JOURNAL OF SOFT COMPUTING AND DATA MINING
e-ISSN: 2716-621X

JSCDM

Vol. 6 No. 3 (2025) 307-322
https://publisher.uthm.edu.my/ojs/index.php/jscdm

An Enhanced Hybrid Binary Grey Wolf and Harris Hawk Optimization Algorithm Based on Cumulative Binomial Probability for Feature Selection in Classification

Manal Othman[1]*, Ku Ruhana Ku-Mahamud[1]

[1] *School of Computing, Universiti Utara Malaysia, 06010 Sintok, Kedah, MALAYSIA*

*Corresponding Author: manal.oshari@taiz.edu.ye
DOI: https://doi.org/10.30880/jscdm.2025.06.03.020

Article Info

Received: 20 September 2025
Accepted: 17 November 2025
Available online: 30 December 2025

Keywords

Classification, feature selection, Harris hawk optimization, grey wolf optimization, hybrid algorithm

Abstract

Feature selection acts as an essential preprocessing step to reduce dimensionality in datasets by eliminating redundant and insignificant features. It substantially reduces computational cost, improves interpretability, and enhances the accuracy of classification models. Hybrid binary grey wolf with Harris hawk optimization (HBGWOHHO) is one of the most promising hybrid algorithms for feature selection in classification. However, the HBGWOHHO algorithm has a limitation in unbalanced exploration and exploitation in achieving the sub-optimal solution. This limitation refers to the linearly declining value of a balancing parameter, which lacks regulation between the exploration and exploitation phases. This paper presents an enhanced HBGWOHHO that employs an adaptive technique based on cumulative binomial probability (CBP) called hybrid grey wolf Harris hawk optimization-CBP (HBGWHHO_CBP) to fine-tune the balancing parameter. This adaptive adjustment technique ensures a more effective balancing the algorithm's exploratory and exploitative behaviors, thus improving the search efficiency and solution quality. Dimension-wise diversity measure is utilized to quantitatively assess this balance during the optimization process. The effectiveness of the proposed HBGWHHO_CBP was evaluated by employing eleven UCI benchmark datasets. The proposed algorithm demonstrated superior performance across the evaluated datasets, yielding an average accuracy of 0.94, a mean of 8.51 selected features, and a mean fitness value of 0.06, while requiring less computational time. According to the Wilcoxon signed-rank test outcomes, the proposed algorithm significantly surpasses the native HBGWOHHO and three other metaheuristic-based feature selection algorithms. The proposed metaheuristic can be applied for addressing the feature selection in classification.

1. Introduction

Classification is a kind of supervised learning in machine learning (ML) where labelled data is utilized to train an algorithm to forecast the class of unseen and new data. There are various classification techniques suitable for different datasets and problem domains, each with its own strengths and weaknesses. The substantial rise in the number of datasets may complicate the management of classifier implementation, and it is essential to differentiate significant features from redundant and irrelevant ones [1]. The process of selecting significant features is known as feature selection (FS).

The FS is a crucial approach for dimensionality reduction. It chooses the significant features thereby minimizing the computational expense, enhancing the performance, simplifying the learned model and improving data understanding in pattern recognition applications [2]. The three fundamental categories of FS approaches are filter, wrapper, and embedded. In the filtering approach, feature subsets are selected independently of the learning algorithm. It offers advantages of low computational cost; however, the drawback of this approach is that disregarding the efficacy of the chosen features in improving the model performance. In the wrapper approach, feature subsets are selected and evaluated through iterative combinations, where the performance of a ML algorithm is employed as the evaluation criterion [3]. Unlike the filter and wrapper approaches, embedded approach selects features during the model training process. It integrates the classifier and the chosen features into a single procedure and selects the optimal subset features from dataset that contains all features. Reviews show that the wrapper method is widely used due to its efficacy in managing larger and more complicated datasets than the filter and embedded approaches; however, acquiring the ideal collection of features is still intractable.

The complexity of the features' interactions/relationships, and the complexity of the FS search space experiences exponential growth in tandem with a rise in feature dimensionality, where a dataset of n features can have a total of 2^n solutions, hence rendering FS an NP-problem. Consequently, it is essential to have a global search technique that efficiently handles FS task. Metaheuristics have emerged as highly effective methods for addressing FS due to their robustness, intelligibility, and efficiency in handling complicated optimization issues such as dipper throated optimization (DTO) [2], teaching learning based optimization (TLBO) [5] and particle swarm optimization (PSO) [4]. Despite metaheuristics perform well in addressing the FS, they have limitations in balancing exploration and exploitation, premature convergence and trapped in local optimum. To mitigate these drawbacks, recent studies have proposed hybrid metaheuristic algorithms as effective methods for FS that combine the strengths of both algorithms, for instance novel hybrid TLBO with salp swarm algorithm (SSA) [6], brain storm optimization with firefly algorithm (FA) [7], improved binary competitive swarm optimization (IBCSO) with whale optimization algorithm (WOA) [8] and Harris hawk optimization (HHO) with WOA (BWOAHHO) [9]. Although hybridization improves algorithmic performance, most hybrid algorithms still rely on iteration-based decays to balance exploration and exploitation, rather than feedback from search process or diversity.

The hybrid binary grey wolf (GWO) with HHO (HBGWOHHO) is one of the most promising hybrid algorithms for FS [10], which capitalizes on the respective advantages of both HHO and GWO for FS. Despite the better performance, the HBGWOHHO algorithm is incapable of efficiently balancing global exploration and local exploitation of the promising areas in the search space because this process has been controlled by using the parameter "a", which decreases linearly. The linear decrement strategy hinders the algorithm's capability of efficiently exploring and exploiting the search area because it does not depend on the search agents' feedback. Finding a suitable value for parameter "a" is crucial to achieve a proper equilibrium between exploring the search space and exploiting promising areas to find good solutions efficiently. This study proposes a novel feedback-based technique, based on cumulative binomial probability (CBP), to adaptively maintain balance between exploration and exploitation of a hybrid HBGWOHHO method. This adaptive technique guides the search more efficiently toward high-quality solutions.

The following sections outline the structure of this study: the review of relevant works is discussed in Section 2, while Section 3 provides the description details of native HBGWOHHO algorithm, the proposed HBGWHHO_CBP description and analyses the computational complexity of the proposed HBGWHHO_CBP. Section 4 provides the exploration and exploitation assessment by using dimension-wise diversity metric. In Section 5 the data and experimental design are provided, and experimental findings are presented and discussed in Section 6. Lastly, the study's conclusions and potential directions for future research are stated in Section 7.

2. Related Work

Hybrid metaheuristic algorithms have emerged as effective methods for FS, as they enhance the investigation of the search space, helping to determine an ideal subset of features, and improve predictive performance. Generally, hybridization algorithms enable algorithms to enhance each other's strengths and alleviate their weaknesses. For instance, the conventional AOA has drawbacks, including the limited local search ability and the balance between exploration and exploitation. Its performance was improved by using the genetic algorithm's (GA) operations [11]. Due to the effectiveness of integrating GA operators, the hybrid AOAGA algorithm achieved better accuracy in FS tasks and explored more regions in the solution space compared to the native AOA algorithm. TLBO was modified by hybridizing it with SSA [6]. This Integration gives TLBO more flexibility to explore the search space and reach the near-optimal solution. Similarly, the exploration ability limitation of WOA is mitigated by incorporating HHO exploration mechanisms into the binary WOA algorithm, where exploration is guided by the HHO, while exploitation remains guided by the WOA [9]. The hybrid strategy considerably improved the performance of WOA, avoiding premature convergence and increasing the search diversity. On the other hand, IBCSOWOA algorithm [8] is proposed by hybridizing the (IBCSO) with (WOA) to enhance the quality of classification results and expedite the procedure of gene chosen. The significant features are selected by IBCSO while WOA is applied to fine-tune the classifier (ANN) hyperparameters. These hybrid algorithms demonstrate varied strategies for solving FS problem by integrating complementary algorithms. However, most of them fail to address the adaptive adjustment of parameters during the optimization process, which is critical to enhance algorithmic robustness and improve overall optimization outcomes [12].

One of the modern bio-inspired metaheuristics is grey wolf optimization (GWO); it has garnered considerable attention in the domain of hybrid metaheuristics for FS like hybrid GWO with butterfly optimization algorithm [13], GWO with DTO [14], GWO with HHO [10]. The efficacy of GWO is influenced by the balance parameter "a". This parameter regulates the exploration and exploitation balance. Its value decreases linearly from a maximum of 2 to a minimum of 0. As a result, the search has uniform behavior [15]. Several studies have proposed different techniques to effectively adjust the balance coefficient "a" including exponential adjustment in hybrid GWO with stochastic fractal search [16] and nonlinear adjustment in hybrid GWO with crow search algorithm [17] and PSO_GWO [18]. However, in these techniques, there is no automatic adjustment of control parameters based on agent fitness.

Recently, researchers have expressed interest in dynamic adjustment of the balance between exploration and exploitation beyond just linear, nonlinear, or exponential adjustments. Shaikh, et al. [19] introduced a balanced and adaptive strategy using hybrid PSO and GWO, where the exploratory capabilities of the GWO algorithm are combined with the exploitative efficiency of PSO and the inertia constant regulates the exploration and exploitation throughout the search process. Despite the fact that this algorithm increases adaptability, it has multiple parameters that must be carefully adjusted for best results, adding to the overall complexity. A dynamical regulation strategy was introduced to fine-tune the convergence factor of the GWO in the hybrid rice optimization algorithm suggested in Ye, et al. [12]. However, this regulation depends on iteration rather than solution, which makes it more dynamic than truly adaptive. Ahmad, et al. [20] proposed the GWO-Employed-Onlooker model and its performance evaluated by using different optimization functions. GWO-Employed-Onlooker incorporates the scout and onlooker bee operators from the ABC algorithm during the positions-update phase of the wolves to raise exploitation ability and enhanced local convergence rates. Nevertheless, it lacks real adaptability, as it fails to change dynamically based on feedback from the optimization process. Consequently, current algorithms need improvement by altering their operators to more effectively in addressing FS and other optimization problems.

3. Methods

3.1 Hybrid Binary Grey Wolf Harris Hawk Optimization Algorithm

The HBGWOHHO algorithm indicates a combination of the GWO algorithm with the HHO algorithm. In essence, this hybridization employs the exploration phase from HHO and the exploitation phase from GWO. It was motivated by the highly accuracy findings generated by HHO and the simplicity and efficiency of GWO [10]. The HHO algorithm, proposed by Heidari, et al. [21], is one of the recent bio-inspired metaheuristic algorithms. The essential concept at the core of HHO involves emulating the coordinated hunting behavior of a hawk team and the evasive actions of prey in nature. Within the HHO, the chasing actions of hawks act as the search agents, with the prey embodying the optimal position [22]. HHO exhibits notable efficacy in addressing various real-world optimization issues [23]. Furthermore, HHO can provide better solution quality and enhance prediction performance [24]. As a result of increasing interest, several researchers have designed hybrid algorithms by integrating HHO with other metaheuristics. It has been hybridized with the binary GWO to tackle the challenges of local stagnation and premature convergence which could lead to an incorrect solution in the native binary GWO.

The GWO has demonstrated efficacy in solving different single and multi-objective optimization issues [25]. It imitates the social structure of gray wolves, which naturally form packs of 5 to 12 individuals. The leadership of a wolf pack is categorized into three hierarchical levels: alpha (α), beta (β), and delta (δ). On the other hand, the remaining members of a group are classified as Omega ($\acute{\omega}$). Mathematically, the implementation of GWO consists of three primary phases, namely encircling, hunting, and attacking prey.

Grey wolves frequently encircle their prey in order to exhaust and slow them down. The encircling phase in GWO is mathematically modelled in Eq. (1-5) [25]:

$$D = \left|C.Y_p(t) - Y(t)\right| \tag{1}$$

$$Y(t+1) = Y_p(t) + A.D \tag{2}$$

Where t represents the current iteration, Y_p refers to the prey's position, and coefficient vectors A and C are calculated as shown in Eq. (3-5) [25]

$$A = 2a.r_1 - a \tag{3}$$

$$a = 2 - \frac{2t}{T} \tag{4}$$

$$C = 2.r_2 \tag{5}$$

Here, r_1 and r_2 are random vectors, with values ranging from 0 to 1. They are utilized to introduce some randomness to the algorithm. In contrast, parameter a is responsible for the exploration-exploitation equilibrium. Its value declined linearly throughout the iterations from "2" to "0". The alpha (α) wolf leads the entire hunting process. The entire pack of grey wolves participates in the hunting depend on the behavior of the top three wolves (α, β, δ), with their positions are updated according to the optimal resultant positions of α, β, δ wolves. The formulation is mathematically presented as follows [25]:

$$\begin{aligned} D_\alpha &= |C_1.Y_\alpha - Y|, \\ D_\beta &= |C_2.Y_\beta - Y|, \\ D_\delta &= |C_3.Y_\delta - Y| \end{aligned} \tag{6}$$

$$\begin{aligned} Y_1 &= Y_\alpha - A_1.(D_\alpha) \\ Y_2 &= Y_\beta - A_2.(D_\beta) \\ Y_3 &= Y_\delta - A_3.(D_\delta) \end{aligned} \tag{7}$$

$$Y_{(t+1)} = \frac{1}{3}\sum_{k=1}^{3} Y_k \tag{8}$$

where Y_α, Y_β and Y_δ denote to the location of the first, second and third optimal solutions respectively, while the positions of the three optimal search agents are updated according to D_α, D_β and D_δ. Eq. (6) and Eq. (7) generate three position vectors for delta, beta, and alpha wolves. The current location of the grey wolf (Y_{t+1}) can be determined using averaging of three vectors in Eq. (8).

To resolve the problem of dropping in local optimal and premature convergence, the HBGWOHHO used the exploration from the HHO by using Eq. (9) and the exploitation phase from GWO by using Eq. (8). The process begins with the initialization of the value of coefficient vector A. Its value is derived from the balance parameter 'a', which regulates the position updates of agents by directing their movements toward the prey, thereby impacting both exploration and exploitation. If absolute value of $A \geq 1$, the exploration phase, of HHO, commences and the search agents used two strategies to determine its position during this phase. The first strategy is to perch on random location inside their group home if $r \geq 0.5$, while the second strategy is to be close to other family members' location if $r < 0.5$ in Eq. (9) [10].

$$Y(t+1) = \begin{cases} Y_{rand}(t) - C|Y_{rand}(t) - 2r_1 Y(t)| & , \ r \geq 0.5 \\ [Y_p(t) - Y_m(t)] - C[r_2(UB - LB) + LB], & r < 0.5 \end{cases} \tag{9}$$

$$C = 2 * r_3 \tag{10}$$

where $Y(t+1)$ indicates the agent's position vector at the next iteration, Y_{rand} represents the random agent's position in population, $Y(t)$ is the selected agent's current position vector, Y_m denotes the search agents' mean location, and Y_p denotes the position of the prey (optimal solution). The coefficient parameter C, calculated by using Eq. (10), improves the randomness of the algorithm to prevent local optima. Lastly, the random variables r, r_1, r_2, and r_3 range between 0 and 1 [10].

A one-dimensional vector is used to demonstrate the solution. This vector's length is equal to the feature size. This vector is binary where selected features are represented by "1" and other features are represented by "0". The classification accuracy can be optimized by eliminating irrelevant and/or redundant features while preserving only the significant ones. The Eq. (11) expresses the fitness function that is utilized in HBGWOHHO, aiming at the reduction of classification error alongside the size of the chosen feature subset [10, 14].

$$Fitness = \alpha * ER + (1 - \alpha)\frac{|S_f|}{|T_f|} \tag{11}$$

where $\alpha \in [0,1]$, ER indicates the classification error rate of the K-nearest neighbor (KNN) classifier. Sf indicates the number of selected feature subsets, while Tf is the whole number of features in datasets. To satisfy the requirement of FS binary nature, the continuous solution space must be transformed into a binary representation. To this end, the sigmoid transfer function is applied according to Eq. (12) [10]:

J. of Soft Computing and Data Mining Vol. 6 No. 3 (2025) p. 307-322

$$Y_{Binary}(t+1) = \begin{cases} 1, & sigmoid\left(\frac{Y_1+Y_2+Y_3}{3}\right) \geq r \\ 0, & otherwise \end{cases} \tag{12}$$

where Y_{Binary} is the binary update position technique for next iteration, r is a random value with a standard uniform distribution ($i.e.,\ r \in [0,1]$), and sigmoid is a function represented in Eq. (13)[10]:

$$sigmoid(Y) = \frac{1}{1+e^{-10(Y-0.5)}} \tag{13}$$

3.2 Proposed Adaptive Technique for Balancing Parameter Control

The exploration and exploitation balancing plays an essential role in determining the effectiveness and capability of the metaheuristic algorithms [26-28]. In the HBGWOHHO algorithm, the strategy of regulating the exploration-exploitation balance is controlled by the GWO parameter "a", with a linearly decreasing value during the search process. This linear decrement technique fails to depict the actual convergence procedure of the algorithm because it does not depend on the search agents' feedback. This limits the algorithm's ability of efficient exploration and exploitation of the solution space [12]. Finding a suitable value for parameter "a" is crucial to balance exploration and exploitation effectively and find good solutions efficiently. This study proposes an adaptive adjustment of the balance parameter "a" using a novel technique, CBP, as shown in Fig. 1, with the aim of achieving an improved equilibrium between exploration and exploitation. Hence, this enhancement makes the algorithm effectively mitigates premature convergence and converge toward the global optimum.

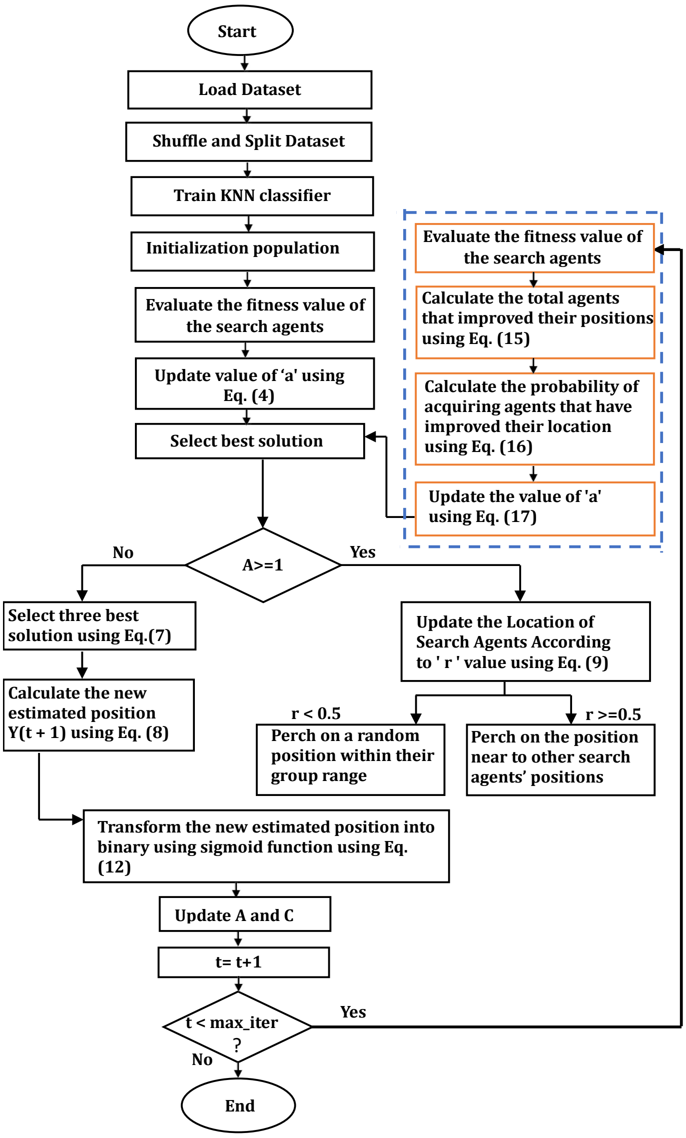

Fig. 1 *General model of the proposed HBGWHHO_CBP*

The pseudocode for the overall procedure of our proposed HBGWHHO_CBP is illustrated in Fig. 2.

```
Algorithm 1 Pseudo-Code of the proposed HBGWHHO_CBP Algorithm
Input: number of search agents (N) and maximum number of iterations (T)
Output: the position of the optimal solution and its value
-----------------------------------------------------------------------------------------------
1: Initialize search agents population Yi (i = 1, 2, ....., n)
2: Initialize parameters a, A, C, and r
3: For (each agent)
4:     Evaluate the fitness value of the search agents
5: End for
6: Identify the best three agents based on fitness value of the search agents Yα, Yβ and Yδ
7: Set the best position (Yα) as the position of prey (YP)
8: Update the value of a using Eq. (4)
9: For (each agent)
10:     Update the absolute value of A using Eq. (3)
11:     If (|A| ≥1) then perform the Exploration phase
12:         If r ≥ 0.5 then  ------------ search agents perched on a random position -------------------
13:          Update the position of the search agent using Eq. (9)
14:         else if r < 0.5  ---------search agents perched near to other search agents positions---
15:         Update the position of the search agent using Eq. (9)
16:         end if
17:          Transform the new estimated position into binary by applying sigmoid function Eq. (12)
18:     else if (|A| < 1) then perform the Exploitation phase
19:         Update the location of Yα, Yβ, and Yδ by using Eq. (7)
20:         Calculate the new estimated position Y(t + 1) using Eq. (8)
21:         Transform the new estimated position into binary by applying sigmoid function Eq. (12)
22:     end if
23:     Update values of coefficient parameters
24:      t = t + 1
25: End for
26: Calculate the total agents that improved their positions using Eq. (15)
27: Calculate the probability of acquiring agents that have improved their location using Eq. (16)
28: Update the value of 'a' using proposed Eq. (17)
29: Repeat steps 9-28 until reach maximum number of iterations
30: Return best solution (YP)
```

Fig. 2 *HBGWHHO_CBP Pseudo-code for FS*

The CBP technique involves determining the population's status after each iteration to calculate the value of the "a" parameter. Eq. (14) shows the proposed indicator function that indicates the search remains focused on superior solutions only if the improvements in the position of agent Y are significant and persist over time t. On other words, solution is considered to have meaningful improved if its new fitness value is better (smaller) than the previous one by more than a predefined tolerance δ.

A tolerance value of 10^{-7} ,which is commonly used in previous studies across various problems like [29], is used in this study where the smaller values ($\delta<10^{-7}$) tend to increase computation time without significant improvements in accuracy.

$$I(Y,t) = \begin{cases} 1, & if \;\; fitness(Y^{t-1}) - fitness(Y^{t}) > \delta \\ 0, & otherwise \end{cases} \tag{14}$$

At iteration t, total number of agents with improved positions was calculated as shown in Eq. (15) [30].

$$sum(Y,t) = \sum_{i=1}^{N} I(Y,t) \tag{15}$$

where t is the current iteration while N represents the population size, it is constant during all iterations. Consider each iteration a separate experiment in which every agent symbolizes a trial that may have either state (a failure or a success). Since the trials are independent, each agent has a probability of success (p) and failure (q) 0.5. The total number of agents with a better position *Sum(Y,t)* is a binomial random variable.

The probability of acquiring a *sum(Y,t)* or fewer agents that have improved their location is known as the cumulative binomial probability and is represented by the Eq. (16) [30].

$$P_r(k \le sum, N, p) = \sum_{k=0}^{sum} \binom{N}{k} p^k \, q^{N-k} \tag{16}$$

where p and q refer to the success probability and failure probability in single trial, k denotes the number of successes, and *sum* is calculated in Eq. (15). In the proposed algorithm, the linear function with the probability P_r was utilized to calculate parameter "a" according to Eq. (17):

$$a = (2 - \frac{2t}{T}) * P_r(k \leq sum, N, p) \quad (17)$$

Here, the current iteration is represented by t, while the maximum number of iterations is denoted by T. Generally, the balancing parameter "a" should not be too low so that agents may not explore the search space effectively and, accordingly, fall into local optimum. In contrast, "a" should not be excessively large so that they could avoid the area that has the global optima. The control parameter "a" is kept relatively high to maintain exploration to find a better area if the probability of finding *sum(Y,t)* or fewer agents whose positions have improved is high. On the other hand, the value of "a" is decreased to cause the search behavior to move toward exploitation when this probability is low, indicating widespread improvement.

3.2.1 Computational Complexity

The proposed HBGWHHO_CBP's computational complexity depends on three basic steps. These steps include population initialization, fitness assessment, and search agents updating. Given that N is the number of agents and T represents the maximum number of iterations, below is described the overall complexity of the algorithm:

- Population and parameter Initialization: $O(N)$.
- Fitness function evaluation for all agents over iterations: $O(N \times T \times C_{fit}(I,d))$ where C_{fit} is the cost of fitness function, I is the number of instances, and d is the feature subset size in dataset.
- Update balancing parameter a by using CBP technique: $O(T)$.
- Update the positions of each agent over iterations: $O(N \times T)$.
- Convert agents positions into binary: $O(N \times T)$.

According to this analysis, the computational complexity of the proposed algorithm is $O(N \times T(C_{fit}(I,d) + 1))$.

4. Quantitative Assessment of Exploration and Exploitation

Exploration and exploitation are two fundamental concepts in metaheuristic optimization; when effectively balanced, these algorithms can achieve superior performance and facilitate successful convergence. Most existing evaluations of metaheuristic algorithms are limited to observing convergence graph and final results comparisons, which fail to examine the quality of exploration-exploitation balance during the search process.

This study uses the dimension-wise diversity metric, adapted from Zhao, et al. [31] and Morales-Castañeda, et al. [32], to conduct a thorough analysis of the algorithm's search behavior.

In metaheuristic optimization, the search agents which have the optimal solutions typically attract the search process towards them. This attraction causes a reduction in the distance between individual agents in the search space and, consequently, the influence of exploitation increases. Conversely, the exploration process has a greater influence when the distance between search agents rises [32]. The spread (or variety) of solutions explored by the algorithm is measured by population diversity. Solutions that are consistently diverse point to a decreased chance of being stuck in local optimum.

More exploration is indicated by higher diversity, whereas exploitation is indicated by reduced diversity. The changes in diversity between search agents, both in terms of increase and decrease, can be determined using the dimension-wise diversity measure. The population diversity is defined as shown in Eq. (18) and Eq. (19) [32]

$$Div_j = \frac{1}{n}\sum_{i=1}^{n} \left| median(Y^j) - Y_i^j \right| \quad (18)$$

$$Div = \frac{1}{m}\sum_{j=1}^{m} Div_j \quad (19)$$

where *median(Y^j)* denotes the median of the jth dimension with respect to the entire population. The variable Y_i^j denotes the jth dimension of the search agent, m is defined as the total number of design variables in the optimization problem, whereas n represents the number of search agents in the population.

The diversity in every dimension Div_j is defined as the average distance between the dimension j of each search agent and the median of that dimension. *Div* is the overall diversity of the population that determined by computing the average of each individual diversity value, Div_j , across all dimensions. Both values are calculated during each iteration.

A specific metaheuristic scheme's percentage of exploration and exploitation is used to characterize the whole balancing response. In every iteration, these values are calculated using Eq. (20) and Eq. (21) [32].

$$XPL\% = \left(\frac{Div}{Div_{max}}\right) * 100 \tag{20}$$

$$XPT\% = \left(\frac{|Div - Div_{max}|}{Div_{max}}\right) * 100 \tag{21}$$

where the exploration and exploitation percentages are represented by *XPL%* and *XPT%*, respectively, while Div_{max} refers to the value of the maximum diversity, discovered during the whole optimization process.

5. Experimental Design

To examine the efficacy of the suggested HBGWHHO_CBP, its performance was evaluated against the native HBGWOHHO algorithm based on the classification accuracy, feature subset size, mean fitness, computational time and exploration percentage. These criteria are essential to evaluate the efficacy of the FS process because they reflect the classification quality, optimization performance, reduction efficiency, and computational cost. The native HBGWOHHO algorithm was re-implemented according to its descriptions in [10] and executed with experimental settings identical to those of the proposed HBGWHHO_CBP. This ensures a fair comparison of improvements in reaching a better exploration-exploitation balance under the same settings, namely the same device, datasets, data splits, and parameter settings. The benchmark datasets, publicly available in the UCI Repository, are utilized for evaluation [32]. Table 1 provides a detailed overview of all datasets, which have been extensively utilized in previous studies to assess the efficacy of FS in classification [8, 14, 23, 33].

Table 1 *Datasets considered for experimental analysis*

Notation	Dataset	#Features	#Instances
D1	Congress-EW	16	434
D2	Exactly	13	1000
D3	Exactly-2	13	1000
D4	Ionosphere-EW	34	351
D5	Krvskp-EW	36	3196
D6	Lymphography	18	148
D7	M_of_n	13	1000
D8	Sonar-EW	60	208
D9	Tic_tac_toe	9	958
D10	Vote	16	300
D11	Zoo	16	101

The specifications of the machine used for the experiments are as follows: a desktop computer with an Intel(R) Core i7-4770 CPU running at 3.4 GHz. Each dataset is partitioned into 80% for training purposes, whereas the remaining 20% is reserved for testing and evaluating algorithm performance. The training and testing subsets are randomly shuffled and repeated to aid in reducing bias that might occur if the data possesses any natural ordering. The wrapper-based KNN classifier with Euclidean separation matric (where K=5) is employed to assess the selected feature subsets. The choice of KNN with k equal to 5 is motivated by its frequent use in previous studies that employed hybrid metaheuristics for FS in classification, as these studies reported in [33]. The proposed algorithm is repeated M times with random seed for T iterations, while the total number of search agents is N. The parameters N, M, LB, UB, α, and K were defined based on the conventional settings established in [10], as presented in Table 2.

J. of Soft Computing and Data Mining Vol. 6 No. 3 (2025) p. 307-322

Table 2 *Parameter configuration for the proposed HBGWHHO_CBP*

Parameter	Value
Population Size (N)	10
Maximum iterations (T)	1000
Number of repeated runs (M)	20
probability of success (p)	0.5
Fitness function constant (α)	0.99
Upper boundary of the search space (UB)	1
lower boundary of the search space (LB)	0

6. Results and Discussion

In this section, the findings of the proposed HBGWHHO_CBP are illustrated and discussed. Table 3 displays the findings in terms of average accuracy, mean size of chosen feature subset, mean fitness, and average computational time after repeating both algorithms 20 times in same device. The high average accuracy indicates that the HBGWHHO_CBP has identified the appropriate feature set that has the ability of training the classifier successfully while achieving a good accuracy on the testing data. The highest recorded average accuracy was 100 %, attained in the D7 dataset due to its well-structured nature, and it has only 13 features and 1000 instances. Followed by D5 and D11 achieved an average accuracy of 99%.

Table 3 *Results of the proposed HBGWHHO_CBP compared with native HBGWOHHO*

Dataset	Algorithm	Accuracy %	Selected feature %	Mean fitness %	Computational time %	Exploration percentage%	Exploitation percentage%
D1	HBGWHHO_CBP	**0.98**	**4.6**	**0.02**	**27.04**	19.36	80.64
	HBGWOHHO	**0.98**	6.6	0.03	29.31	25.81	74.19
D2	HBGWHHO_CBP	**0.96**	**5.9**	**0.04**	**59.54**	21.48	78.52
	HBGWOHHO	0.78	10	0.22	76.50	25.46	74.54
D3	HBGWHHO_CBP	**0.77**	**3.7**	**0.22**	**52.24**	31.16	68.84
	HBGWOHHO	**0.77**	6.1	**0.22**	63.67	36.43	63.57
D4	HBGWHHO_CBP	**0.96**	9.1	**0.04**	**25.60**	15.38	84.62
	HBGWOHHO	0.95	**4.8**	**0.04**	26.41	28.18	71.82
D5	HBGWHHO_CBP	**0.99**	**21.5**	**0.02**	**385.63**	15.13	84.87
	HBGWOHHO	0.97	31.7	0.04	387.01	21.18	78.82
D6	HBGWHHO_CBP	**0.92**	**6.7**	**0.08**	**14.39**	19.73	80.27
	HBGWOHHO	0.90	8.4	0.10	15.09	25.10	74.90
D7	HBGWHHO_CBP	**1**	**6.2**	**0.00**	**59.02**	19.83	80.17
	HBGWOHHO	0.93	10.4	0.08	77.98	23.49	76.51
D8	HBGWHHO_CBP	**0.94**	**18.1**	**0.06**	19.03	13.48	86.52
	HBGWOHHO	0.90	29.9	0.10	**18.61**	22.74	77.26
D9	HBGWHHO_CBP	**0.84**	**7.2**	**0.17**	**59.91**	24.16	75.84
	HBGWOHHO	**0.84**	8.8	**0.17**	66.65	21.43	78.57
D10	HBGWHHO_CBP	**0.98**	**4.4**	**0.02**	**21.15**	21.39	78.61
	HBGWOHHO	0.97	5.2	0.03	21.81	27.67	72.33
D11	HBGWHHO_CBP	**0.99**	**6.2**	**0.01**	12.99	19.11	80.89
	HBGWOHHO	0.96	9.3	0.05	**12.73**	24.98	75.02
Avg. all	HBGWHHO_CBP	**0.94**	**8.51**	**0.06**	**66.96**	20.02	79.98
	HBGWOHHO	0.90	11.93	0.10	72.34	25.68	74.32

Accuracy and feature subset size comparison: Comparative analysis with native HBGWOHHO demonstrates that the suggested HBGWHHO_CBP achieved improved classification accuracy in 8 out of the total

11 datasets considered in this work. Despite the suggested HBGWHHO_CBP achieved similar classification accuracy to the native HBGWOHHO on three datasets (D1, D3, and D9), it utilized fewer features, suggesting a more efficient FS strategy that maintains competitive performance. The HBGWHHO_CBP chose fewer features than BHGWOHHO across all datasets except one dataset (D4), where the differences were relatively small the HBGWHHO_CBP selected 9.1 features on average compared to 4.8 by HBGWOHHO. In contrast to these small differences, HBGWHHO_CBP exhibited significantly greater performance in other datasets, such as selecting only 21.5 and 18.1 features compared to HBGWOHHO's 31.7 and 29.9 in D5 and D8 datasets respectfully, indicating a considerable reduction in feature set size.

Mean fitness and computational time comparison: HBGWHHO_CBP outperformed BHGWOHHO on most datasets, achieving lower mean fitness values in seven out eleven datasets and same results in three datasets (D3, D4, and D9). The proposed HBGWHHO_CBP required less computing time than the native HBGWOHHO. This reduction is essentially due to the utilize the adaptive technique, as adaptive adjustment of the balancing parameter "a" has been proven to reduce in unnecessary exploration, save computational time, and improve solutions compared to linear decrement technique [34]. By adaptively tuning the balance parameter "a" depending on the feedback from search process, HBGWHHO_CBP has a lower percentage of exploration compared to HBGWOHHO, as shown in Table 3, suggesting that the algorithm tends to intensively search within local regions to uncover potential global optimal solutions.

Convergence behavior comparison: The proposed HBGWHHO_CBP exhibits better convergence and reaches better fitness values than the native HBGWOHHO across almost all datasets. For clarity, Fig. 3 shows the convergence curves for the first 600 iteration. Although the average accuracy of HBGWHHO_CBP and HBGWOHHO in D1 and D3 are similar, HBGWHHO_CBP converges more effectively, achieves higher stability, and selected fewer features than HBGWOHHO. This improved convergence is due to the incorporation of adaptive technique CBP which guides the search process to explore additional regions when the search agents succeed in improving their positions, indicating that the algorithm continues investigating further promising areas of the search space.

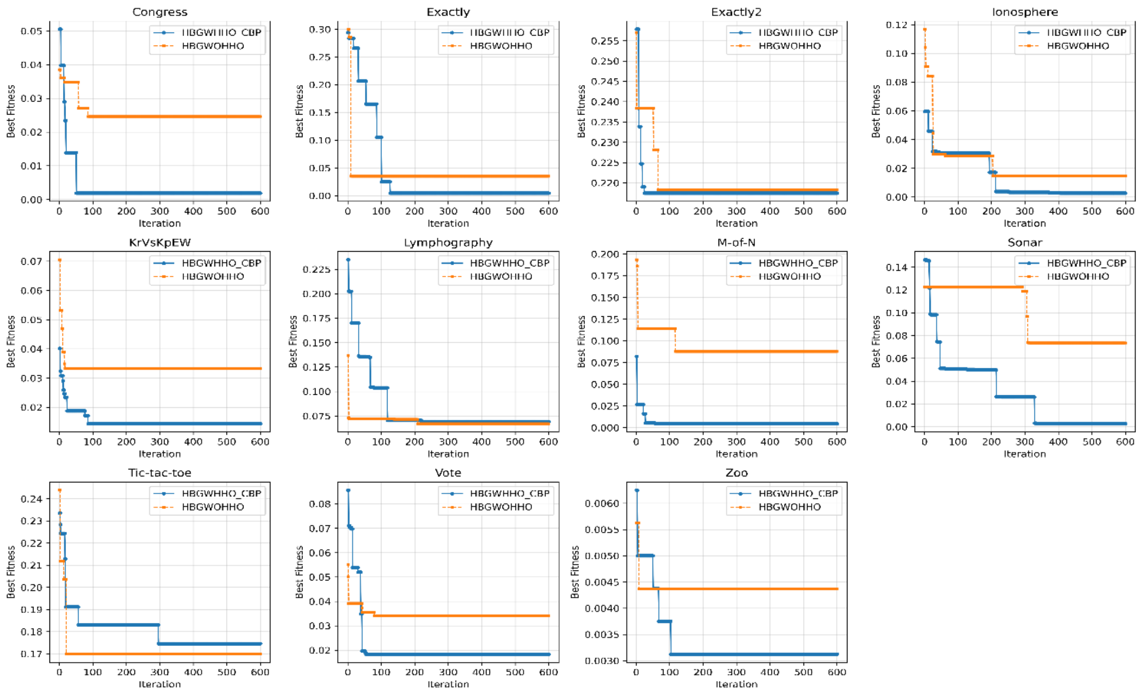

Fig. 3 *Convergence curve of proposed HBGWHHO_CBP compared with native HBGWOHHO*

Diversity comparison: In all experiments, the diversity (refer to Eq. (19)) is computed and reported throughout the optimization procedure. To analyze the diversity, exploration and exploitation behavior, we chose only the graphical results of four datasets (D5, D7, D8 and D11) that demonstrated a significant performance disparity between the proposed HBGWHHO_CBP and the native HBGWOHHO algorithms, facilitating a clearer and more meaningful comparison. According to Fig. 4, both algorithms exhibit a high degree of diversity due to random initialization, whilst the diversity of population diminishes as the number of iterations rises.

The proposed HBGWHHO_CBP algorithm exhibits high oscillations in its behavior, reflecting efficiently diversify the search in early phases and subsequently focus its search on promising areas in later phases. Such large oscillations in algorithmic behavior indicate improved control over the exploration and exploitation process [32]. Although HBGWOHHO showed higher starting oscillations, its balance is not dynamically modified, which raised the risk of being stuck in local optimal.

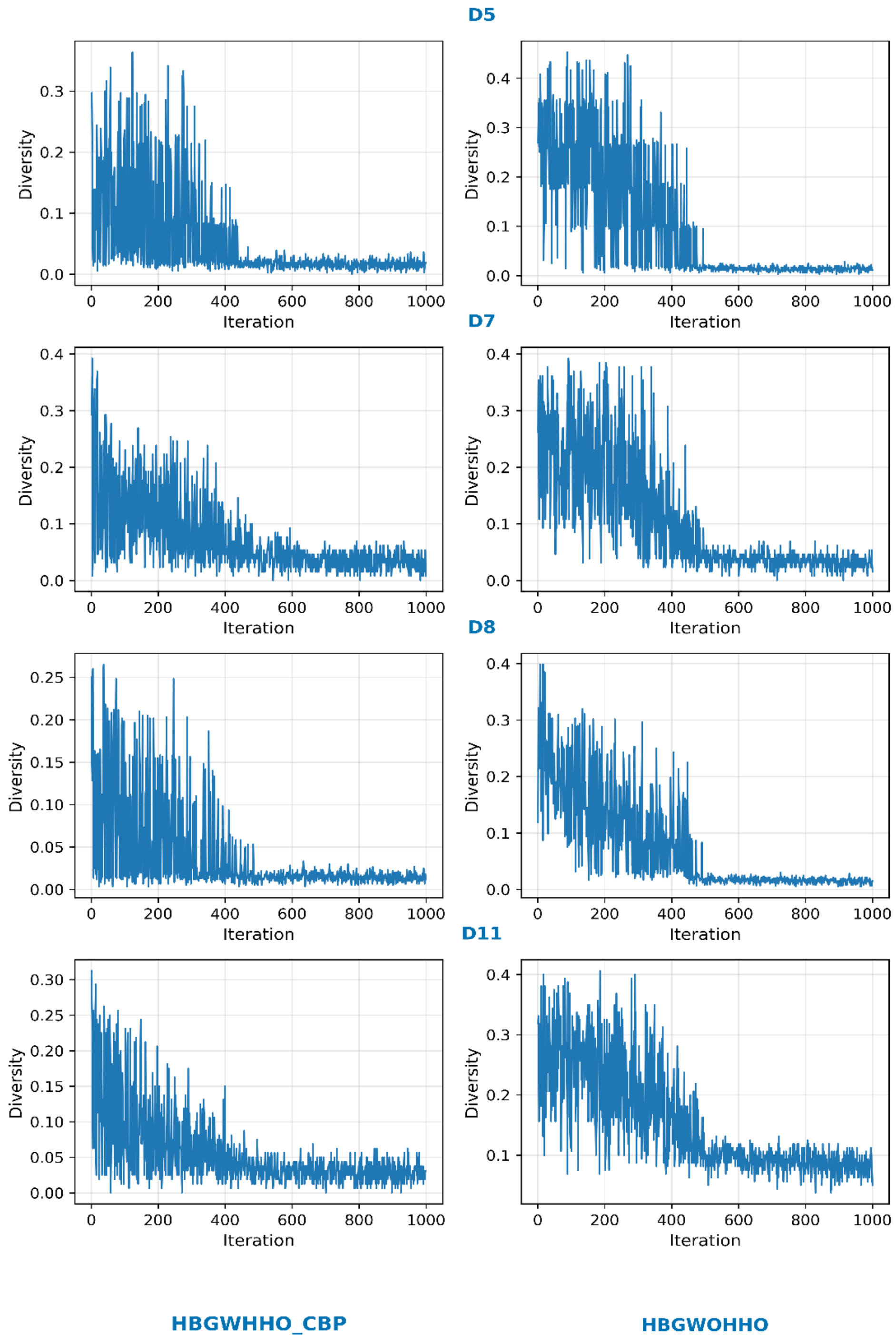

Fig. 4 *Diversity behavior of proposed HBGWHHO_CBP compared with native HBGWOHHO*

Assessing the generalizability of selected feature subsets: Due to the fact that the wrapper-based FS is frequently closely associated with the selected learner, the generalizability of the feature subsets selected by the proposed HBGWHHO_CBP-based KNN was evaluated using alternative classifiers: artificial neural network (ANN), support vector machine (SVM), and decision tree (DT). The proposed HBGWHHO_CBP-based KNN demonstrated significant robustness, as it achieved perfect robustness on the D7, with all classifiers (SVM, ANN, and DT) attaining 100% accuracy using the same selected subset. As shown in Table 4, the accuracy decrease among alternative classifiers was typically less than 5% in 8 of 11 datasets, which means that the model generalizes well. The drop was more than 10% in only two datasets (D6 and D8), where KNN can exploit distance-based discrimination without being more sensitive to noise and feature correlations. Overall, the findings show that the chosen feature subsets are both robust and adaptable across different learners.

Table 4 *Average accuracy by other classifiers evaluated on the proposed HGWHHO_CBP-KNN-selected subset*

Dataset	HBGWHHO_CBP-KNN	HBGWHHO_CBP-KNN-selected subset		
		SVM	ANN	DT
D1	**0.9782**	0.9586	0.9598	0.9598
D2	0.9567	0.9440	**0.9690**	**0.9690**
D3	**0.7720**	0.7600	0.7617	0.7595
D4	**0.9585**	0.9352	0.9317	0.8993
D5	**0.9883**	0.9814	0.9870	0.9861
D6	**0.9156**	0.8317	0.8417	0.7683
D7	**1.00**	**1.00**	**1.00**	**1.00**
D8	**0.9405**	0.8500	0.8631	0.7310
D9	0.8448	0.8422	**0.8945**	0.8495
D10	**0.9758**	0.9475	0.9483	0.9483
D11	**0.9905**	0.9500	0.9619	0.9548

Comparison with Existing FS metaheuristics: The efficacy of the suggested algorithm HBGWHHO_CBP is validated against related metaheuristics in the FS literature including BGWO [15], BWOAHHO [9], and GWOCSA [17]. Compared to the other metaheuristics, HBGWHHO_CBP achieves the best performance—in terms of classification accuracy, selected feature subset and mean fitness—in most of the datasets, as shown in Table 5. HBGWHHO_CBP demonstrates enhanced performance by choosing fewer features while keeping its effective classification across large, medium, and small datasets. This illustrates the HBGWHHO_CBP 's capability to handle optimization objectives and can be considered a candidate for FS with a minimized quantity and good classification accuracy. In comparison to the state-of-the-art methods, the suggested HBGWHHO_CBP algorithm exhibited superior performance. This is mainly owing to the robust exploration mechanisms of HHO, limited set of parameters in GWO and the combination of their strengths and the adaptive technique CBP.

Table 5 *Comparison of the results of the proposed HBGWHHO_CBP with other FS metaheuristics*

Dataset	Evaluation metrics	HBGWHHO_CBP	BGWO	GWOCSA	BWOAHHO
D1	Accuracy	**0.98**	0.96	0.96	**0.98**
	Selected feature	4.6	**4.0**	5	5.3
	Mean fitness	**0.02**	0.04	0.04	**0.02**
D2	Accuracy	0.96	0.96	**0.99**	0.85
	Selected feature	**5.9**	6.1	6.4	8.7
	Mean fitness	0.04	0.04	**0.01**	0.15
D3	Accuracy	**0.77**	0.76	0.75	0.76
	Selected feature	**3.7**	**3.7**	4.6	5
	Mean fitness	**0.22**	0.24	0.25	0.23
D4	Accuracy	**0.96**	0.90	0.91	0.93
	Selected feature	9.1	15.2	13	**7.7**
	Mean fitness	**0.04**	0.10	0.09	0.07
D5	Accuracy	**0.99**	0.97	0.95	0.97
	Selected feature	21.5	**11.9**	18.6	30.2
	Mean fitness	**0.02**	0.03	0.05	0.04
D6	Accuracy	**0.92**	0.85	0.87	0.89
	Selected feature	**6.7**	7.8	8	9.2
	Mean fitness	**0.08**	0.15	0.13	0.11
D7	Accuracy	**1**	0.97	**1**	0.94
	Selected feature	**6.2**	7.2	6.4	9.9
	Mean fitness	**0.00**	0.03	0.01	0.09
D8	Accuracy	**0.94**	0.86	0.91	0.93
	Selected feature	**18.1**	32	29.6	25
	Mean fitness	**0.06**	0.15	0.10	0.07

J. of Soft Computing and Data Mining Vol. 6 No. 3 (2025) p. 307-322

Dataset	Evaluation metrics	HBGWHHO_CBP	BGWO	GWOCSA	BWOAHHO
D9	Accuracy	0.84	0.77	0.80	**0.85**
	Selected feature	7.2	**5**	**5**	6.8
	Mean fitness	0.17	0.23	0.20	**0.14**
D10	Accuracy	**0.98**	0.96	0.95	0.97
	Selected feature	**4.4**	4.8	4.6	5.1
	Mean fitness	**0.02**	0.05	0.05	**0.02**
D11	Accuracy	**0.99**	0.95	0.97	0.98
	Selected feature	6.2	8.5	**5.2**	10.2
	Mean fitness	**0.01**	0.05	0.03	0.02

Statistical Validation of Results: As a nonparametric statistical test, the Wilcoxon test is utilized to evaluate whether the suggested HBGWHHO_CBP algorithm significantly outperforms the baseline algorithms. A p-value of less than 0.05 indicates that the null hypothesis is rejected. In other words, there exists a statistically significant difference in performance. In contrast, if the p-value is greater than 0.05, we can then conclude that the observed differences between the compared values fail to demonstrate statistical significance. The analysis procedures in this study are adopted from [7]. A rank is independently assigned according to the performance of the algorithm. This performance is measured by the classification accuracy and fitness across the 11 datasets. The Findings of the Wilcoxon statistical test at significance level 0.05 are presented in Table 6, where the p-values for all comparable algorithms are below 0.05. These findings indeed demonstrate a statistically significant difference between suggested HBGWHHO_CBP and the baseline existing algorithms and due to the better trade-off between exploration and exploitation procedures when guided by the knowledge gained throughout optimization process.

Table 6 *Wilcoxon signed-rank tests results across all datasets*

HBGWHHO_CBP vs	Fitness p-value	Accuracy p-value
HBGWOHHO	5.81E-03	3.91E-03
BGWO	2.53E-03	9.80E-04
GWOCSA	4.88E-03	6.84E-03
BWOAHHO	2.18E-02	6.84E-03

The proposed HBGWHHO_CBP can be employed in a diversity of real-world application domains such as in the healthcare industry where it can be used to identify the most informative biomarkers that can be useful in the forecasting of various illnesses. For example, in the forecasting of heart disease, diabetes, or cancer, choosing the most informative biomarkers or genetic features can enhance the effectiveness of the diagnosis and support more reliable prognosis prediction.

7. Conclusions

The proposed HBGWHHO_CBP has enhanced the performance of HBGWOHHO for FS in classification. The enhancement was achieved using an adaptive technique based on feedback from the optimization process that was formulated to adjust the balancing parameter value for exploration and exploitation. This technique facilitates an effective investigation of the solution space for identifying promising regions and refine the search within those regions to determine the optimal solution. 11 UCI benchmark datasets were employed to assess the efficiency and efficacy of the suggested algorithm. The efficacy of the proposed HBGWHHO_CBP as wrapper FS method is compared with native HBGWOHHO and other metaheuristics, including BGWO, BWOAHHO, and GWOCSA. The findings demonstrated that the performance of the HBGWOHHO algorithm was successfully enhanced by the proposed method. The proposed HBGWHHO_CBP outperformed other algorithms on most datasets with respect to the classification accuracy, mean fitness, and selected feature size. The superior outcomes of the proposed HBGWHHO_CBP indicate its capability to increase the population diversity, regulate the exploration-exploitation trade-off and converge to a near-optimal solution more efficiently than the other algorithms.

For future work, HBGWHHO_CBP can be employed for real-world problems and can also be used with other popular classifiers to assess whether performance stays the same or changes. Another possible future work is to examine the efficacy of HBGWHHO-CBP when utilized on various high-dimensional datasets.

Acknowledgement

I would like to thank Taiz University for its financial support during my studies at Universiti Utara Malaysia. I also thank the reviewers and the editorial team for their valuable comments and constructive feedback, which greatly improved the quality of this paper.

Conflict of Interest

Authors declare that there is no conflict of interest regarding the publication of the paper.

Declaration of AI Use in Manuscript Preparation

In preparing this manuscript, the authors used QuillBot (Premium) and Grammarly to assist with grammar checking and improve the quality of the writing while preserving the original meaning of the content. All content generated was reviewed and verified by the authors, who take full responsibility for the final submission.

Author Contribution

Study conception and design, analysis and interpretation of results, draft manuscript preparation: *Manal Othman;* ***writing review, editing, and supervision:*** *Ku Ruhana Ku-Mahamud. All authors reviewed the results and approved the final version of the manuscript.*

References

[1] Alhussan, A. A., Abdelhamid, A. A., Towfek, S. K., Ibrahim, A., Eid, M. M., Khafaga, D. S., & Saraya, M. S. (2023) Classification of diabetes using feature selection and hybrid Al-Biruni Earth Radius and Dipper Throated Optimization, Diagnostics, 13(12), 2038, https://doi.org/10.3390/diagnostics13122038

[2] Khafaga, D. S., Alhussan, A. A., El-Kenawy, E. S. M., Takieldeen, A. E., Hassan, T. M., Hegazy, E. A., & Abdelhamid, A. A. (2022) Meta-heuristics for feature selection and classification in diagnostic breast cancer, Computers, Materials & Continua, 73(1), 749–765, https://doi.org/10.32604/cmc.2022.029605

[3] Safar, N. Z. M., Mujilahwati, S., Halif, K. M. N. K., & Ghazalli, N. (2024) Optimizing sentiment analysis of Indonesian texts: Enhancing deep learning models with genetic algorithm-based feature selection, Journal of Soft Computing and Data Mining, 5(2), 208–222, https://doi.org/10.30880/jscdm.2024.05.02.016

[4] Tama, B.A., Im, S. & Lee, S. (2020) Improving an intelligent detection system for coronary heart disease using a two-tier classifier ensemble, BioMed Research International, 2020(1), 9816142, https://doi.org/10.1155/2020/9816142

[5] Allam, M. & Nandhini, M. (2022) Optimal feature selection using binary teaching learning based optimization algorithm, Journal of King Saud University-Computer and Information Sciences, 34(2), 329-341, https://doi.org/10.1016/j.jksuci.2018.12.001

[6] Thawkar, S. (2021) A hybrid model using teaching–learning-based optimization and Salp swarm algorithm for feature selection and classification in digital mammography, Journal of Ambient Intelligence and Humanized Computing, 12(9), 8793-8808, https://doi.org/10.1007/s12652-020-02662-z

[7] Bezdan, T., Zivkovic, M., Bacanin, N., Chhabra, A. & Suresh, M. (2022) Feature selection by hybrid brain storm optimization algorithm for COVID-19 classification, Journal of Computational Biology, 29(6), 515-529, https://doi.org/10.1089/cmb.2021.0256

[8] Phogat, M. & Kumar, D. (2023) A Hybrid Metaheuristics based technique for Mutation Based Disease Classification, International journal of electrical and computer engineering systems, 14(6), 635-646, https://doi.org/10.32985/ijeces.14.6.3

[9] Alwajih, R., Abdulkadir, S. J., Al Hussian, H., Aziz, N., Al-Tashi, Q., Mirjalili, S. & Alqushaibi, A. (2022) Hybrid binary whale with harris hawks for feature selection, Neural Computing and Applications, 34(21) , 19377-19395, https://doi.org/10.1007/s00521-022-07522-9

[10] Al-Wajih, R., Abdulkadir, S.J., Aziz, N., Al-Tashi, Q. & Talpur, N. (2021) Hybrid binary Grey Wolf with Harris Hawks optimizer for feature selection, IEEE Access, 9, 31662-31677, https://doi.org/10.1109/ACCESS.2021.3060096

[11] Ewees, A. A., Al-Qaness, M. A., Abualigah, L., Oliva, D., Algamal, Z. Y., Anter, A. M. & Abd Elaziz, M. (2021) Boosting arithmetic optimization algorithm with genetic algorithm operators for feature selection: case study on Cox proportional hazards model, Mathematics, 9(18), 2321, https://doi.org/10.3390/math9182321

[12] Ye, Z., Huang, R., Zhou, W., Wang, M., Cai, T., He, Q. & Zhang, Y. (2024) Hybrid rice optimization algorithm inspired Grey Wolf optimizer for high-dimensional feature selection, Scientific Reports, 14(1), 30741, https://doi.org/10.1038/s41598-024-80648-z

[13] Aly, M. & Alotaibi, A.S. (2025) Hybrid Butterfly–Grey Wolf Optimization (HB-GWO): A novel metaheuristic approach for feature selection in high-dimensional data, Statistics, Optimization & Information Computing, 13(6), 2575-2600. https://doi.org/10.19139/soic-2310-5070-2617

[14] Khafaga, D. S., El-Kenawy, E. S. M., Karim, F. K., Abotaleb, M., Ibrahim, A., Abdelhamid, A. A. & Elsheweikh, D. L. (2023) Hybrid Dipper Throated and Grey Wolf Optimization for feature selection applied to life benchmark datasets, Computers, Materials & Continua, 74(2), 4531-4545, https://doi.org/10.32604/cmc.2023.033042

[15] Too, J., Abdullah, A. R., Saad, N. M., Ali, N. M. & Tee, W. (2018) A new competitive binary Grey Wolf optimizer to solve the feature selection problem in EMG signals classification, Computers, 7(4), 58, https://doi.org/10.3390/computers7040058

[16] El-Kenawy, E. S. M., Eid, M. M., Saber, M. & Ibrahim, A. (2020) MbGWO-SFS: Modified binary Grey Wolf optimizer based on stochastic fractal search for feature selection, IEEE Access, 8, 107635-107649, https://doi.org/10.1109/ACCESS.2020.3001151.

[17] Arora, S., Singh, H., Sharma, M., Sharma, S. & Anand, P. (2019) A new hybrid algorithm based on Grey Wolf optimization and Crow Search algorithm for unconstrained function optimization and feature selection, IEEE Access, 7, 26343-26361, https://doi.org/10.1109/ACCESS.2019.2897325.

[18] Teng, Z. J., Lv, J. L. & Guo, L. W. (2019) An improved hybrid Grey Wolf optimization algorithm, Soft Computing, 23(15), 6617-6631, https://doi.org/10.1007/s00500-018-3310-y

[19] Shaikh, M. S., Lin, H., Zheng, G., Wang, C. & Dong, X. (2024) Innovative hybrid grey wolf–particle swarm optimization for calculating transmission line parameter, Heliyon, 10(19), e38555. https://doi.org/10.1016/j.heliyon.2024.e38555

[20] Ahmad, I., Qayum, F., Rahman, S. U. & Srivastava, G. (2024) Using improved hybrid Grey Wolf Algorithm Based on Artificial Bee Colony Algorithm Onlooker and Scout Bee Operators for Solving Optimization Problems, International Journal of Computational Intelligence Systems, 17(1), 111, https://doi.org/10.1007/s44196-024-00497-6

[21] Heidari, A. A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M. & Chen, H. (2019) Harris hawks optimization: Algorithm and applications, Future Generation Computer Systems, 97, 849-872, https://doi.org/10.1016/j.future.2019.02.028

[22] Ouyang, C., Liao, C., Zhu, D., Zheng, Y., Zhou, C. & Li, T. (2024) Integrated improved Harris hawks optimization for global and engineering optimization, Scientific Reports, 14(1), 7445, https://doi.org/10.1038/s41598-024-58029-3

[23] Wang, X., Dong, X., Zhang, Y. & Chen, H. (2023) Crisscross Harris hawks optimizer for global tasks and feature selection, Journal of Bionic Engineering, 20(3), 1153-1174, https://doi.org/10.1007/s42235-022-00298-7

[24] Yasear, S. A. & Ku-Mahamud, K. R. (2021) Fine-Tuning the Ant Colony System Algorithm Through Harris's Hawk Optimizer for Travelling Salesman Problem, International Journal of Intelligent Engineering and Systems, 14(4), 136-145, https://doi.org/10.22266/ijies2021.0831.13

[25] Al-Tashi, Q., Shami, T. M., Abdulkadir, S. J., Akhir, E. A. P., Alwadain, A., Alhussain, H. & Mirjalili, S. (2023) Enhanced Multi-Objective Grey Wolf Optimizer with Lévy Flight and Mutation Operators for Feature Selection, Computer Systems Science and Engineering, 47(2), 1937-1966, http://dx.doi.org/10.32604/csse.2023.039788

[26] Hashim, F. A., Houssein, E. H., Mostafa, R. R., Hussien, A. G. & Helmy, F. (2023) An efficient adaptive-mutated Coati optimization algorithm for feature selection and global optimization, Alexandria Engineering Journal, 85, 29-48, https://doi.org/10.1016/j.aej.2023.11.004

[27] Pham, T. H. & Raahemi, B. (2023) Bio-inspired feature selection algorithms with their applications: a systematic literature review, IEEE Access, 11, 43733–43758, https://doi.org/10.1109/ACCESS.2023.3272556

[28] Piri, J., Mohapatra, P., Dey, R., Acharya, B., Gerogiannis, V. C. & Kanavos, A. (2023) Literature Review on Hybrid Evolutionary Approaches for Feature Selection, Algorithms, 16(3), 167, https://doi.org/10.3390/a16030167

[29] Gebhardt, C. G., Lange, S. & Steinbach, M. C. (2024) Formulating and heuristic solving of contact problems in hybrid data-driven computational mechanics, Communications in Nonlinear Science and Numerical Simulation, 134, 108031, https://doi.org/10.1016/j.cnsns.2024.108031

[30] Agrawal, A. & Tripathi, S. (2021) Particle swarm optimization with adaptive inertia weight based on cumulative binomial probability, Evolutionary Intelligence, 14(2), 305-313, https://doi.org/10.1007/s12065-018-0188-7

[31] Zhao, Z., Yu, H., Guo, H. & Chen, H. (2024) Multi-strategy augmented Harris Hawks optimization for feature selection, Journal of Computational Design and Engineering, 11(3), 111-136, https://doi.org/10.1093/jcde/qwae030

[32] Morales-Castañeda, B., Zaldivar, D., Cuevas, E., Fausto, F. & Rodríguez, A. (2020) A better balance in metaheuristic algorithms: Does it exist?, Swarm and Evolutionary Computation, 54, 100671, https://doi.org/10.1016/j.swevo.2020.100671

[33] Othman, M. M. & Ku-Mahamud, K. R. (2024) Hybrid Metaheuristic Algorithms for Feature Selection in Classification: A Systematic Literature Review, Research Square, https://doi.org/10.21203/rs.3.rs-4286826/v1

[34] Meidani, K., Hemmasian, A., Mirjalili, S. & Barati Farimani, A. (2022) Adaptive grey wolf optimizer, Neural Computing and Applications, 34(10), 7711-7731, https://doi.org/10.1007/s00521-021-06885-9