



International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Hybrid Marine Predators and Parrot Optimization Approach For Mammographic Mass Classification

Navreet Kaur¹, Rahul Hans², Harpreet Singh³

^{1,2}Department of Computer Science and Engineering, DAV University, Jalandhar, Punjab, India.

¹Department of Computer Science and Engineering, Chandigarh University, Gharuan Mohali-140413, Punjab, India.

³Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, India.

¹reetnav1995@gmail.com, ²rahulhans@gmail.com, ³harpreet99.nitj@gmail.com

Abstract

Breast cancer is regarded as among the most serious health issues impacting women globally. Mammography is the benchmark standard for diagnosing this deadly disease, but the inherent subjectivity of human interpretation may introduce biases that adversely affect outcomes. In recent years, deep learning-based medical image analysis techniques have emerged to assist clinicians in precise disease diagnosis. In this research, a Chaotic Binary Marine Predators Algorithm integrated with Parrot Optimization (MPPA) model, which uses chaotic maps in order to provide better exploitation and exploration balance throughout the search process, is proposed. To conduct the research, regions of interest (ROIs) are extracted from the CBIS-DDSM mammogram image dataset and are pre-processed using various steps. Further, the deep features are extracted using the AlexNet and ResNet model and optimised using the proposed Hybrid versions. The proposed algorithm aims to enhance classification accuracy while reducing feature dimensionality. In this regard, an extreme learning machine (ELM)-based fitness function is considered. Further, the performance of the proposed algorithm is also validated on standard datasets from the UCI repository. Experimental results signify that the proposed algorithm outperforms the other algorithms considered for performance comparison for the efficient classification of breast mass in mammogram images.

Keywords: Breast Mass Classification; Nature-Inspired Metaheuristic Algorithms; Marine Predators Optimization Algorithm; Parrot Optimization; Feature Selection.

1 Introduction

Breast cancer is one of the biggest problems in global healthcare and the most common type of cancer diagnosed in women. As reported by WHO and GLOBOCAN, about 2.3 million cases occurred in 2022, representing 11.6% of all cancers [1,2]. The disease happens because of the unlimited growth of breast cells and forms benign and malignant tumors. Timely diagnosis is very important for the effective treatment of breast cancer. Mammography is a gold-standard tool for its diagnosis; however, detecting tumors can be difficult in dense breasts [3–5].

Recent developments in deep learning allowed automatic analysis of medical images to diagnose diseases more accurately [6–8]. Convolutional neural networks (CNNs) are often used for breast cancer detection and classification; however, the training process requires vast amounts of data and a high level of computation [9]. Using transfer learning helps in overcoming these issues by taking advantage of pre-trained models, thus reducing training time and improving performance, especially on small datasets [10–12]. Even though deep learning algorithms automatically extract discriminative features from medical images, a number of them are unnecessary. Selecting only informative features is usually solved as a multi-objective problem [13,14].

Metaheuristics inspired from nature have exhibited great efficiency in tackling difficult optimization problems based on two essential procedures, namely exploration and exploitation [15]. While the former procedure searches for solutions in the global search space, the latter procedure improves selected solutions in the local space. However, striking a good balance between both procedures is essential to prevent premature convergence and obtain optimal results efficiently [16,17]. With such benefits in mind, the BMPPA-PO method will be proposed for feature selection to classify breast mass from mammograms using the ELM-based fitness function.

1.1 Research contributions

Some of the important research contributions of this article.

- A robust pre-processing pipeline is proposed to enhance mammographic image quality.
- Deep features are extracted from AlexNet and Resnet.
- An enhanced version of Chaotic Binary Marine Predators Algorithm integrated with Parrot Optimization is proposed for selecting minimum number of features while maximizing classification accuracy.
- ELM-based fitness evaluation for high classification accuracy and reduced computational time.
- Performance evaluation of proposed algorithm on various standard datasets of diverse domains.

1.2 Structuring of paper

The rest of the article is structured as: Section 2 presents the review of existing work done in the domain of disease diagnosis using deep learning and nature-inspired metaheuristic algorithms. Section 3 discusses the proposed architecture in detail. The results and discussion are presented in section 4. Lastly, section 5 briefly presents the conclusions and future directions.

2 Literature Review

It can be acknowledged from literature, that a significant work has been accomplished in the domain of feature selection by utilizing nature-inspired metaheuristic algorithms and deep transfer learning algorithms in recent times, some of the recent work done is summarized below:

For instance, in [18], an optimized hybrid LSTM-GRU neural network model that used BPSO for feature selection and Grey Wolf-Whale Optimization Algorithm to tune the hyperparameters attained 100% and 99.33% accuracy for lung cancer diagnosis using BPSO-GWO approach. Moreover, ODeep-NN architecture in [19] combined transfer learning, feature fusion, and Biogeography-Based Optimization for diabetic retinopathy diagnosis and acquired 98.04% average accuracy. A combination of Liquid Neural Network algorithm with Ant Colony Optimization was used for breast tumor detection, which had an accuracy rate of 93.75% on the BreakHis dataset [20]. An optimization algorithm for mammographic breast density classification through a Biogeography-Based Optimization and Genetic Algorithm was purposed [21]. In [22], CNN models and Grey Wolf Optimization were combined with transfer learning algorithms to detect breast cancer with a high accuracy rate of 96.5%. Also, a CAD model using wavelets for image processing with ResNet-50, Inception, and Xception architecture achieved an accuracy, sensitivity, specificity, and AUC score of 99% on the INbreast and MIAS dataset [23].

This paper [24] proposed a breast cancer classification framework using stacked machine learning models optimized with a hybrid PSO-GWO algorithm. Feature selection was performed using BPSO, BGWO, and BWO, with BPSO achieving the best results. A stacked model with XGBoost as the meta-classifier reached 93.91% accuracy, which further improved to 99.49% after PSO-GWO optimization. Another study [25] introduced a pancreatic cancer diagnosis method that converts CT images into graphs using Harris Corner Detection and applies transfer learning with DenseNet121 and InceptionV3. The Whale Optimization Algorithm (WOA) was used for parameter tuning, and the k-NN classifier achieved the best performance with 92.10% accuracy and a 92.74% F1 score. In [26], the Q-BGWO-SQSVM classifier was proposed that uses Quantum-Inspired Binary Grey Wolf Optimizer along with the SqueezeNet CNN architecture and SVM for breast cancer detection with an impressive accuracy of 99%, sensitivity of 98%, and specificity of 100% on the CBIS-DDSM database. In [27] and [28], deep learning and meta-heuristics hybrid models using ResNet50 and CS/GWO classifiers were introduced that have shown outstanding performance in terms of up to 99.68% accuracy in detecting COVID-19. In [29], hybrid deep-transfer models such as Xception, especially XRF and MXC, performed excellently on BreakHis and BHID databases by achieving 94.44% and 88.50% accuracy, respectively.

3. Proposed Architecture

This research aims to propose a pipeline architecture for breast mass classification using various phases. The proposed architecture is shown in Fig. 1 considering image preprocessing phase, transfer learning-based feature extraction phase, feature subset generation phase and ELM based classification phase.

To conduct the experiments this research considers the CBIS-DDSM [30] dataset, which is a curated and standardized subset of the original Digital Database for Screening Mammography (DDSM). The DDSM consists of 2,620 scanned film mammography studies containing normal, benign, and malignant cases. In CBIS-DDSM, radiologists provide precisely annotated region-of-interest patches for breast masses, all originally in single-channel grayscale images with different resolutions and various contrast levels. The detailed explanation of various phases of the proposed architecture has been given below.

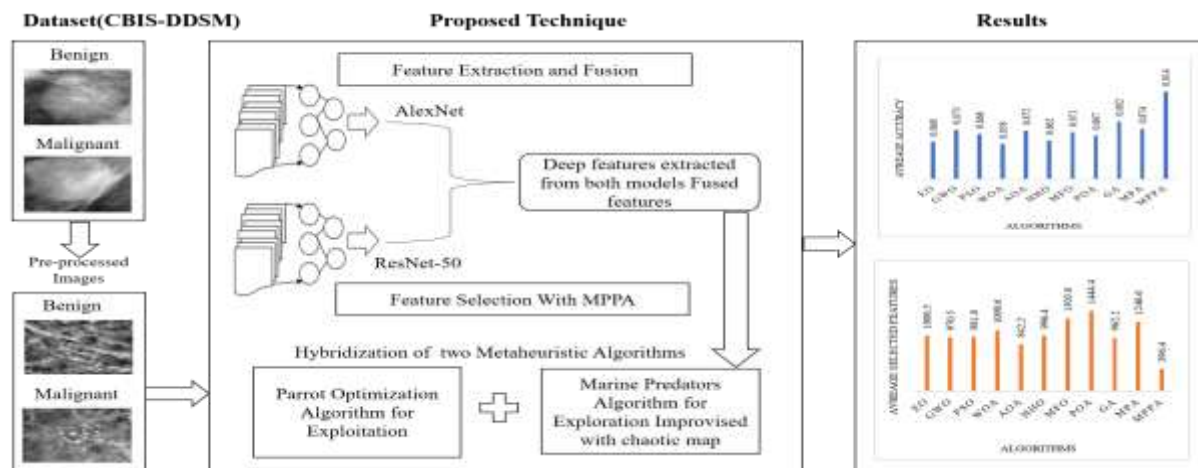


Figure 1. Proposed Workflow

3.1 Preprocessing

In order to preprocess CBIS-DDSM mammographic ROI images to be fed into the deep learning model, an image preprocessing pipeline was used to increase image quality and preserve relevant diagnostic information within each image. Due to the presence of noise in the raw grayscale image along with inconsistencies in the contrast and sharpness, the images underwent preprocessing to help visualize the lesions and to enable proper feature extraction through AlexNet. The preprocessing pipeline was made up of seven major stages, which included resizing the image, removing the noise, increasing the contrast, normalizing intensities, replicating channels, sharpening edges, and morphological filtering [31].



Figure 2. Mammographic ROI Benign and Malignant Images before preprocessing.

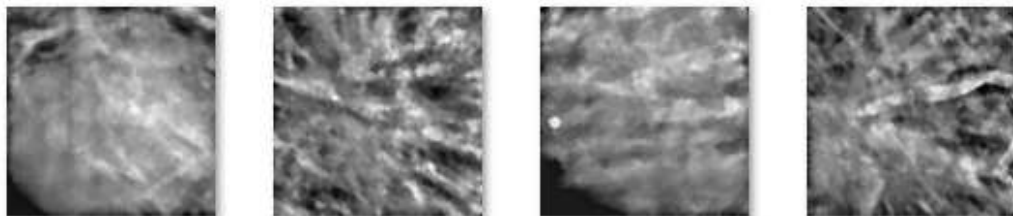


Figure 3. Mammographic ROI Benign and Malignant Images After preprocessing

3.2 Feature Extraction using AlexNet and ResNet

Both pretrained deep learning models, namely, AlexNet and ResNet, are fed the pre-processed CBIS-DDSM mammographic ROIs independently for extracting their features. Both pretrained models automatically learn discriminatory attributes from the ROIs, making it possible to distinguish between the malignant and benign cases based on these attributes. The images are resized based on the size specifications needed by each model and sent through the layers of each model independently. The deep features are thus extracted from the last layer of both AlexNet and ResNet separately, which correspond to higher-level attributes learned by each model. Finally, after the feature extraction process, the deep features extracted by AlexNet and ResNet are merged together into one single feature dataset. In this way, the proposed deep feature-based approach will ensure robustness and lower dependence on handcrafted descriptors, together with increasing the discriminative power of the proposed classification framework.

3.3 Feature Selection and classification

This research considers one of the latest nature-inspired metaheuristic algorithms, the Marine Predators Algorithm, which is a bio-inspired optimization technique. To enhance this algorithm, the study incorporates two concepts: Chaotic maps and Extreme Learning Machine (ELM) Machine [32].

3.3.1. Marine Predators Algorithm (MPA)

This algorithm mimics the foraging behaviour of marine predators like sharks and tunas while hunting prey. It excels in solving continuous optimisation problems and is employed to solve various optimization

problems. Each agent in the population represents a potential solution, categorized as Top, Mid-level, or Lower-level predators. The MPA [190] utilizes Lévy flights and Brownian motion, mirroring marine animals' random and directional movements. Brownian motion facilitates local searches, while Lévy flight supports global searches. Three stages inspired by predator-prey interactions are defined as.

a) Initialization

Let the population consist of search agents, where each agent represents a potential solution in a d -dimensional search space. The initial prey population can be represented as:

$$prey = [prey_1, prey_2, \dots, prey_N]$$

where the position of the prey is represented by:

$$prey = [x_{i1}, x_{i2}, \dots, x_{id}]$$

The initial positions are generated within the predefined lower and upper bounds of the search space. At each iteration, the fitness of every solution is calculated. The best solution obtained is stored in the Elite matrix, which guides the movement of the remaining search agents. The elite solution is updated whenever a better solution is found

1. High-Speed Phase (Exploration): Prey outpaces predators, leading to Brownian motion for a broader search area and avoiding local optima.

The step size is calculated as:

$$step_i = R_B \otimes (Elite_i - R_B \otimes prey_i) \quad (1)$$

The prey position is subsequently updated as:

$$prey_i^{t+1} = prey_i^t + P \cdot R \otimes step_i \quad (2)$$

Where R_B is a random vector generated according to Brownian motion, R is a uniformly distributed random vector, $P=0.5$ is a constant, denotes element-wise multiplication and $Elite_i$ represents the corresponding elite solution. t represents the current iteration.

2. Unit-Speed Phase (Exploitation + Exploration): In this phase, the population is divided into two groups.

Both predator and prey move at the same speed, combining Brownian motion for elites and Lévy flight for others to focus on promising regions.

$$step_i = R_L \otimes (Elite_i - R_L \otimes prey_i) \quad (3)$$

the position is updated as:

$$prey_i^{t+1} = prey_i^t + P \cdot R \otimes step_i \quad (4)$$

where R_L represents a random vector generated according to Lévy flight.

For the second half of the population, Brownian motion is applied:

$$step_i = R_B \otimes (R_B \otimes Elite_i - prey_i) \quad (5)$$

$$prey_i^{t+1} = Elite_i + P \cdot CF \otimes step_i \quad (6)$$

Here CF is an adaptive convergence factor to control the magnitude of movement during optimization. The decreasing value of CF gradually reduces the search step and supports the transition from exploration to exploitation. This phase provides a transition between broad exploration and focused exploitation.

3. Low-Speed Phase (Exploitation): Predators are faster than prey, emphasising refinement and solution exploitation through Lévy flight for elite fine-tuning.

$$step_i = R_L \otimes (R_L \otimes Elite_i - prey_i) \quad (7)$$

and the prey position is updated using:

$$prey_i^{t+1} = Elite_i + P \cdot CF \otimes step_i \quad (8)$$

This mechanism enables the algorithm to progressively concentrate the search around high-quality solutions during the later iterations.

b) Fish Aggregating Devices (FADs) Effect

To maintain population diversity and reduce premature convergence, MPA incorporates the effect of Fish Aggregating Devices (FADs). The FAD probability is generally set to 0.2. The FAD mechanism introduces additional perturbations or moves solutions toward different regions of the search space.

c) Elite and Memory Mechanism

MPA also incorporates a memory mechanism through the Elite matrix. The best solutions obtained during previous iterations are retained and compared with newly generated solutions. If a newly generated solution provides better fitness, it replaces the corresponding stored solution. This enables the algorithm to preserve historically good solutions throughout the optimization process.

3.3.2. Parrot Optimization Algorithm

The Parrot Optimization (POA) algorithm [34] is an evolutionary computational approach that uses the behaviors of *Pyrrhura* parrots. This method uses four different behaviors that include foraging, settling down, communication, and fear of strangers. These behaviors help the algorithm achieve a balance between exploration and exploitation. Firstly, a set of candidate solutions or particles is initialized in the search area. Each of the particles acts as a solution to the optimization problem. Their fitness is calculated, and the optimal solution is determined as best [36]

1. Foraging Behavior (Exploring): The particles evaluate the position of the food source using the best solution or the position of the owner and head towards it. The Lévy flight pattern is also used to improve global exploration and prevent premature convergence.

$$X_i^{t+1} = (X_i^t - X_{best}). Levy(dim) + rand(0,1). (1 - t/max_iter)^{2t/max_iter}. X_{mean}^t \quad 9$$

2. Staying Behavior (Local Exploiting): In order to simulate a stay behavior pattern, the particles randomly select a position close to the owner and stay at that point. This step improves local exploiting skills of the particles.

$$X_i^{t+1} = (X_i^t + X_{best}). Levy(dim) + rand(0,1). ones(1, dim) \quad 10$$

3. Communicating Behavior (Exchange of Information): Information exchange occurs among particles in the flock. In other words, each particle chooses either to approach or move away from the population mean randomly.

$$X_i^{t+1} = \begin{cases} 0.2.rand(0,1). (1 - t/max_iter). X_i^t - X_{mean}^t, & p \leq 0.5 \\ 0.2.rand(0,1). exp(-(t/rand(0,1). max_iter)), & p > 0.5 \end{cases} \quad 11$$

4. Fear of Strangers Behavior (Prevention of Diversity Loss): Particles try not to enter strange areas and rather move towards safer areas (towards the best solution or owner). In addition to that, they try moving away from bad solutions.

$$X_i^{t+1} = X_i^t + rand(0,1).cos(0.5\pi.t/max_iter). (X_{best} - X_i^t). \\ -cos(rand(0,1). \pi. (t/max_iter)^{2/max_iter}). (X_i^t - X_{best}). \quad 12$$

Algorithm 1: Marine Predators Algorithm (MPA)

```

1. Initialize parameters and population (solutions)
2. Evaluate fitness of each predator (solution)
3. Identify Elite (best solution) and construct Elite matrix
4. Set iteration t = 1
5. While (t ≤ Max_iter)
6. Calculate CF
   CF = (1 - t/Max_iter)^(2t/Max_iter)
   For each predator i=1 to N: do
     a. If t ≤ Max_iter/3
       Calculate step size using Eq. (1)
       Update prey position using Eq. (2)
     b. Else if Max_iter/3 < t ≤ 2*Max_iter/3
       If elite group Eq (3)
         Update using Eq. (4)
       Else
         Calculate using Eq (5)
         Update position using Eq (6)
     c. Else
       Calculate step size using Eq. (7)
       Update Xi using Eq. (8)
     d. Apply boundary constraints
   End For
   e. Evaluate fitness of updated prey solutions
     Update Elite
     Apply FDAs
     Apply boundary constraints
     Update Elite
     t = t + 1
   End While
6. Return the best solution (Elite)

```

Algorithm 2: Parrot Optimizer Algorithm (POA)

```

1. Initialize parameters and population (solutions)
2. Evaluate fitness of each predator (solution)
3. Identify best Parrot (global best solution)
4. Set iteration t = 1
5. While (t ≤ Max_iter)
   For each parrot Xi:
     a. Generate a random number p ∈ [0,1]
     b. If p < 0.25
       Update Xi using Eq. (9)
     c. Else if 0.25 ≤ p < 0.50
       Update Xi using Eq. (10)
     d. Else if 0.50 ≤ p < 0.75
       Update Xi using Eq. (11)
     e. Else
       Update Xi using Eq. (12)
     f. Apply boundary constraints
     g. Evaluate fitness of updated Xi
     h. If Xi is better than previous
       Update Xi
   End For
   i. Update global best Xbest
   j. t = t + 1
   End While
6. Return the best (global best parrot)

```

3.3.3. Proposed Binary Marine Predators Algorithm integrated with Parrot Optimization Algorithm.

This section showcases the improvised version of marine predator algorithm by using chaotic maps and Extreme Learning Machine is used as fitness function. The proposed algorithm considers $R = \text{chaos}(\text{Iter})$; instead of the standard $R = \text{rand}()$; and introduces chaotic behaviour into the algorithm's random decision-making process. Chaotic maps [35] are mathematical functions known for their chaotic behavior, which implies they are deterministic but extremely sensitive to initial conditions.

In this study in the Binary Marine Predators Algorithm, using $R = \text{chaos}(\text{Iter})$; instead of the standard $R = \text{rand}()$; introduces chaotic behavior into the algorithm's random decision-making process. The Tent chaotic maps used in this algorithm. Chaotic maps generate deterministic yet non-repetitive sequences within [0,1], offering more diverse and unpredictable values than conventional random functions. In the

proposed model, such chaotic value is typically used for probabilistic decisions, such as updating binary positions, thereby boosting performance in complex binary optimization tasks like feature selection. Furthermore, to enhance the exploitation capability of the proposed framework, the Binary Marine Predators Algorithm is integrated with the Parrot Optimizer (POA), resulting in a hybrid BMPA-POA model. Even though BMPA with the inclusion of chaotic maps can effectively cover the entire search space via Brownian motion, Lévy flight, and phase transition, there could be some constraints during the optimization process to fine-tune the solution towards the end. In order to overcome such problems, POA is introduced as a local search method. In the proposed hybrid approach, after the position update of search agents using BMPA dynamics, each solution is further refined using the POA-based update rule defined as:

$$x_i^{new} = x_i + F \cdot (x^* - x_r) \quad 13$$

where x_i represents the current solution, x^* denotes the global best solution obtained from MPA, x_r is a randomly selected solution from the population, and F is the flight factor controlling the step size. Here Flight factor reduced over iterations as:

$$F_{t+1} = \alpha \cdot F_t \quad \alpha \in (0,1)$$

This gives the algorithm a better ability to explore the solution space widely at an initial stage, with subsequent steps being dominated by exploration of already known good solutions in the later iterations. Combining POA after MPA will lead to refinement of globally searched solutions, thus achieving more efficient convergence and higher stability. Moreover, because our model works within a binary search space for selecting features, two different kinds of transfer functions have been used for converting continuous variables to binary values. Firstly, a sinusoidal S-shaped transfer function is employed to create binary solutions stochastically, while secondly, a V-shaped transfer function is used to adaptively flip bits of the generated solutions.

Algorithm 3: Hybrid Chaotic MPPA for Optimal Feature Selection

Input:

Population size N, Maximum iterations T, Dimension d, Fitness function $f(x)$

Output:

Best solution x^* , Best fitness $f(x^*)$

1. Initialize population
2. Evaluate fitness of each solution
3. Identify global best solution
4. Initialize parameters α
- For $t = 1$ to T
 5. Update adaptive parameters
 6. Generate chaotic value
 - (A) CBMPA Exploration Phase
 7. For each solution
 - a. Compute Brownian motion and Lévy flight
 - b. Phase 1 to Phase 3 equation (1-8)
 - (B) POA Exploitation Phase
 8. For each solution
 - a. Select random solution
 - b. Update as equation (9-12).
 - c. Apply boundary control
 - d. Evaluate fitness
 - e. Better found update
 9. Update flight factor
 10. Binary Conversion
 11. Memory Update
 12. End For
 13. Return best solution x^*

3.3.4 Extreme Learning Machine (ELM)

Extreme Learning Machine (ELM) [32] is a fast-learning algorithm for Single-Layer Feedforward Neural Networks (SLFNs). This algorithm is widely utilized for solving classification and regression problems and comprises three layers, namely, the input layer, the hidden layer having weights and bias assigned randomly, and the output layer. In the present paper, ELM is applied as the classifier in the feature selection procedure. For each of the feature sets generated using EFOA, the classification results are evaluated by employing ELM. Given that ELM is characterized by high training speed and low computational expense as well as its ability to generalize successfully, this algorithm may be considered appropriate for processing high-dimensional images. Thus, the aim is to increase classification accuracy while selecting a minimal set of features. The equation of the proposed fitness function is given in Equation (13).

$$Fitness = CP_1 \cdot (1 - Acc) + CP_2 \cdot \frac{|S|}{m} \quad 13$$

Here Acc is the accuracy obtained from classifier, |S| is the number of selected features and m is the total number of features available. CP_1 and CP_2 are two control parameters: $CP_1 = 1 - \alpha$ and $CP_2 = \alpha$. Here, α is a weighting factor (set to 0.01 in our experiment), it adjusts the balance between classification error and subset size. A lower fitness value indicates a better trade-off between accuracy and feature reduction. The classifiers are responsible for creating prediction labels on the validation set, further from these predictions, multiple performances metrics can be computed. Here $(1 - Acc)$ is the Error rate which used in the fitness function

$$Acc = \frac{\text{Correct predictions}}{\text{Total predcations}}$$

Table 2. Evaluation Metrics

S. No.	Metric	Explanation
1	Average Classification Accuracy	The overall average of classification accuracy obtained over multiple independent runs.
2	Average Fitness	Represents the mean fitness score, reflecting both solution quality and feature subset size.
3	Standard Deviation	Measures the consistency or variation in results across multiple runs.
4	Average Number of Selected Features	Indicates how many features are typically selected by the model on average.
5	Average Time	The average time required to complete one full execution of the algorithm.

4. Results and Discussions

The proposed algorithm outperforms the other nature-inspired metaheuristic algorithms on the basis of various parameters as shown in table 2. The figure 4 to 9 represents the results of on CBIS-DDSM dataset.

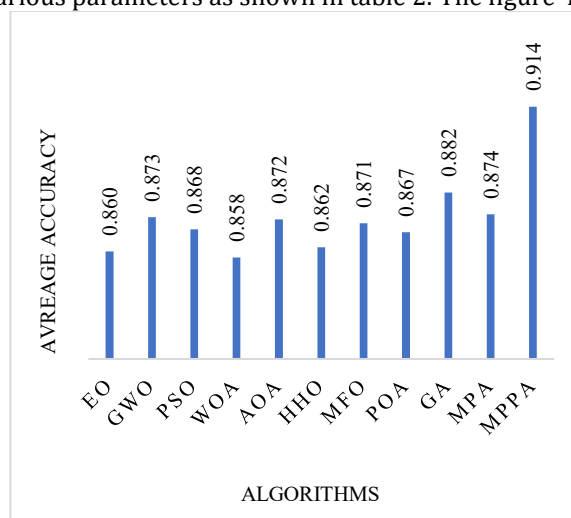


Figure 4. Average Classification Accuracy

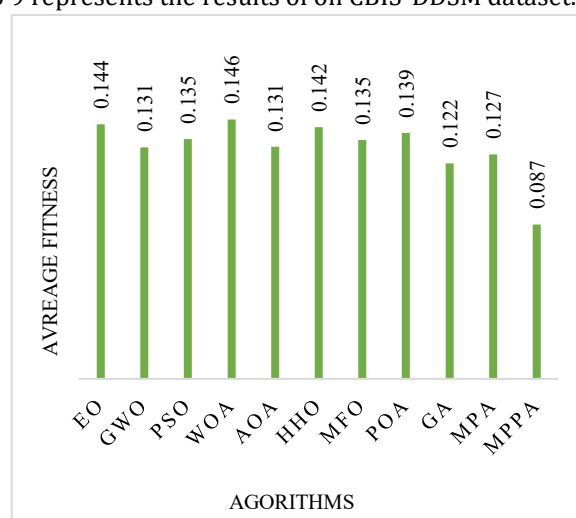


Figure 5. Average Fitness

Figure 4 clearly shows that proposed algorithm provides the better accuracy than other algorithms considered for comparative analysis. Further fitness values also present the stability of the proposed algorithm.

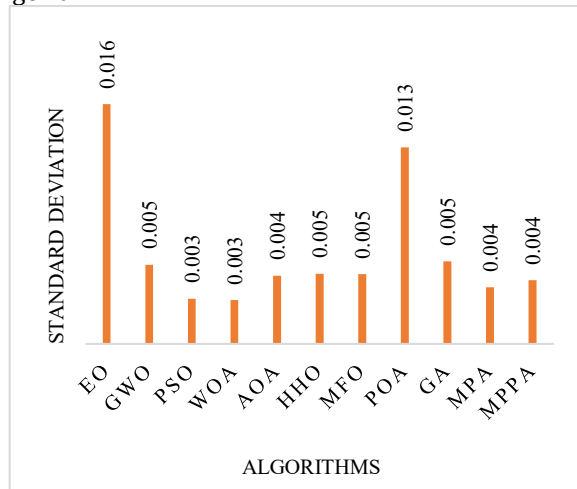


Figure 6. Standard Deviation

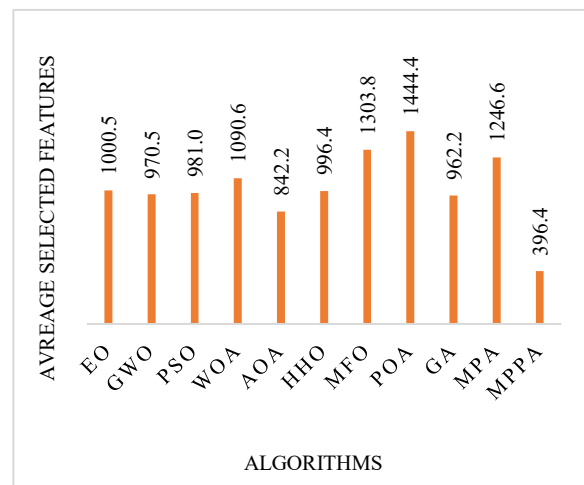


Figure 7. Average Feature selected

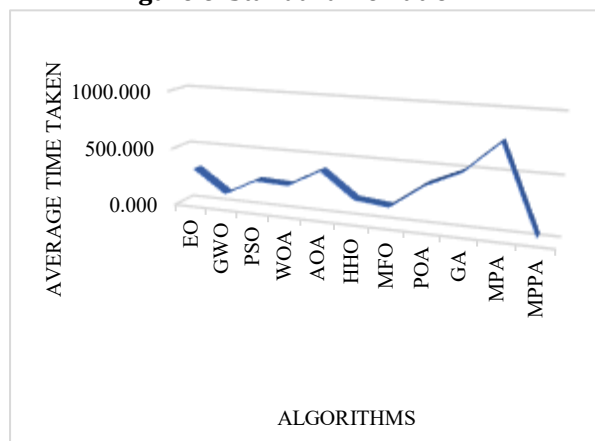


Figure 8. Average Computational time

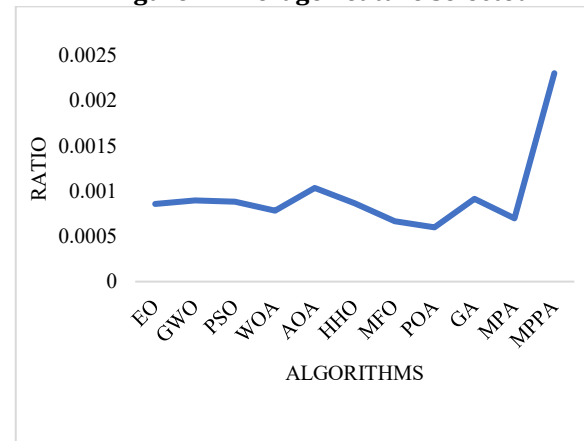


Figure 9. Ratio of Accuracy to Selected

Features

Similarly, figure 6 and 7 shows the results obtained for standard deviation and average number of features selected. MPPA outperforms the other algorithms in terms of selected features and furthermore figure 8 Shows the outperformance of the proposed algorithms in terms of average computational time, where MPPA has taken less time as compared to others. Figure 9 provides the results on ratio of average classification accuracy to average number of selected features. Furthermore, to validate the performance of proposed algorithms, eight standard datasets are considered for the experimentation, which are taken from the UCI [29] machine learning repository as shown in the table 3. All the datasets are from different domains and with different sizes in terms of number of features and instances.

Table 3. Datasets Considered

	Dataset	Features	Instances
D1	Waveform	40	5000
D2	breast-cancer	10	699
D3	Spect	22	267
D4	Tic-Tac-Toe	9	958
D5	banknote_authentication	5	1372
D6	Bupa (liver disorder)	7	345
D7	Dermatology	33	366
D8	Seeds	7	210

The proposed MPPA achieved the highest accuracy on most datasets. Table 4 presents the average classification accuracy indicating that the integration of chaotic behavior with BMFA and POA substantially enhanced the exploration and exploitation capabilities of the optimization process.

Table 4 Average Classification Accuracy on different datasets

	EO	GWO	PSO	WOA	AOA	HHO	MFO	GA	POA	MPA	MPPA
D1	0.781	0.781	0.787	0.788	0.779	0.777	0.803	0.826	0.798	0.840	0.847
D2	0.964	0.966	0.970	0.975	0.977	0.972	0.981	0.969	0.979	0.979	0.993
D3	0.721	0.709	0.713	0.721	0.728	0.743	0.757	0.743	0.738	0.795	0.937
D4	0.469	0.473	0.449	0.494	0.494	0.481	0.514	0.522	0.512	0.542	0.518
D5	0.996	1.000	0.996	0.998	0.999	1.000	0.999	0.996	0.997	1.000	1.000
D6	0.659	0.640	0.657	0.644	0.658	0.666	0.687	0.697	0.676	0.681	0.710
D7	0.700	0.697	0.751	0.718	0.738	0.718	0.774	0.754	0.738	0.768	0.794
D8	0.748	0.739	0.737	0.735	0.735	0.748	0.748	0.754	0.750	0.806	0.806

Tables 5 show the average fitness achieved through the proposed CBMFA-POA and competing optimization techniques. As a result of the fitness function, which is smaller for better optimizations, it can be seen that the proposed technique can be effectively applied to balance classification accuracy and feature selection.

Table 5 Average Fitness obtained on different datasets.

	EO	GWO	PSO	WOA	AOA	HHO	MFO	GA	POA	MPA	MPPA
D1	0.223	0.224	0.2194	0.218	0.227	0.228	0.2017	0.1770	0.2072	0.165	0.157
D2	0.042	0.040	0.0360	0.031	0.030	0.034	0.0244	0.0351	0.0262	0.026	0.010
D3	0.283	0.295	0.2905	0.282	0.277	0.263	0.2469	0.2591	0.2657	0.209	0.064
D4	0.531	0.526	0.5506	0.504	0.505	0.518	0.4853	0.4780	0.4875	0.458	0.482
D5	0.013	0.010	0.0140	0.012	0.010	0.009	0.0099	0.0118	0.0111	0.010	0.008
D6	0.345	0.360	0.3453	0.360	0.346	0.338	0.3152	0.3059	0.3258	0.321	0.294
D7	0.303	0.307	0.2536	0.286	0.266	0.285	0.2294	0.2476	0.2649	0.235	0.206
D8	0.256	0.264	0.2661	0.269	0.267	0.254	0.2558	0.2486	0.2538	0.198	0.198

Tables 6-8 present the standard deviation, average number of selected features, and F-measure results obtained by the proposed model and the other comparative algorithms considered for this work. The results demonstrate the effectiveness, stability, and feature selection capability of the proposed approach. Table 6 showcase the results of standard deviation which is indicating stable convergence behavior and consistent performance across multiple runs. The reduced variation in results demonstrates the robustness of the proposed optimization framework and its ability to avoid unstable solutions. Further table 7 shows the results of average selected features on various datasets, which indicates the proposed model removes the redundant and irrelevant features, leading to a more compact and efficient feature subset. The integration of chaotic maps and POA improves the search capability and enhances feature selection efficiency.

Table 6. Standard Deviation obtained on different datasets.

	EO	GWO	PSO	WOA	AOA	HHO	MFO	GA	POA	MPA	MPPA
D1	0.012	0.015	0.005	0.010	0.005	0.004	0.006	0.009	0.005	0.005	0.009
D2	0.003	0.007	0.004	0.006	0.005	0.007	0.003	0.008	0.007	0.000	0.000
D3	0.032	0.036	0.027	0.011	0.045	0.024	0.014	0.052	0.037	0.012	0.014
D4	0.034	0.024	0.018	0.045	0.021	0.034	0.023	0.032	0.021	0.000	0.018
D5	0.005	0.001	0.005	0.001	0.001	0.001	0.001	0.006	0.002	0.000	0.000
D6	0.026	0.020	0.016	0.043	0.054	0.022	0.014	0.030	0.023	0.000	0.000
D7	0.046	0.054	0.029	0.024	0.019	0.033	0.023	0.032	0.030	0.026	0.017
D8	0.025	0.012	0.028	0.014	0.027	0.021	0.020	0.023	0.018	0.000	0.000

Table 7. Average Number of Selected features obtained on different datasets.

	EO	GWO	PSO	WOA	AOA	HHO	MFO	GA	POA	MPA	MPPA
D1	26.20	28.40	33.00	31.40	34.20	31.20	28.60	18.80	29.60	26.40	22.80
	0	0	0	0	0	0	0	0	0	0	0
D2	6.000	5.800	6.000	6.600	7.200	6.400	5.200	4.600	5.800	5.000	3.000

D3	15.80 0	14.80 0	13.60 0	13.00 0	16.40 0	17.40 0	15.00 0	9.200	14.60 0	12.60 0	4.200
D4	3.800	3.600	4.000	3.000	3.000	3.400	3.000	3.800	3.400	3.000	4.000
D5	3.800	3.800	4.000	3.800	3.600	3.600	3.600	3.200	3.400	4.000	3.000
D6	4.200	2.600	3.200	4.400	4.200	4.400	3.000	3.600	3.200	3.000	4.000
D7	10.60 0	13.40 0	14.00 0	13.00 0	13.40 0	11.20 0	11.40 0	7.400	11.40 0	9.000	3.600
D8	1.800	1.600	1.800	1.800	1.400	1.400	1.800	1.600	2.000	2.000	2.000

The results confirm that the proposed framework enhances classification reliability and predictive performance. Overall, the obtained results demonstrate that the proposed MPPA provides stable optimization performance, efficient feature reduction, and improved classification quality compared to the existing metaheuristic algorithms.

5. Conclusions and Future Directions

This research introduced an innovative optimization strategy based on Chaotic Binary Marine Predators Algorithm integrated with Parrot Optimization algorithm (MPPA) in order to achieve optimal feature selection. The rationale behind this combination is the utilization of the powerful exploration skills of MPA combined with the intelligent exploitation behaviors of POA. Through the use of Brownian motion and Lévy flight techniques, MPA algorithm searches for the best solutions in the solution space; meanwhile, chaotic maps improve population diversification and avoid premature convergence problems. On the other hand, POA technique refines the candidate solutions by applying different behaviors that lead to more effective exploitation and better accuracy in achieving convergence. Experimental results show a higher level of efficiency of the introduced hybrid strategy as compared to independent algorithms in terms of various measures such as classification accuracy, Fitness, Average time, Average selected features, Time taken and F-measure.

For future endeavours, work can be done in enhancing the MPA by integrating it with other deep learning models such as Vision Transformers and adapting it for multi-objective optimization. The adaptive selection of chaotic maps could lead to improved convergence. Moreover, utilizing this method with multimodal datasets and real-time systems presents promising opportunities.

Author Contributions

All authors have contributed equally.

Funding

No funding was received.

Declarations

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. World Health Organization. (2022). Global cancer burden: Disparities and projections. International Agency for Research on Cancer. Retrieved from <https://www.iarc.who.int>
2. International Agency for Research on Cancer. (2022). GLOBOCAN 2022: Global cancer data and statistics. Union for International Cancer Control. Retrieved from <https://www.uicc.org>.
3. Smolarz, B., Nowak, A. Z., & Romanowicz, H. (2022). Breast cancer—epidemiology, classification, pathogenesis and treatment (review of literature). *Cancers*, 14(10), 2569.
4. Milosevic, M., Jankovic, D., Milenkovic, A., & Stojanov, D. (2018). Early diagnosis and detection of breast cancer. *Technology and Health Care*, 26(4), 729-759.
5. Nassif, A. B., Talib, M. A., Nasir, Q., Afadar, Y., & Elgendy, O. (2022). Breast cancer detection using artificial intelligence techniques: A systematic literature review. *Artificial intelligence in medicine*, 127, 102276.
6. Dromain, C., Boyer, B., Ferre, R., Canale, S., Delalogue, S., & Balleyguier, C. (2013). Computed-aided diagnosis (CAD) in the detection of breast cancer. *European journal of radiology*, 82(3), 417-423.
7. O'shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.

8. Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS journal of photogrammetry and remote sensing*, 173, 24-49.
9. Chauhan, R., Ghanshala, K. K., & Joshi, R. C. (2018, December). Convolutional neural network (CNN) for image detection and recognition. In 2018 first international conference on secure cyber computing and communication (ICSCCC) (pp. 278-282). IEEE.
10. Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, 3, 1-40.
11. Torrey, L., & Shavlik, J. (2010). Transfer learning. In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques* (pp. 242-264). IGI global.
12. Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43-76.
13. Lavanya, D., & Rani, D. K. U. (2011). Analysis of feature selection with classification: Breast cancer datasets. *Indian Journal of Computer Science and Engineering (IJCSE)*, 2(5), 756-763.
14. Aličković, E., & Subasi, A. (2017). Breast cancer diagnosis using GA feature selection and Rotation Forest. *Neural Computing and applications*, 28, 753-763.
15. Zang, H., Zhang, S., & Hapeshi, K. (2010). A review of nature-inspired algorithms. *Journal of Bionic Engineering*, 7(4), S232-S237.
16. Fister Jr, I., Yang, X. S., Fister, I., Brest, J., & Fister, D. (2013). A brief review of nature-inspired algorithms for optimization. *arXiv preprint arXiv:1307.4186*.
17. Sharma, M., & Kaur, P. (2021). A comprehensive analysis of nature-inspired meta-heuristic techniques for feature selection problem. *Archives of Computational Methods in Engineering*, 28, 1103-1127.
18. Amrir, M. M. S., Ayid, Y. M., Elshewey, A. M., & Fouad, Y. (2026). A hybrid LSTM-GRU framework for lung cancer classification using GWO-WOA algorithm for hyperparameter tuning and BPSO for feature selection. *Scientific Reports*.
19. Hans, R., Sharma, S. K., & Aickelin, U. (2024). Optimised deep k-nearest neighbour's based diabetic retinopathy diagnosis (ODeep-NN) using retinal images. *Health Information Science and Systems*, 12(1), 23.
20. Sappa, N., & Lingam, G. (2026). An Improved LNN-ACO Framework with Optimal Feature Selection for Breast Tumor Detection. *Engineering, Technology & Applied Science Research*, 16(1), 30710-30715.
21. Hans, R., & Kaur, H. (2022). Hybrid biogeography-based optimization and genetic algorithm for feature selection in mammographic breast density classification. *International Journal of Image and Graphics*, 22(03), 2140007.
22. Hassan, E., Saber, A., El-Sappagh, S., & El-Rashidy, N. (2025). Optimized ensemble deep learning approach for accurate breast cancer diagnosis using transfer learning and grey wolf optimization. *Evolving Systems*, 16(2), 59.
23. Attallah, O. (2026). MERGE: Mammogram-Enhanced Representation via Wavelet-Guided CNNs for Computer-Aided Diagnosis of Breast Cancer. *Machine Learning and Knowledge Extraction*, 8(2), 40.
24. Elshewey, A. M., Hassan, S. A., Youssef, R. Y., El-Bakry, H. M., & Osman, A. M. (2026). Enhancing breast cancer disease classification based on stacked machine learning models with hybrid Particle Swarm and Grey Wolf Optimization algorithms. *Iran Journal of Computer Science*, 9(1), 30.
25. Tiriyaki, V. M., & Tutkun, N. (2024). Breast cancer mass classification using machine learning, binary-coded genetic algorithms and an ensemble of deep transfer learning. *The Computer Journal*, 67(3), 1111-1125.
26. Alaca, Y., & Akmeşe, Ö. F. (2025). Pancreatic Tumor Detection From CT Images Converted to Graphs Using Whale Optimization and Classification Algorithms With Transfer Learning. *International Journal of Imaging Systems and Technology*, 35(2), e70040.
27. Bilal, A., Alkhatlan, A., Kateb, F. A., Tahir, A., Shafiq, M., & Long, H. (2025). A quantum-optimized approach for breast cancer detection using SqueezeNet-SVM. *Scientific Reports*, 15(1), 3254.
28. Saini, H., & Saini, D. S. (2025). Bio-inspired hybrid algorithms in deep learning model for the detection of COVID-19 and lung diseases. *International Journal of Bio-Inspired Computation*, 25(2), 104-112.
29. Gupta, M., Verma, N., Sharma, N., Singh, S. N., Brojen Singh, R. K., & Sharma, S. K. (2025). Deep transfer learning hybrid techniques for precision in breast cancer tumor histopathology classification. *Health Information Science and Systems*, 13(1), 20.
30. Falconi, L. G., Perez, M., Aguilar, W. G., & Conci, A. (2020). Transfer learning and fine tuning in breast mammogram abnormalities classification on CBIS-DDSM database. *Adv. Sci. Technol. Eng. Syst. J.*, 5(2), 154-165.

31. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, 42, 60-88.
32. Huang, G. B., Ding, X., & Zhou, H. (2010). Optimization method based extreme learning machine for classification. *Neurocomputing*, 74(1-3), 155-163.
33. Faramarzi, A., Heidarinejad, M., Mirjalili, S., & Gandomi, A. H. (2020). Marine Predators Algorithm: A nature-inspired metaheuristic. *Expert systems with applications*, 152, 113377.
34. Lian, J., Hui, G., Ma, L., Zhu, T., Wu, X., Heidari, A. A., ... & Chen, H. (2024). Parrot optimizer: Algorithm and applications to medical problems. *Computers in Biology and Medicine*, 172, 108064.
35. Sayed, G. I., Tharwat, A., & Hassanien, A. E. (2019). Chaotic dragonfly algorithm: an improved metaheuristic algorithm for feature selection. *Applied Intelligence*, 49, 188-205.
36. Chen, T., & Yi, Y. (2024). Multi-strategy enhanced parrot optimizer: global optimization and feature selection. *Biomimetics*, 9(11), 662.