92.33%
		Ensemble	89.32%
Alotaibi et al. [16]	BUSI (780)	VGG19	87.8%
Ismael et al. [18]	BUSI (780)	HOG + GWO	95.23%
Mohammed et al. [19]	WDBC (569)	PSO + RF	97.2%
Zafar et al. [22]	BUSI (780)	EO	96.79%
Kumar and Singh [23]	WDBC (569)	EGWO-SVM	98.24%
Ayanaet al. [24]	Mendeley (200)	MSTL	98.7%
Talukder and Ray [28]	WOBC + WPBC (569)	Logistic Regression	98.07%
Singh et al. [29]	WDBC (569)	EPO and GSOA	97.66%
Thawkar et al. [30]	BUSI (780)	Hybrid WOADA	97.84%
Eroglu et al. [31]	BUSI (780)	Alexnet, MobilenetV2, and Resnet50	95.6%
Lu et al. [32]	BUSI (780)	SAFNet	89.6%
Moon et al. [33]	BUSI (780)	CNN	94.62%
Proposed	BUSI (780)	GWO-CM + LR	99.12%

As observed in Table 4, GWO-CM + LR, EGWO-SVM, and MSTL achieved the best classification results. EGWO-SVM enhanced the classification accuracy by improving the SVM classifier tuning. Nevertheless, this approach makes the performance depends heavily on the selecting the SVM kernels and the parameter sensitivity, thus lead for a higher computational complexity. However, the representation capability is improved in MSTL by capturing the multi-level patterns for better classification results. Nonetheless, the model needs further training time due to the increasing in the complexity and the preprocessing requirements. In another hand, GWO-CM + LR improved the convergence and robustness by making a better balance between the exploration and the exploitation strategies. The proposed system,

improves the search diversity and the convergence behavior to make a stable optimization performance by reducing the dependency on the classifier tuning parameters. Furthermore, GWO-CM + LR maintain the competitive predictive capability by employing a simple optimization driven structure comparing to MSTL. A complex hierarchical information can be capture by the MSTL lead for a higher architecture complexity. In contrast, GWO-CM + LR reduced the premature convergence risk and emphasizes efficient optimization performance.

7. CONCLUSIONS

In this paper, an accurate breast cancer detection system is proposed. The system employs several ML methods, and metaheuristic algorithm in order to obtain the best detection results. Five essential stages are included in the proposed system that work together to achieve the best performance. Median filter was used in pre-processing to remove the unwanted noise, and the desired features were extracted by GLCM, statistical moments from the input images. A proposed features selection algorithm based on GWO, chaotic maps, and CM called GWO-CM is presented. The GWO-CM is inspired by the original GWO for selecting the best possible relevant features of the extracted features then employing the Logistic, Tent maps in population initialization and enhancing the exploration phase of the original GWO, by making a large jump to the alpha wolf, thus avoiding the local optima issue. The system was tested and evaluated using standard ultrasound breast images BUSI and achieved better results with an accuracy 99.12%, precision 99%, recall 100%, and f-score 99% with linear regression. Under the experimental conditions, LR achieved results superior to several published methods.

REFERENCES

- [1] Joseph, K., Zebak, S., Alba, V., Mah, K., et al. (2021). Adjuvant breast radiotherapy, endocrine therapy, or both after breast conserving surgery in older women with low-risk breast cancer: Results from a population-based study. *Radiotherapy and Oncology*, 154: 93-100. <https://doi.org/10.1016/j.radonc.2020.09.017>
- [2] Hong, B., Ying, F., Zhaoqing, F., Xichun, H., et al. (2023). Consensus on clinical diagnosis and medical treatment of HER2-low breast cancer (2022 edition). *Journal of the National Cancer Center*, 3(4): 266-272. <https://doi.org/10.1016/j.jncc.2023.09.002>
- [3] Burstein, H.J., Curigliano, G., Thürlimann, B., Weber, W.P., et al. (2021). Customizing local and systemic therapies for women with early breast cancer: The St. Gallen international consensus guidelines for treatment of early breast cancer 2021. *Annals of Oncology*, 32(10): 1216-1235. <https://doi.org/10.1016/j.annonc.2021.06.023>
- [4] Wen, X., Guo, X., Wang, S., Lu, Z., Zhang, Y. (2024). Breast cancer diagnosis: A systematic review. *Biocybernetics and Biomedical Engineering*, 44(1): 119-148. <https://doi.org/10.1016/j.bbe.2024.01.002>
- [5] Fiorica, J.V. (2016). Breast cancer screening, mammography, and other modalities. *Clinical Obstetrics and Gynecology*, 59(4): 688-709. <https://doi.org/10.1097/GRF.0000000000000246>
- [6] Li, Q., Chen, H., Huang, H., Zhao, X., et al. (2017). An enhanced grey wolf optimization based feature selection wrapped kernel extreme learning machine for medical diagnosis. *Computational and Mathematical Methods in Medicine*, 2017(1): 9512741. <https://doi.org/10.1155/2017/9512741>
- [7] Akay, M.F. (2009). Support vector machines combined with feature selection for breast cancer diagnosis. *Expert Systems with Applications*, 36(2): 3240-3247. <https://doi.org/10.1016/j.eswa.2008.01.009>
- [8] Okwu, M.O., Tartibu, L.K. (2020). Grey wolf optimizer. In *Metaheuristic Optimization: Nature-Inspired Algorithms Swarm and Computational Intelligence, Theory and Applications*, pp. 43-52. https://doi.org/10.1007/978-3-030-61111-8_5
- [9] Long, W., Liang, X., Cai, S., Jiao, J., Zhang, W. (2017). A modified augmented Lagrangian with improved grey wolf optimization to constrained optimization problems. *Neural Computing and Applications*, 28(Suppl 1): 421-438. <https://doi.org/10.1007/s00521-016-2357-x>
- [10] Saremi, S., Mirjalili, S.Z., Mirjalili, S.M. (2015). Evolutionary population dynamics and grey wolf optimizer. *Neural Computing and Applications*, 26(5): 1257-1263. <https://doi.org/10.1007/s00521-014-1806-7>
- [11] Zhang, Y., Jin, Z., Chen, Y. (2020). Hybridizing grey wolf optimization with neural network algorithm for global numerical optimization problems. *Neural Computing and Applications*, 32(14): 10451-10470. <https://doi.org/10.1007/s00521-019-04580-4>
- [12] Emary, E., Zawbaa, H.M., Hassanien, A.E. (2016). Binary grey wolf optimization approaches for feature selection. *Neurocomputing*, 172: 371-381. <https://doi.org/10.1016/j.neucom.2015.06.083>
- [13] Sahoo, A., Chandra, S. (2017). Multi-objective grey wolf optimizer for improved cervix lesion classification. *Applied Soft Computing*, 52: 64-80. <https://doi.org/10.1016/j.asoc.2016.12.022>
- [14] Kumar, N., Kumar, D. (2021). An improved grey wolf optimization-based learning of artificial neural network for medical data classification. *Journal of Information and Communication Technology*, 20(2): 213-248. <https://e-journal.uum.edu.my/index.php/jict/article/view/13151>
- [15] Halder, R.K., Lima, M.A., Uddin, M.N., Islam, M.A., Saha, A. (2025). Integrated feature selection-based stacking ensemble model using optimized hyperparameters to predict breast cancer with smart web application. *Clinical eHealth*, 8: 146-161. <https://doi.org/10.1016/j.ceh.2025.08.001>
- [16] Alotaibi, M., Aljouie, A., Alluhaidan, N., Qureshi, W., et al. (2023). Breast cancer classification based on convolutional neural network and image fusion approaches using ultrasound images. *Heliyon*, 9(11): e22406. <https://doi.org/10.1016/j.heliyon.2023.e22406>
- [17] Benghazouani, S., Nouh, S., Zakrani, A. (2025). Optimizing breast cancer diagnosis: Harnessing the power of nature-inspired metaheuristics for feature selection with soft voting classifiers. *International Journal of Cognitive Computing in Engineering*, 6: 1-20. <https://doi.org/10.1016/j.ijcce.2024.09.005>
- [18] Ismael, O.F., Imran, M., Al-Khuzai, M., Kherallah, M., Kammoun, F. (2025). Deep WolfNet model: Enhancing medical disease diagnosis using gray wolf technology

- and deep neural networks. *International Journal of Online & Biomedical Engineering*, 21(4): 61-78. <https://doi.org/10.3991/ijoe.v21i04.53811>
- [19] Mohammed, S.A., Abeysinghe, S.D., Ralescu, A.L. (2023). Feature selection and comparative analysis of breast cancer prediction using clinical data and histopathological whole slide images. *Advances in Artificial Intelligence and Machine Learning*, 3(3): 1494-1525. <https://doi.org/10.54364/aaiml.2023.1187>
- [20] Yavuz, G., Moghanjoughi, M.K., Dumlu, H., Çakir, H.İ. (2025). A feature selection method combining filter and wrapper approaches for medical dataset classification. *Vietnam Journal of Computer Science*, 12(4): 375-393. <https://doi.org/10.1142/S2196888824500246>
- [21] Mustafa, M.A., Erdem, O.A., Söğüt, E. (2025). Hybrid optimization and explainable deep learning for breast cancer detection. *Applied Sciences*, 15(15): 8448. <https://doi.org/10.3390/app15158448>
- [22] Zafar, A., Tanveer, J., Ali, M.U., Lee, S.W. (2023). BU-DLNet: Breast ultrasonography-based cancer detection using deep-learning network selection and feature optimization. *Bioengineering*, 10(7): 825. <https://doi.org/10.3390/bioengineering10070825>
- [23] Kumar, S., Singh, M. (2021). Breast cancer detection based on feature selection using enhanced grey wolf optimizer and support vector machine algorithms. *Vietnam Journal of Computer Science*, 8(2): 177-197. <https://doi.org/10.1142/S219688882150007X>
- [24] Ayana, G., Park, J., Jeong, J.W., Choe, S.W. (2022). A novel multistage transfer learning for ultrasound breast cancer image classification. *Diagnostics*, 12(1): 135. <https://doi.org/10.3390/diagnostics12010135>
- [25] Mirjalili, S., Mirjalili, S.M., Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software*, 69: 46-61. <https://doi.org/10.1016/j.advengsoft.2013.12.007>
- [26] Ali, M., Pant, M. (2011). Improving the performance of differential evolution algorithm using Cauchy mutation. *Soft Computing*, 15(5): 991-1007. <https://doi.org/10.1007/s00500-010-0655-2>
- [27] Al-Dhabyani, W., Gomaa, M., Khaled, H., Fahmy, A. (2020). Dataset of breast ultrasound images. *Data in Brief*, 28: 104863. <https://doi.org/10.1016/j.dib.2019.104863>
- [28] Talukder, P., Ray, R. (2024). Analysis of breast cancer classification using machine learning techniques and hyper parameter tuning. *Biocatalysis and Agricultural Biotechnology*, 58: 103195. <https://doi.org/10.1016/j.bcab.2024.103195>
- [29] Singh, L.K., Khanna, M., Singh, R. (2024). Efficient feature selection for breast cancer classification using soft computing approach: A novel clinical decision support system. *Multimedia Tools and Applications*, 83(14): 43223-43276. <https://doi.org/10.1007/s11042-023-17044-8>
- [30] Thawkar, S., Katta, V., Parashar, A.R., Singh, L.K., Khanna, M. (2023). Breast cancer: A hybrid method for feature selection and classification in digital mammography. *International Journal of Imaging Systems and Technology*, 33(5): 1696-1712. <https://doi.org/10.1002/ima.22889>
- [31] Eroğlu, Y., Yildirim, M., Çinar, A. (2021). Convolutional neural networks based classification of breast ultrasonography images by hybrid method with respect to benign, malignant, and normal using mRMR. *Computers in Biology and Medicine*, 133: 104407. <https://doi.org/10.1016/j.compbiomed.2021.104407>
- [32] Lu, S.Y., Wang, S.H., Zhang, Y.D. (2022). SAFNet: A deep spatial attention network with classifier fusion for breast cancer detection. *Computers in Biology and Medicine*, 148: 105812. <https://doi.org/10.1016/j.compbiomed.2022.105812>
- [33] Moon, W.K., Lee, Y.W., Ke, H.H., Lee, S.H., Huang, C.S., Chang, R.F. (2020). Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks. *Computer Methods and Programs in Biomedicine*, 190: 105361. <https://doi.org/10.1016/j.cmpb.2020.105361>