

## Research Article

# High-Dimensional Feature Selection Based on Improved Binary Ant Colony Optimization Combined with Hybrid Rice Optimization Algorithm

A. Zhiwei Ye,<sup>1</sup> B. Ruihan Li <sup>1</sup>, C. Wen Zhou,<sup>1</sup> D. Mingwei Wang,<sup>1</sup> E. Mengqing Mei,<sup>1</sup>  
F. Zhe Shu,<sup>1</sup> and G. Jun Shen <sup>2</sup>

<sup>1</sup>School of Computer Science, Hubei University of Technology, 430068 Wuhan, China

<sup>2</sup>School of Computing and Information Technology, University of Wollongong, Wollongong, Australia

Correspondence should be addressed to B. Ruihan Li; 102111083@hbut.edu.cn

Received 28 February 2023; Revised 3 June 2023; Accepted 26 June 2023; Published 10 July 2023

Academic Editor: Said El Kafhali

Copyright © 2023 A. Zhiwei Ye et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the realm of high-dimensional data analysis, numerous fields stand to benefit from its applications, including the biological and medical sectors that are crucial for computer-aided disease diagnosis and prediction systems. However, the presence of a significant number of redundant or irrelevant features can adversely affect system accuracy and real-time diagnosis efficiency. To mitigate this issue, this paper proposes two innovative wrapper feature selection (FS) methods that integrate the ant colony optimization (ACO) algorithm and hybrid rice optimization (HRO). HRO is a recently developed metaheuristic that mimics the breeding process of the three-line hybrid rice, which is yet to be thoroughly explored in the context of solving high-dimensional FS problems. In the first hybridization, ACO is embedded as an evolutionary operator within HRO and updated alternately with it. In the second form of hybridization, two subpopulations evolve independently, sharing the local search results to assist individual updating. In the initial stage preceding hybridization, a problem-oriented heuristic factor assignment strategy based on the importance of the knee point feature is introduced to enhance the global search capability of ACO in identifying the smallest and most representative features. The performance of the proposed algorithms is evaluated on fourteen high-dimensional biomedical datasets and compared with other recently advanced FS methods. Experimental results suggest that the proposed methods are efficient and computationally robust, exhibiting superior performance compared to the other algorithms involved in this study.

## 1. Introduction

High-dimensional data are increasingly prevalent in diverse fields such as medical diagnosis and genomics, presenting significant challenges to the construction of intelligent systems. Data from medical imaging generally comprise a variety of features that delineate patient conditions. Similarly, microarray data from high-throughput gene expression experiments concurrently measure the expression levels of thousands of genes [1]. Such high dimensionality creates difficulties in terms of storage, computation, and analysis [2]. Specifically, the abundance of features may lead to the “curse of dimensionality,” where data sparsity occurs and the distance between sample points loses significance.

This situation, in turn, can lead to the overfitting of data mining models and increased computation time. Therefore, selecting representative features that positively impact intelligent systems becomes critical. Feature selection (FS) is employed to minimize the feature space, thereby not only conserving storage space but also facilitating information discovery and mitigating the potential for model overfitting.

FS techniques directly identify the optimal feature subset (OFS) from the original feature space, preserving the interpretability of the chosen features and enhancing the comprehension of underlying data patterns and relationships. The broad applicability of FS spans areas such as text classification [3], image processing [4, 5], finance [6, 7], and other fields. As an NP-hard problem in nature, the search

space of FS grows exponentially with the number of features, making traditional search methods ineffective in finding the OFS within polynomial time. Existing FS methods primarily fall into three categories: filter, wrapper, and embedded, each distinguished by its selection process. The filter method independently evaluates the relevance and redundancy of each feature using statistical or ranking measures, irrespective of the particular classifier employed in subsequent steps. The wrapper method integrates a specific learning algorithm within the FS process, treating FS as a search problem and seeking the OFS by assessing the performance of each feature subset on the classifier. Although wrappers entail a higher computational cost compared to highly efficient filter methods, they can handle nonlinear feature interactions and dependencies. Furthermore, wrappers can determine the OFS for a specific learning algorithm, potentially enhancing predictive performance [8]. The embedded method incorporates the FS process within the training process of the learning algorithm. The superior performance of the wrapper method has attracted increasing research interest, marking it as a popular FS technique. To tackle the computational intensity of wrapper methods, researchers are investigating the use of metaheuristics in FS to boost search efficiency and OFS quality.

Metaheuristics have surfaced as promising alternatives for FS in high-dimensional data, efficiently balancing exploration and exploitation in the search space to converge towards near-optimal solutions. Various high-performance metaheuristics have been successfully employed across diverse FS tasks [9, 10]. For instance, Jia et al. [11] enhanced the slime mould algorithm (SMA) [12] for FS, wherein the FS process and the parameter optimization of SVM occur concurrently. To improve population diversity, a composite mutation strategy was introduced, and a trial-based restart strategy was designed to circumvent the local optima. Qu et al. [13] proposed a novel gene selection method based on Harris hawk optimization (HHO). The F-score is initially employed for preliminary feature filtering to condense the feature space, followed by a variable neighborhood learning strategy for balancing the exploration and exploitation of HHO. Ghosh et al. [14] studied eight different transfer functions from two series (S-shaped and V-shaped) and suggested an improved binary manta ray foraging optimization (MRFO) [15] for FS. Ewees et al. [16] modified the seagull optimization algorithm (SOA) using the Levy flight mechanism to overcome its linear search deficiency in the search space, expanding the search area and enhancing the capability of individuals to escape from local optima. To select the optimal gene combination effectively in microarray data, Pashaei [17] utilized mRMR to filter the top  $m$  promising genes in the initial stage to reduce the feature space and then applied the Aquila Optimizer (AO) with a mutation mechanism and TVMS transfer function to search the OFS. Awadallah et al. [18] developed a new enhanced binary rat swarm optimizer (RSO), in which the local historical optimal strategy of PSO was introduced to enhance individual exploitation. Furthermore, three crossover operators, namely, one-point crossover, two-

point crossover, and uniform crossover, are incorporated into the individual update process and randomly selected with equal probability. A multi-strategy integrated grey wolf optimizer (MSGWO) was presented for biological data classification [19]. It adopted the convergence factor concept to adjust the transition between exploration and exploitation explicitly. Additionally, multiple exploration and exploitation strategies were employed to boost the global search and local exploitation processes.

The no-free lunch (NFL) theorem suggests that no single algorithm can resolve all optimization problems due to the inability to strike a perfect balance between global search and local exploitation. Consequently, recent research has increasingly focused on hybrid metaheuristics to address the limitations of individual algorithms. For instance, Pashaei [2] introduced a hybrid FS method that integrates the dragonfly algorithm (DFA) and the black hole algorithm (BHA), in which the optimal solution derived from DFA serves as the initial solution for BHA. A hybrid FS method, merging the binary arithmetic optimization algorithm (BAOA) and simulated annealing (SA), was proposed in [20], utilizing SA as a local search operator to discover potential solutions near the optimal solution of BAOA. Stephan et al. [21] put forward a method combining the artificial bee colony (ABC) and whale optimization algorithm (WOA) to concurrently search the OFS of breast cancer data and optimize the artificial neural network parameters. 'In addition, there are several studies attempting to integrate the HHO with other metaheuristic algorithms.' We hope this revision is more clear and accurate [22–24]. Beyond metaheuristic hybrids, the combination of filters and wrappers has also been extensively researched, aiming to determine the OFS at a lower computational cost. Zhu et al. [25] combined the Fisher filtering method with the artificial immune optimizer for high-dimensional FS, introducing a lethal mutation mechanism and a Cauchy mutation operator with an adaptive adjustment factor to enhance population diversity. A new high-dimensional FS method in conjunction with mRMR was developed in [26]. This method improved the recently proposed COOT algorithm using a crossover operator and employed the hyperbolic tangent transfer function for continuous numerical binarization. The improved COOT was hybridized with SA and applied to FS of microarray data. While these studies have demonstrated the performance of metaheuristic-based FS methods, the risk of local optima entrapment persists due to the stochastic strategies inherent in metaheuristics that ensure the global search of the algorithm. Therefore, the development of a computationally efficient and robust high-dimensional FS method remains a crucial and ongoing research endeavor.

Hybrid rice optimization (HRO) algorithm, proposed by Ye et al. [27], is a novel population-based metaheuristic inspired by the real-world breeding process of three-line hybrid rice. The heterosis theory suggests that the first generation of hybrids manifests superior physical, reproductive, and behavioral characteristics compared to their parent generation. Echoing this concept, HRO has exhibited several desirable properties, including high search efficiency, fewer control parameters compared to other metaheuristics, and ease of implementation. Its high flexibility and reduced

parameter dependency have encouraged researchers to apply it to the 0-1 knapsack problem [28] as well as the FS problem [29].

While most metaheuristics necessitate conversion to binary form via a specific transfer function to indicate whether the feature is selected, this approach may be appropriate for all situations, as they were initially designed to manage continuous optimization problems. As a classic swarm intelligence algorithm, ant colony optimization (ACO) is more suitable for FS tasks, given its origin as a combinatorial optimization algorithm. ACO strikes a well-balanced ratio between exploration and exploitation by dynamically adjusting pheromone density. Numerous studies have demonstrated the robustness and adaptability of ACO in resolving FS problems. Wang et al. [30] introduced a novel approach for FS, namely, the probabilistic sequence-based graphical representation ACO, incorporating symmetric uncertainty (SU) into the algorithm. Paniri et al. [31] presented an innovative multi-label FS method based on ACO, which used both unsupervised and supervised heuristic functions to seek features with minimal redundancy and maximal correlation with class labels. Moreover, a recent study [32] put forth a semisupervised FS method based on ACO that employs a nonlinear heuristic function trained using temporal difference (TD) reinforcement learning instead of the traditional linear heuristic functions. Despite the standard ACO and HRO having demonstrated commendable optimization performance in low-dimensional optimization problems, their performance may encounter difficulty in achieving ideal performance when dealing with high-dimensional FS problems. This challenge motivates us to combine the strengths of both algorithms, aiming to create a more robust and efficient method for high-dimensional FS.

This paper proposes a two-stage hybrid technique for solving high-dimensional FS problems, leveraging the strengths of both ACO and HRO. Considering the limitations of standard ACO, this paper seeks to enhance it in the first stage prior to hybridization. A novel problem-oriented heuristic factor assignment strategy based on the importance of the knee point feature is designed to augment the search capability of ACO in high-dimensional FS. In the second stage, two hybrid models merging the improved binary ACO (IBACO) and HRO are presented and applied to high-dimension FS tasks. The hybridization manifests in two forms: low-level relay hybrid (LRH) and high-level teamwork hybrid (HTH) [33]. In the case of LRH, IBACO serves as an operator within HRO to guide the evolution of the maintainer line. This hybrid mode has been introduced due to the notable absence of suitable maintainer line update strategies within HRO, a factor of particular importance for high-dimensional FS tasks, while in HTH, subpopulations of HRO and IBACO evolve independently and share their local search results after each iteration. A comprehensive list of acronyms and symbol annotations used throughout this paper is provided in Table 1. The main contributions of the paper are summarized as follows:

TABLE 1: The meaning of the used symbols.

Acronym	Meaning
$F$	The feature space
$D$	Dimensionality of feature
OFS	The optimal feature subset
PD	Pheromone density
HF	Heuristic factor
$\tau_{i,j}(k)$	The pheromone density from bit $i$ to bit $j$ on the $k_{th}$ path
$g_{best}$	The global optimal solution
$p_{best}$	The best solution obtained during the current iteration
$V_{max}^d$	The maximum value in the $d_{th}$ dimension
$V_{min}^d$	The minimum value in the $d_{th}$ dimension
$\rho$	The evaporation factor of pheromone density
$W$	The descending order of feature correlation
$\lambda$	The weight of error rate
$\mu$	The weight of selection rate
SC	The number of selfing of the restorer line in HRO

- (1) Two unique hybridizations of IBACO and HRO, namely, R-IBACO and C-IBACO, are presented to effectively leverage the advantages of the two algorithms and obtain a more promising solution.
- (2) A new problem-oriented assignment strategy for the heuristic factor (HF) based on the feature correlation of the knee point feature is proposed.
- (3) The proposed methods are applied in high-dimensional FS tasks and compared with existing state-of-the-art metaheuristic-based FS methods.
- (4) The performance of the proposed methods is evaluated from multiple perspectives, and the results are subjected to the Wilcoxon signed-rank test and Friedman test.

The remainder of this paper is structured as follows. Section 2 presents an overview of the related work. Section 3 discusses the mathematical model of FS and provides some theoretical background on BACO and HRO. The specifics of the proposed methods are elucidated in Section 4. Section 5 outlines the experiments conducted and analyzes the obtained results. Finally, Section 6 draws conclusions from this study and indicates potential areas for future research.

## 2. Related Work

*2.1. Metaheuristic-Based FS.* In recent years, the application of metaheuristics for FS problems has seen a significant rise. Notably, many of the proposed wrapper FS methods draw upon GWO due to its remarkable performance in solving continuous optimization problems. In the work of Hu et al. [34], an enhanced variant of the binary GWO was introduced, incorporating a novel strategy for updating the parameter governing exploration and exploitation, along with five transfer functions for mapping continuous values to their binary counterparts, thereby enhancing the quality of candidate solutions. Likewise, Abdel-Basset et al. [35] proposed three distinct binary GWO variants, each utilizing different transfer functions. In addition to GWO-based

research, PSO has been a focus in prior studies. For example, Song et al. [36] proposed an improved integer PSO with a fast correlation-guided feature clustering strategy that markedly reduces the computation cost for FS. The gaining-sharing knowledge (GSK) optimization algorithm is a newly devised metaheuristic that draws its inspiration from the human process of acquiring and sharing knowledge [37]. Its robustness and convergence properties have proven to be highly effective in solving continuous optimization problems. Some enhanced versions of GSK have been successfully employed for FS tasks. As a continuous optimization algorithm, Agrawal et al. [38] implemented eight S-shaped and V-shaped transfer functions to map the individual codes of GSK into the binary search space. In addition, a dynamic population reduction strategy is introduced to facilitate the adjustment of population size during the pursuit of OFS. Another improved GSK integrating chaotic map strategies was proposed in [39], which utilized a probability estimation operator to represent its binary variant. Hanbay [40] proposed an innovative standard error-based ABC for FS, incorporating new solution search mechanisms based on standard error into the original ABC algorithm and utilizing the Shannon condition entropy value for FS. Moreover, an advanced salp swarm algorithm (SSA) was also suggested in [41], utilizing opposition-based learning to augment population diversity and implementing a novel local search mechanism to circumvent the local optima.

*2.2. ACO-Based FS.* ACO has displayed exceptional performance in tackling a wide range of discrete optimization challenges. For instance, Owuor et al. [42] performed a detailed evaluation of three population-based optimization techniques in the context of mining gradual patterns. The findings suggested that ACO surpassed the other two techniques and traditional counterparts. In another study, Zhang et al. [43] introduced a knowledge-based local search for multi-population ant colony systems to address the multi-objective supply chain configuration problem in supply chain management. By leveraging two independent ant colonies, the cost of goods sold and the lead time were concurrently minimized, facilitating an efficient search of the target space. Moreover, Zhao et al. [44] enhanced ACO using horizontal and vertical crossover search and applied it in the field of image segmentation. The superior combinatorial optimization capability of ACO has encouraged researchers to extend its application to various types of FS tasks, some of which have been successfully implemented [45].

The heuristic factor (HF) and the pheromone density (PD) are two crucial parameters in ACO that can significantly influence its performance. However, some FS methods based on ACO have not appropriately set the HF, even generating it randomly. Appropriately configuring the HF and PD parameters can enhance the global and local search capabilities of ACO. Recently, the impact of these parameters on the performance of ACO has been explored. Manbari et al. [46] introduced a circular graph-based ACO algorithm for FS, wherein each feature is interconnected

with the subsequent one through a pair of select/deselect edges. The heuristic information pertaining to each feature is determined based on its corresponding term variance. Ghosh et al. [47] proposed an ACO-based wrapper-filter FS method that employs a filter to assess feature subsets instead of a wrapper, significantly reducing the computational complexity. The HF is estimated by calculating the similarity between the last added feature and the feature whose addition probability is being calculated. Additionally, a novel multi-label FS method based on ACO was proposed in [48], which employed a heuristic learning approach instead of a static heuristic function. The heuristic function of ACO is learned from experiences directly using TD reinforcement learning, which significantly improves the quality of the selected feature subset. Moreover, Hashemi et al. [49] proposed an ensemble FS approach based on ACO that utilized multiple heuristic information determined by a multi-criteria decision-making procedure. Such distinctive heuristic information supplies further insights about subsequent nodes. Experimental results demonstrated that the proposed method significantly outperforms other methods across various evaluation indicators.

Despite the proven effectiveness of ACO in solving FS problems, it still encounters issues such as limited global search capability and slow convergence rates, particularly when handling high-dimensional problems. To mitigate these limitations, the concept of hybridizing with complementary metaheuristics has emerged as a promising solution. As detailed by Talbi [33], there exists a two-tiered hybridization system encompassing low-level and high-level hybridization, each comprising two distinct mechanisms of hybridization. Building on this concept, Wan et al. [50] proposed the VMBACO, a hybrid FS method that blends a modified binary ACO algorithm with a genetic algorithm (GA). This hybrid method employs the solution derived from GA as the assignment of the HF in BACO, resulting in substantial improvements in both the quality of feature subsets and classifier accuracy. In a similar vein, Li et al. [51] introduced a hybrid FS model that combines ACO with the antlion optimizer, which includes a novel mutation operator to enhance exploration capability. Furthermore, Ma et al. [52] suggested a two-stage hybrid ACO algorithm for high-dimensional FS. This method employs an interval strategy to determine the size of the OFS and integrates a hybrid model that harnesses the inherent relevance attributes of features and classification performance to direct the OFS search. This advanced hybrid ACO assigns HF to a feature by calculating its correlation with the chosen feature subset after softmax normalization. Despite the improved performance of these hybrid algorithms, they remain susceptible to the local optima and often exhibit relatively high computational complexity.

The heterosis-inspired HRO has demonstrated high search efficiency and robust global search capabilities, as illustrated by its successful application to the band selection problem [29]. Moreover, the diverse set of operators employed by HRO facilitates the effective maintenance of population diversity in high-dimensional problem spaces. Considering these benefits of HRO, it is hypothesized that

integrating the improved, problem-oriented ACO with HRO may yield a more efficient and robust search for OFS in high-dimensional FS tasks.

### 3. Preliminaries

**3.1. The Binary Coded Ant Colony Optimization Algorithm.** Binary ant colony optimization (BACO) is a metaheuristic inspired by ant behavior, and it is designed to find the optimal solutions for binary optimization problems. BACO is particularly useful in optimizing scenarios such as FS and intrusion detection [53]. BACO commences with a randomly generated set of candidate solutions, depicted as binary strings. Throughout each iteration, the algorithm adjusts the probability of including each feature, drawing from the information acquired from previously established solutions. In BACO, the PD serves as a directive for subsequent ants to decide whether a feature should be incorporated into their solution. Figure 1 presents a binary directed acyclic graph, symbolizing the potential paths ants can follow to find a solution. Figure 2 elaborates on the path of an ant to construct a corresponding solution to a feature subset. In this example, five features ( $F_i, i = 1, 2, \dots, 5$ ) are considered, and the bold black solid line represents the path constructed by an ant. The selected feature subset is  $\{F_2, F_3, F_5\}$ , while features  $F_1$  and  $F_4$  are deselected.

The node selection for the next bit is determined by the state transition probability, a function derived from both the PD and HF. The mathematical expression for the state transition probability is given by equations (1) and (2),

$$p_{i,j}^k(1) = \frac{\tau_{i,j}^\alpha(1) \cdot \eta_{i,j}^\beta(1)}{\tau_{i,j}^\alpha(0) \cdot \eta_{i,j}^\beta(0) + \tau_{i,j}^\alpha(1) \cdot \eta_{i,j}^\beta(1)}, \quad (1)$$

$$p_{i,j}^k(0) = 1 - p_{i,j}^k(1), \quad (2)$$

where  $p_{i,j}^k(1)$  denotes the probability that the  $k_{th}$  ant selects path 1 from bit  $i$  to bit  $j$  and  $\tau_{i,j}(1)$  and  $\eta_{i,j}(1)$  are the corresponding PD and HF values for this path. Similarly,  $p_{i,j}^k(0)$ ,  $\tau_{i,j}(0)$ , and  $\eta_{i,j}(0)$  represent the probability, PD, and HF values for the case where the ant selects path 0 from bit  $i$  to bit  $j$ . The parameters  $\alpha$  and  $\beta$  control the relative importance of the PD and HF, respectively. Initially, the PD for each path is set to be equivalent and then decreases as ants make their selections. The update rule for the PD from bit  $i$  to bit  $j$  is described by equations (3) and (4).

$$\tau_{i,j}(1)(t+1) = (1 - \rho) \cdot \tau_{i,j}(1)(t) + \Delta\tau, \quad (3)$$

$$\tau_{i,j}(0)(t+1) = (1 - \rho) \cdot \tau_{i,j}(0)(t). \quad (4)$$

The evaporation factor, denoted as  $\rho$ , signifies the rate of pheromone decay on unselected paths. Conversely, the chosen path experiences an increase in PD as described by  $\Delta\tau$ , which is defined by the fitness value of the optimal candidate solution,  $f_{best}$ , procured in the current iteration. BACO also maintains a record of the optimal solution found to date, denoted as  $g_{best}$ . Additionally, each iteration records its own optimal solution,  $p_{best}$ , with its fitness value being

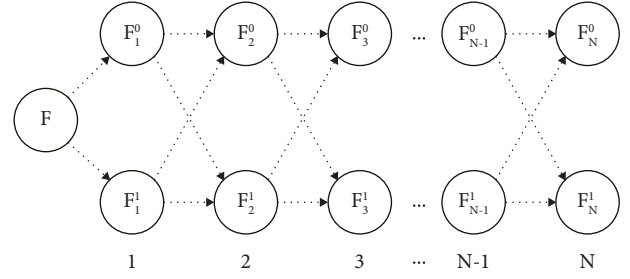


FIGURE 1: DAG of path constructed by BACO.

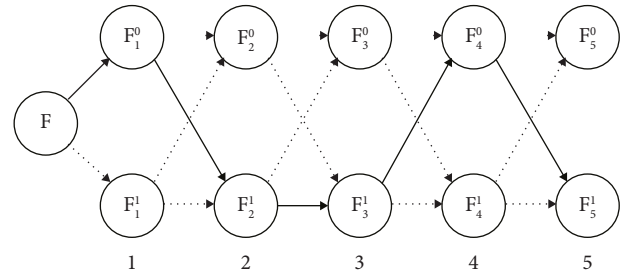


FIGURE 2: The construction process of a feature subset  $\{F_2, F_3, F_5\}$ .

compared against that of  $g_{best}$ . If the fitness value of  $p_{best}$  surpasses that of  $g_{best}$ ,  $g_{best}$  is updated with  $p_{best}$  at iteration  $t$ . Upon the completion of all iterations,  $g_{best}$  serves as the global optimal solution.

BACO is limited by the absence of a suitable HF for the FS task, especially when the dimension is high. This makes BACO tend to converge to a local optimal solution rather than a global optimal solution. To fully exploit the performance of BACO in FS, it is essential to introduce a suitable HF assignment strategy specifically for the FS task and to improve the relevant update strategy. The improved strategies will be presented in detail in Section 4.2.

**3.2. Hybrid Rice Optimization Algorithm.** Hybrid rice optimization (HRO), proposed by Ye et al. [27], is an innovative metaheuristic that draws upon the advantages of heterosis, offering superior search efficiency and robust global search capabilities. HRO divides the population into three lines based on their fitness values. Let  $X = (X_1, X_2, \dots, X_n)$  denote the population sorted by fitness value, where  $n$  is the size of the population. The first subpopulation, representing the best fitness within the entire population, is referred to as the maintainer line and is denoted by  $X_1, X_2, \dots, X_p$  ( $p = \lfloor n/3 \rfloor$ ). The second subpopulation represented by  $X_{2p+1}, X_{2p+2}, \dots, X_n$  possesses the poorest fitness value and constitutes the sterile line, which requires hybridization with the maintainer line to enhance its fitness quality. The remaining subpopulation is the restorer line, denoted as  $X_{p+1}, X_{p+2}, \dots, X_{2p}$ . It aims to evolve into the maintainer line through a process referred to as *selfing*. The principle of HRO will be further elaborated in the following sections.

**3.2.1. Hybridization.** The hybridization process involves crossing the maintainer line and sterile line, which have the greatest difference in fitness value. This process aims to update the sterile line by substituting the original individual with a newly generated hybrid individual if it displays superior fitness. The new individual within the sterile line can be obtained using the following equation:

$$\begin{cases} X_{\text{new}(i)}^d(t) = r_1 \cdot X_s^d(t) + (1 - r_1) \cdot X_m^d(t), \\ m \in \{1, 2, \dots, p\}; i, s \in \{2p + 1, 2p + 2, n\}, \end{cases} \quad (5)$$

where  $X_{\text{new}(i)}^d(t)$  denotes the  $d_{\text{th}}$  gene of the  $i_{\text{th}}$  hybrid individual of the sterile line at iteration  $t_{\text{th}}$ .  $X_s^d(t)$  represents the  $d_{\text{th}}$  gene of a randomly selected individual from the sterile line, while  $X_m^d(t)$  is the  $d_{\text{th}}$  gene of the individual randomly selected from the maintainer line.  $r_1$  is a random number generated from  $[0, 1]$ .

**3.2.2. Selfing.** The selfing process is designed to update the restorer line, with the intention of steering the individual towards the global optimal solution. This behavior can be mathematically modeled using the following equation:

$$\begin{cases} X_{\text{new}(i)}^d(t) = r_2(X_{\text{best}}^d(t) - X_{r(j)}^d(t)) + X_{r(i)}^d(t), \\ i, j \in \{p + 1, p + 2, \dots, 2p\}, j \neq i, \end{cases} \quad (6)$$

where  $X_{\text{new}(i)}^d(t)$  represents the new gene produced by selfing between the  $i_{\text{th}}$  and  $j_{\text{th}}$  restorer ( $j \neq i$ ).  $X_{\text{best}}^d(t)$  denotes the  $d_{\text{th}}$  gene of the best individual found so far and  $X_{r(j)}^d(t)$  represents the  $d_{\text{th}}$  gene of the restorer randomly selected from the restorer line. The variable  $r_2$  is a random number generated within the range of  $[0, 1]$ .

After generating a new individual through hybridization and selfing, it is evaluated and compared to the original candidate solution. The substitution process, as defined in equation (7), replaces the old individual with the new one only if the fitness value of the new individual surpasses that of the old one.

$$X_i(t + 1) = \begin{cases} X_{\text{new}(i)}(t), & \text{if } f(X_{\text{new}(i)}(t)) > f(X_i(t)), \\ X_i(t), & \text{otherwise.} \end{cases} \quad (7)$$

**3.2.3. Renewal.** The selfing process in HRO involves tracking the number of iterations during which a restorer has not undergone updates using a parameter called SC (self-crossing). If a restorer's SC reaches the upper limit, denoted as  $t_{\text{max}}$  iterations without updates, a reset operation is performed on that individual. This reset behavior is mathematically encapsulated in the following equation:

$$X_{r(i)}^d(t + 1) = r_3(V_{\text{max}}^d - V_{\text{min}}^d) + X_{r(i)}^d(t) + V_{\text{min}}^d, \quad (8)$$

where  $X_{r(i)}^d(t)$  represents the  $d_{\text{th}}$  gene of the  $i_{\text{th}}$  restorer that has not been updated and  $V_{\text{max}}^d$  and  $V_{\text{min}}^d$  represent the maximum and minimum values of the  $d_{\text{th}}$  dimension, respectively.  $r_3$  is a random number drawn from  $[0, 1]$ . It is

worth noting that the solution obtained by HRO is continuous, which needs to be mapped into the solution space of FS through the application of the transfer function. In this study, the sigmoid function is employed as a binary map, which is defined by

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}, \quad (9)$$

$$X_i^d = \begin{cases} 1, & \text{if } \text{sigmoid}(X_i^d) > \text{rand}, \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

## 4. The Proposed Method

Although the effectiveness of traditional ACO-based methods in solving low-dimensional combinatorial optimization problems is well established, it is necessary to further improve the strategies employed by BACO to tackle high-dimensional FS problems. In pursuit of this aim, two hybrid models have been proposed, integrating IBACO and HRO to improve the efficiency of exploring the global optimal solution and to enhance the rate of convergence. Moreover, given the considerable impact of the HF assignment strategy on both the population update process and the overall performance, a novel problem-oriented HF assignment scheme is introduced, which is based on the significance of the knee point feature. This scheme takes into account the importance of the knee point feature to ensure that the population is updated effectively. Lastly, the objective function that needs to be optimized has been finessed, taking into account not only the classification accuracy but also the size of the feature subset.

**4.1. The Hybrid Models.** To fully harness the convergence rate and global search capability benefits of HRO, IBACO is melded with HRO in two unique hybrid models: the relay model (R-IBACO) and the collaborative model (C-IBACO). The details of these two hybrid models are discussed in the following.

**4.1.1. The Relay Model.** As described in Section 3.2, the maintainer line in HRO symbolizes the optimal candidate solutions discovered by the population. The quality of the maintainer line is crucial in determining the evolution of the population and influencing both the convergence rate and the final result. Moreover, the sterile line is updated via crossing with the maintainer line, underscoring the importance of enhancing the quality of the maintainer line in HRO. To this end, a LRH model is proposed, applying the update strategies of IBACO to the update process of the HRO maintainer line. Specifically, the  $k_{\text{th}}$  bit of the binary string corresponding to each maintainer candidate solution is evolved using the PD and HF of IBACO. Each individual of the maintainer line selects a path determined by the transition probability (equations (1) and (2)). This integration of the IBACO operator effectively enhances numerous dimensions in the maintainer line during the early

iterations, enabling R-IBACO to concentrate on global search in discrete spaces and quickly converge towards the vicinity of the global optimum. Furthermore, the best individual from the maintainer line is employed to update the PD. The detailed implementation of the relay model is outlined in Algorithm 1.

**4.1.2. The Collaborative Model.** In this section, a co-evolutionary hybrid model, referred to as C-IBACO, is proposed. This model maintains the efficiency of IBACO in tackling high-dimensional FS problems while integrating the effective optimization performance of HRO. C-IBACO consists of two subpopulations, with IBACO and HRO independently executing their respective update strategies in each iteration. Specifically, ants select a path based on the transition probabilities specified in equations (1) and (2), while the  $k_{th}$  gene of the  $i_{th}$  rice individual is updated using equations (5), (6), and (8). The HF of the  $0_{th}$  path assigns importance to the knee point to measure the potential of each feature to be selected. Features with a correlation greater than the knee point are more likely to be included in the feature subset while those with a low feature correlation not losing the opportunity to be selected.

To promote co-evolution between these two parallel algorithms, the search results from each subpopulation are shared after each iteration. The global optimal solution is determined by comparing the fitness values of the best candidate solutions from both subpopulations. If the best individual from HRO outperforms that of IBACO, the pheromone updated originally with the best candidate solution from the ant colony is substituted by the best individual from HRO. Conversely, if the best candidate solution from IBACO prevails, the worst solution of the maintainer line in HRO is supplanted by the best candidate solution of IBACO.

By incorporating these improvement strategies and information sharing mechanisms, C-IBACO can potentially identify a superior feature subset, selecting the most promising features effectively. The specifics of the C-IBACO procedure are detailed in Algorithm 2.

**4.2. The Improved Heuristic Factor.** The random assignment of the HF in BACO has been identified as a limitation in its capability to optimize high-dimensional FS problems. This limitation calls for a task-specific HF assignment to enhance performance. Nevertheless, traditional FS methods, such as those based solely on high-correlation features, may be arbitrary and do not consider the potential influence of combined features with low correlation on the final classification result. To mitigate this issue, this paper proposes the use of the weight of the knee point [54] as a threshold in the HF assignment process. This approach obviates the need for complex experimental verification and avoids significant loss of class label information.

Initially, the proposed methods employ the random forest (RF) algorithm to calculate the importance of each feature. All features are then sorted based on their RF results, and the knee point is defined as the feature with the

maximum distance from the straight-line projection connecting the features with the maximum and minimum importance. The detailed selection process for the knee point is described in Algorithm 3 and is graphically represented in Figure 3.

Equation (11) presents the HF assignment strategy in the proposed IBACO method, where  $I_k$  denotes the importance of the  $k_{th}$  feature and  $I_{knee\_point}$  represents the importance of the knee point feature.  $\eta_k(0)$  indicates that the HF of the  $0_{th}$  path adopts  $I_{knee\_point}$  while  $\eta_k(1)$  means that the HF of the  $1_{th}$  path adopts  $I_k$ . According to equations (1) and (2), this assignment strategy prioritizes selecting features with higher correlation while not disregarding less significant features.

$$\begin{cases} \eta_k(1) = I_k, \\ \eta_k(0) = I_{knee\_point} \end{cases} \quad (11)$$

**4.3. Time Complexity Analysis.** The time complexity of the proposed hybrid algorithms primarily hinges on four aspects: initialization, individual evaluation and update, population sorting, and pheromone update. Table 2 contrasts the time complexity of the hybrid algorithms with those of standalone algorithms at each stage. In this context,  $N$  symbolizes the population size, while  $N_1$  and  $N_2$  denote the subpopulation sizes of HRO and IBACO within C-IBACO, respectively.  $D$  signifies the dimension of  $F$ . The detailed analysis of time complexity reveals that, in comparison to ACO, R-IBACO increases the computational complexity merely at the population sorting phase. Conversely, in relation to HRO, it heightens the computational overhead solely in updating the maintainer line and PD. Owing to the cooperative updating of two subpopulations engaged in C-IBACO, it manifests a higher time complexity than the single algorithms. This increased complexity emerges during the population initialization, population sorting, and individual update process. The total time complexities of R-IBACO and C-IBACO stand at  $O(T \times (N \times (1 + D + \log N) + D))$  and  $O(T \times ((2/3N_1 + N_2 + 1) \times D + N_1 \times (1 + \log N_1) + N_2))$ , respectively, where  $T$  represents the maximum number of iterations. Overall, it can be suggested the hybrid algorithm significantly improves the overall performance of the model within an acceptable range of increased computational overhead.

**4.4. The Objective Function.** The objective function serves as the guiding principle in providing guidance for the update process and determining the optimization direction of the algorithms. In the context of FS, the objective is to identify the most informative feature subset that can enhance the performance of classifiers. The primary criterion for evaluating the efficacy of a classifier is its classification accuracy, which is defined in equation (12). However, since the objective function is designed to be minimized, the error rate of classification, which is the complement of accuracy (1-accuracy), is employed as a primary factor in the objective function. Moreover, it is desirable to maintain a minimal

**Input:** Dataset  $D$  and objective function Fitness  
**Output:** OFS and corresponding classification accuracy

- (1) Initialize population  $P$  and set the maximum number of iterations  $T$
- (2) Calculate feature importance and knee point feature using Algorithm 3
- (3) Set heuristic factor and initialize PD
- (4) **while**  $t < T$  **do**
- (5) Calculate the fitness of the population  $f = \text{Fitness}(P)$
- (6) Sort  $P$  in descending order and divide it into three lines:  $M$  (Maintainer),  $R$  (Restorer), and  $S$  (Sterile)
- (7) **for** Individual  $X_i$  in  $P$  **do**
- (8) **if**  $X_i$  in  $M$  **then**
- (9) Generate trial solution  $X_{\text{new}(i)}$  using equations (1) and (2)
- (10) **else if**  $X_i$  in  $S$  **then**
- (11) Generate trial solution  $X_{\text{new}(i)}$  using equation (5)
- (12) **else**
- (13) **if**  $SC < SC_{\text{max}}$  **then**
- (14) Generate trial solution  $X_{\text{new}(i)}$  using equation (6)
- (15) **else**
- (16) Generate trial solution  $X_{\text{new}(i)}$  using equation (8)
- (17) **end if**
- (18) **end if**
- (19) Compute  $f_{X_{\text{new}(i)}} = \text{Fitness}(X_{\text{new}(i)})$
- (20) Update  $X_i$  using equation (7)
- (21) **if**  $X_i$  is a binary vector with all bits set to 0 **then**
- (22) Reinitialize the binary vector corresponding to  $X_i$
- (23) **end if**
- (24) **end for**
- (25) Update  $g_{\text{best}}$ ,  $p_{\text{best}}$ , and PD
- (26) **end while**

ALGORITHM 1: FS optimized by R-IBACO.

feature count in the subset, prompting the inclusion of the selection ratio in the objective function, as specified in equation (13). By minimizing the objective function, the proposed methods can identify the least redundant and most informative feature subset with the highest classification accuracy.

$$\text{Accuracy} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N}, \quad (12)$$

$$\begin{aligned} \text{Fitness} &= \lambda \cdot (1 - \text{Accuracy}) + \mu \cdot \frac{n}{N} \\ \text{s.t. } \lambda &\in [0, 1], \lambda + \mu = 1, \\ 1 &\leq n \leq N, n = 1, 2, \dots, N. \end{aligned} \quad (13)$$

In equation (12),  $T_P$  and  $T_N$ , respectively, denote instances where the classifier correctly identified a test sample as positive or negative, whereas  $F_P$  and  $F_N$  represent instances where the classifier incorrectly classified a test sample as positive or negative. In equation (13), the parameters  $n$  and  $N$  refer to the count of features included in the selected feature subset and the total number of features, respectively. The fraction  $n/N$  signifies the feature selection rate. The fitness value, denoted by Fitness, is calculated by weighing the error rate and the selection rate using the weights  $\lambda$  and  $\mu$ , respectively. As the classifier's accuracy is the primary component of the fitness value,  $\lambda$  is typically assigned a larger value, such as 0.9.

## 5. Experimental Results and Discussion

In this section, the performance of the proposed methods is evaluated on fourteen high-dimensional biomedical datasets, with feature sizes ranging from 2000 to 12533 dimensions. The primary performance metrics considered are classification accuracy, size of feature subset, and running time. The K-nearest neighbor (KNN) classifier serves as the evaluation metric for the performance of the feature subset selected by each algorithm. The effectiveness of two methods for computing feature importance on model performance is also investigated. Five-fold cross-validation is employed to avoid model overfitting due to limited sample sizes, providing a more accurate assessment of model performance metrics. To thoroughly analyze the FS performance of the proposed methods, comparison experiments with thirteen other metaheuristic-based FS methods are conducted in three distinct groups. In the first group, the proposed methods are compared with standard HRO and IBACO, which are components of the proposed hybrid methods. The standard ACO is also included in this group to validate the effectiveness of the proposed heuristic factor. In the second group, the proposed methods are compared with five well-known FS methods based on basic metaheuristics, such as FPA, BQPSO, ABC, SSA [41], and GWO. In the final group, the effectiveness of the proposed hybrid methods is validated against five state-of-the-art methods reported in recent studies, which are CMSRSSMA [11], MBO [17], MSGWO



**Input:** Dataset  $D$  and objective function Fitness  
**Output:** OFS and corresponding classification accuracy

- (1) Initialize the subpopulations  $P_X$  and  $P_Y$  and set the maximum number of iterations  $T$
- (2) Calculate feature importance and knee point feature using Algorithm 3
- (3) Set heuristic factor and initialize PD
- (4) **while**  $t < T$  **do**
- (5)   Calculate  $f_X = \text{Fitness}(P_X)$  and  $f_Y = \text{Fitness}(P_Y)$
- (6)   Find  $p_{\text{best}x}$  and  $p_{\text{best}y}$ : individuals with min fitness in  $P_X$  and  $P_Y$
- (7)   Sort  $P_X$  in descending order and divide into three lines:  $M$  (Maintainer),  $R$  (Restorer),  $S$  (Sterile)
- (8)   **for** Individual  $X_i$  in  $P_X$  **do**
- (9)     **if**  $X_i$  in  $R$  **then**
- (10)       **if**  $SC < SC_{\text{max}}$  **then**
- (11)          Generate candidate solution  $X_{\text{new}(i)}$  using equation (6)
- (12)       **else**
- (13)          Generate candidate solution  $X_{\text{new}(i)}$  using equation (8)
- (14)       **end if**
- (15)     **else if**  $X_i$  in  $S$  **then**
- (16)       Generate candidate solution  $X_{\text{new}(i)}$  using equation (5)
- (17)     **end if**
- (18)     Calculate  $f_{X_{\text{new}(i)}} = \text{Fitness}(X_{\text{new}(i)})$
- (19)     Update  $X_i$  using equation (7)
- (20)     Check if  $X_i$  is an all-zero binary vector
- (21)   **end for**
- (22)   **for** Individual  $Y_i$  in  $P_Y$  **do**
- (23)     Update path of  $Y_i$  using equations (1) and (2)
- (24)     Check if  $Y_i$  is an all-zero binary vector
- (25)   **end for**
- (26)   Update  $p_{\text{best}x}$  and  $p_{\text{best}y}$
- (27)   Update the PD and the worst individual in maintainer line
- (28) **end while**
- (29)  $g_{\text{best}}$  is the best of  $p_{\text{best}x}$  and  $p_{\text{best}y}$

ALGORITHM 2: FS optimized by C-IBACO.

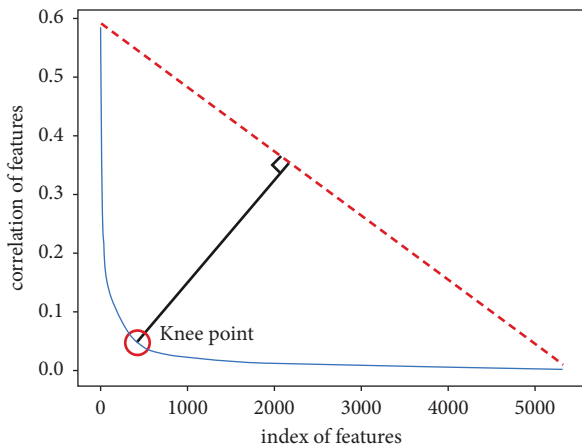


FIGURE 3: Knee point-based heuristic factor assignment.

[19], SCHHO [24], and HFSIA [25]. These advanced hybrid algorithms allow the proposed methods to be benchmarked against current leading solutions in the field.

All algorithms employed in this research were implemented using Python language version 3.6.9. The experimental results presented in this paper were obtained on a personal computer equipped with an Intel (R) Core (TM)

i7-8700 CPU operating at 3.2 GHz, 16.0 GB RAM, under a Windows 10 system.

**5.1. Description of Datasets.** Table 3 outlines the key characteristics of the datasets employed in this research. These datasets span a range of medical disciplines, predominantly serving binary or multi-class classification tasks. Notably, each dataset is characterized by a large number of features, varying from 2000 to 12533. A shared trait among these datasets is the presence of numerous redundant and irrelevant features, compounded by a typically limited sample size. It is imperative to apply dimensionality reduction to such data, as the presence of unnecessary features can compromise model performance.

**5.2. Parameter Settings.** The parameters for all algorithms are configured as follows. The maximum number of iterations is fixed at 100 for all algorithms. To ensure fairness in terms of maximum evaluation times of the fitness function ( $FE\_MAX = 3000$ ), the population size of all algorithms is set to 30 except for HRO, C-IBACO, and HFSIA. In consideration of the fact that only the sterile line is updated during the hybridization process of HRO and taking into account that the subpopulation sizes in C-IBACO are required to

<p><b>Input:</b> Dataset <math>D</math> and feature space <math>F</math></p> <p><b>Output:</b> Knee point and its weight <math>I</math></p> <ol style="list-style-type: none"> <li>(1) Set maximum vertical projection <math>d_{\max} \leftarrow 0</math></li> <li>(2) Initialize index of knee point <math>k</math> and its weight <math>I \leftarrow 0</math></li> <li>(3) Compute feature correlation for each <math>D_i</math> using RF</li> <li>(4) Sort the feature correlation in descending order, denoted by <math>W = (w_1, w_2, \dots, w_n)</math></li> <li>(5) Connect features with largest and smallest correlations to form line <math>L</math> (shown as red dotted line in Figure 3)</li> <li>(6) <b>for</b> each feature <math>f_j</math> in <math>F</math> <b>do</b></li> <li>(7)   Calculate vertical projection distance from <math>f_j</math> to line <math>L</math>, denoted as <math>d_j</math></li> <li>(8)   <b>if</b> <math>d_j &gt; d_{\max}</math> <b>then</b></li> <li>(9)     <math>k \leftarrow j</math></li> <li>(10)    <math>I \leftarrow w_j</math></li> <li>(11)    <b>end if</b></li> <li>(12) <b>end for</b></li> <li>(13) The knee point is the <math>k_{\text{th}}</math> feature in <math>F</math> (as marked with a red circle in Figure 3) and its weight is <math>I</math></li> </ol>
--

ALGORITHM 3: The selection process of the knee point.

TABLE 2: Comparative analysis of time complexity.

Stage	ACO	HRO	R-IBACO	C-IBACO
Initialization	$O(N)$	$O(N)$	$O(N)$	$O(N_1 + N_2)$
Evaluate and update population	$O(N \times D)$	$O(2/3N \times D)$	$O(N \times D)$	$O((2/3N_1 + N_2) \times D)$
Sort population	—	$O(N \times \log N)$	$O(N \times \log N)$	$O(N_1 \times \log N_1)$
Update PD	$O(D)$	—	$O(D)$	$O(D)$

TABLE 3: General information of the fourteen datasets.

Dataset	Number of features	Instances	Number of classes
Colon	2000	62	2
WarpAR10P	2400	130	10
Lung	3312	203	5
Lymphoma	4026	96	9
GLIOMA	4434	50	4
Leukemia_1	5327	72	3
DLBCL	5469	77	2
Brain_Tumor_1	5920	90	5
Prostate_GE	5966	102	2
ALLAML	7129	72	2
Brain_Tumor_2	10367	50	4
Prostate_Tumors	10509	102	2
Leukemia_2	11225	72	4
11_Tumor	12533	174	11

satisfy equation (14), where  $N_p$  stands for the total population size (fixed at 30), the population size for the single standard HRO is adjusted to 45. Additionally, the sub-population sizes for HRO and IBACO in C-IBACO are set to 27 and 12, respectively. For HFSIA, the population size is set at 14 and the crossover probability is set at 0.35 to maintain consistency in the evaluation times of the fitness function. The filter ratio of Fisher in HFSIA is set at 0.3 to preselect the top 30% of relevant features. Moreover, in the initial stage of MBAO, the number of features filtered by mRMR is fixed at 100, aligning with the existing literature. Lastly, to mitigate the effects of randomness and bolster the stability of the

experimental outcomes, each algorithm is executed independently on each dataset ten times.

$$N_p = \frac{2}{3}N_{\text{HRO}} + N_{\text{IBACO}}. \quad (14)$$

### 5.3. Experiment Analysis

**5.3.1. Experimental Evaluation of Methods for Computing Feature Correlation.** In alignment with the high-dimensional FS research framework delineated in Section 4, this section endeavors to assess and compare two feature correlation computation methods, including ReliefF and RF. The results derived from the application of KNN with all features serve as a control group. Classification accuracy and average number of selected features (AvgN) are designated as the principal evaluation metrics to provide a comprehensive performance appraisal of different feature correlation computation methods. The experimental findings are detailed in Table 4.

The evaluation results underscore the efficiency and effectiveness of RF and ReliefF in determining feature correlation and eliminating redundant features. It is noted that ReliefF tends to select more features than RF across all datasets, indicating the superior selection efficiency of RF. For example, the selection rate of RF in the lymphoma dataset reaches 95.85%, which is 90.59% higher than that of ReliefF. A similar outcome is observed in the warpAR10P dataset, where the selection rate of RF outperforms ReliefF by 88.32%. Moreover, both methods exhibit varying degrees of improvement in classification accuracy compared to using

TABLE 4: Experimental results for feature correlation.

Datasets	Original		RF		ReliefF	
	AvgN	Acc (%)	AvgN	Acc (%)	AvgN	Acc (%)
Colon	2000	83.97	<b>67</b>	87.05	232	<b>88.46</b>
WarpAR10P	2400	56.67	<b>192</b>	<b>68.67</b>	2142	54.67
Lung	3312	95.03	<b>108</b>	<b>95.50</b>	311	92.02
Lymphoma	4026	91.47	<b>167</b>	89.78	3663	<b>92.72</b>
GLIOMA	4434	76.77	<b>60</b>	<b>82.41</b>	384	67.14
Leukemia_1	5327	88.75	<b>57</b>	<b>91.90</b>	446	89.97
DLBCL	5469	89.46	<b>45</b>	<b>96.25</b>	316	93.57
Brain_Tumor_1	5920	84.48	<b>89</b>	<b>84.55</b>	340	83.37
Prostate_GE	5966	80.38	<b>76</b>	89.24	396	<b>92.14</b>
ALLAML	7129	76.38	<b>39</b>	91.52	539	<b>95.81</b>
Brain_Tumor_2	10367	70.45	<b>60</b>	<b>64.82</b>	533	70.64
Prostate_Tumors	10509	76.62	<b>71</b>	<b>88.24</b>	332	79.52
Leukemia_2	11225	<b>84.96</b>	<b>74</b>	82.56	409	82.19
11_Tumor	12533	<b>82.77</b>	<b>279</b>	71.02	954	81.03
Average	6180.07	81.30	<b>98.86</b>	<b>84.53</b>	785.5	83.09

Bold numbers represent the best average number of selected features and classification accuracy under each dataset.

all original, unfiltered features. In the ALLAML dataset, the classification accuracies of RF and ReliefF are recorded at 92.92% and 95.81%, respectively, marking improvements of 19.74% and 25.43% over the baseline. RF also delivers a classification accuracy of 68.67% in the warpAR10P dataset, which is 21.18% higher than the baseline. The enhancement in classification accuracy, coupled with an average dimensionality reduction rate of 98.4%, testifies to the ability of RF to effectively discard irrelevant features and calculate feature correlation in high-dimensional data.

However, the selected feature subset does not invariably guarantee improved classification accuracy. For instance, in the Brain\_Tumor\_2 dataset, RF selects a small feature subset of 60 features, achieving a dimensionality reduction rate of 99.42%. However, the classification accuracy merely attains 64.82%, marking a decrease of 7.99% from the baseline. This trend is similarly observed in lymphoma, Leukemia\_2, and 11\_Tumor datasets. The diminished classification accuracy underscores the limitations of FS methods that rely solely on feature correlation ranking, as crucial information tied to features with lower correlations than the knee point may be overlooked. Consequently, it might be more judicious to consider the correlation of all features as the criteria for high-dimensional feature ordering instead of discarding less correlated features.

The classification experiment results, detailed in Tables 5 and 6, demonstrate the effectiveness of employing the feature correlation determined by RF and ReliefF as the heuristic factor in the proposed IBACO. The findings suggest that C-IBACO and R-IBACO surpassed standalone RF or ReliefF in classification accuracy. For instance, when  $k$  is set to 3, the classification accuracies of RF-based C-IBACO and R-IBACO on the Brain\_Tumor\_2 dataset register at 88.82% and 86.82%, respectively, which are 22.28% and 18.51% superior to the accuracies of single RF. Furthermore, the classification accuracy obtained in the DLBCL dataset by C-IBACO using the feature correlation calculated by RF achieved 100%, which

indicates the ability of the proposed methods to identify samples that are challenging to distinguish between KNN and RF-based KNN. Compared to  $K = 5$ , the proposed methods exhibit an enhanced search capability when  $K$  is fixed at 3. As a result, in subsequent comparative experiments, KNN ( $K = 3$ ) will serve as the final classifier to evaluate the performance of feature subsets.

*5.3.2. Comparison of the Proposed Methods with HRO and IBACO.* To substantiate the effectiveness of the proposed heuristic factor assignment and hybrid strategies, the proposed methods were contrasted with the standard single algorithms. Table 7 summarizes the performance of the proposed methods and the basic methods in terms of classification accuracy, average number of selected features, and average runtime over ten independent runs. The results distinctly demonstrate that the mean accuracy of the proposed methods surpasses that of the original single algorithm across all datasets. Moreover, the introduction of an improved problem-oriented HF propels IBACO to outperform the standard ACO on ten of the fourteen datasets. Regarding the average number of selected features, the proposed methods obtained the smallest feature subsets on nine datasets. Although the standard ACO recorded the shortest average runtime across all datasets except the lung dataset, exhibiting its efficient selection capability, its lack of enhanced strategies culminated in inferior classification outcomes.

*5.3.3. Comparison of the Proposed Methods with Other Basic Metaheuristics.* To further investigate the superiority of the proposed methods, a fair comparison was conducted with five well-known basic metaheuristic-based FS methods, including FPA, BQPSO, ABC, ISSA [41], and GWO. The comparison maintained an equal maximum number of fitness evaluations, and the comparative outcomes of this group of methods are presented in Table 8. The proposed methods achieve the highest average and maximum classification accuracy across all datasets. BQPSO and GWO also deliver impressive results on most datasets, with classification accuracy on the Leukemia\_1 dataset even surpassing that of C-IBACO, a pattern also observed on the 11\_Tumor dataset.

Regarding the size of the selected feature subset, the proposed methods significantly outperform their counterparts. For instance, the feature selection rate of C-IBACO on the Prostate\_Tumors dataset is remarkably low, at only 0.41%. This is in stark contrast to the substantially higher feature selection rate of 49.47% displayed by FPA and IBSSA, a difference of 99.17%, which underscores the superior effectiveness of the proposed methods. Furthermore, C-IBACO selected an average of 143.6, 32.2, 53.1, and 23.6 features on the warpAR10P, GLIOMA, Prostate\_GE, and ALLAML datasets, respectively, with a minimum dimensionality reduction rate of 94.02%. R-IBACO selects the smallest average number of features on the Colon dataset, with only 57 features selected, a reduction of 94.09% compared to the average number of features chosen by IBSSA. The proposed methods record the shortest runtime on seven out of fourteen datasets. Notably, IBSSA

TABLE 5: Classification accuracy of IBACO using different feature correlations (KNN,  $K = 5$ ).

Datasets	Original	RF	Relieff	RF+		Relieff+	
				C-IBACO	R-IBACO	C-IBACO	R-IBACO
Colon	83.97	87.05	88.46	96.27	94.50	95.86	<b>96.83</b>
WarpAR10P	56.67	68.67	54.67	79.90	<b>80.23</b>	69.33	70.10
Lung	95.03	95.50	92.02	<b>98.91</b>	98.34	97.54	97.64
Lymphoma	91.47	89.78	92.72	95.97	<b>96.08</b>	95.44	95.28
GLIOMA	76.77	82.41	67.14	<b>90.14</b>	88.86	88.21	88.23
Leukemia_1	88.75	91.90	89.97	<b>99.33</b>	<b>99.33</b>	98.73	98.87
DLBCL	89.46	96.25	93.57	<b>100.00</b>	99.75	98.00	98.75
Brain_Tumor_1	84.48	84.55	83.37	<b>91.00</b>	89.98	89.54	89.89
Prostate_GE	80.38	89.24	92.14	95.29	94.90	94.39	<b>95.67</b>
ALLAML	76.38	91.52	95.81	99.00	<b>99.14</b>	94.77	98.44
Brain_Tumor_2	70.45	64.82	70.64	<b>88.82</b>	86.82	84.90	85.85
Prostate_Tumors	76.62	88.24	79.52	<b>95.31</b>	94.92	89.42	89.80
Leukemia_2	84.96	82.56	82.19	<b>94.00</b>	93.53	91.77	90.71
11_Tumor	82.77	71.02	81.03	<b>91.09</b>	90.42	90.76	90.02
AVERAGE	81.30	84.53	83.71	<b>93.97</b>	93.34	91.33	91.86

Bold numbers represent the best average classification accuracy for each dataset.

TABLE 6: Classification accuracy of IBACO using different feature correlations (KNN,  $K = 3$ ).

Datasets	Original	RF	Relieff	RF+		Relieff+	
				C-IBACO	R-IBACO	C-IBACO	R-IBACO
Colon	80.77	88.59	90.26	95.27	93.71	94.97	<b>95.94</b>
WarpAR10P	60.00	74.33	58.00	<b>83.07</b>	82.87	76.33	76.30
Lung	96.01	95.06	94.02	98.44	<b>98.54</b>	97.48	97.93
Lymphoma	91.89	94.25	91.89	<b>96.33</b>	<b>96.33</b>	95.51	95.64
GLIOMA	78.95	86.23	68.95	91.02	<b>91.09</b>	86.90	88.21
Leukemia_1	87.21	91.70	85.67	99.15	<b>99.60</b>	98.75	99.30
DLBCL	87.06	96.25	91.07	<b>100.00</b>	99.88	97.89	99.13
Brain_Tumor_1	87.88	83.50	85.66	<b>92.29</b>	92.08	91.18	90.45
Prostate_GE	80.52	90.24	90.19	<b>95.71</b>	95.52	94.22	94.70
ALLAML	79.05	92.95	93.05	<b>99.30</b>	98.86	95.73	98.29
Brain_Tumor_2	74.45	72.95	74.64	89.44	<b>90.15</b>	86.41	87.70
Prostate_Tumors	73.71	90.24	83.63	<b>95.31</b>	94.71	89.21	89.71
Leukemia_2	82.39	85.48	82.95	<b>95.47</b>	95.60	93.39	92.56
11_Tumor	83.84	75.08	78.52	<b>91.56</b>	90.90	90.06	89.13
AVERAGE	81.59	86.92	83.77	<b>94.45</b>	94.27	92.00	92.50

Bold numbers represent the best average classification accuracy for each dataset.

demonstrated higher computational efficiency than other methods on high-dimensional datasets (the last four datasets with over ten thousand dimensions).

In general, as the problem dimension increases, the performance gap between the proposed methods and other algorithms widens, indicating that C-IBACO and R-IBACO exhibit greater robustness and are more suitable for high-dimensional FS tasks.

*5.3.4. Comparative Study with the State of the Art.* Apart from comparing the proposed methods with single standard algorithms and basic metaheuristic-based FS methods, this study also probes into the performance comparison of C-IBACO and R-IBACO with five advanced metaheuristic-based FS methods recently expounded in the literature, including MBAO [17], HFSIA [25], SCHHO [24], MSGWO

[19], and CMSRSSMA [11]. Table 9 presents the comparison results of multiple independent runs for each advanced algorithm on each dataset. As evidenced in the table, C-IBACO and R-IBACO surpassed their counterparts in terms of classification accuracy on thirteen datasets, except for the lymphoma dataset. In the case of the lymphoma dataset, MBAO achieved the best classification accuracy while selecting the least number of features. This hinted at the limitations of the proposed methods in eliminating redundant and irrelevant features while bolstering classification accuracy. Although the application of filters substantially diminished the time consumed in the FS process, it consequently led to a drop in classification performance. This is illustrated by the average classification accuracy of MBAO, amounting to only 65.4 on the 11\_Tumor dataset. In this case, the classification accuracy of our proposed C-IBACO is 40.61% higher than that of

TABLE 7: Results of comparison between the proposed hybrid methods and their original counterparts.

Datasets	Algorithm	Max	Min	Mean + STD	AvgN	Time
Colon	C-IBACO	<b>98.33</b>	<b>91.92</b>	<b>95.27 + 1.94</b>	73.90	49.81
	R-IBACO	95.26	<b>91.92</b>	93.71 + 1.32	<b>57.00</b>	51.51
	BHRO	93.72	88.85	90.37 + 1.52	358.80	40.08
	IBACO	91.92	90.26	91.76 + 0.50	59.40	39.92
	ACO	93.59	90.26	92.04 + 1.42	902.10	<b>30.72</b>
WarpAR10P	C-IBACO	<b>84.67</b>	<b>81.33</b>	<b>83.07 + 0.98</b>	<b>143.60</b>	133.88
	R-IBACO	84.00	81.00	82.87 + 0.79	153.90	87.36
	BHRO	78.00	73.00	75.23 + 1.40	900.70	85.89
	IBACO	83.67	81.00	82.43 + 0.93	156.30	62.56
	ACO	53.08	51.54	52.08 + 0.52	1121.90	<b>42.98</b>
Lung	C-IBACO	98.50	98.00	98.44 + 0.15	323.00	278.02
	R-IBACO	<b>98.99</b>	<b>98.02</b>	<b>98.54 + 0.26</b>	367.40	276.07
	BHRO	98.49	97.97	98.24 + 0.25	818.10	256.76
	IBACO	98.00	96.06	96.79 + 0.54	<b>93.40</b>	<b>100.60</b>
	ACO	97.52	96.05	97.08 + 0.49	1509.40	102.01
Lymphoma	C-IBACO	<b>96.33</b>	<b>96.33</b>	<b>96.33 + 0.00</b>	145.20	132.89
	R-IBACO	<b>96.33</b>	<b>96.33</b>	<b>96.33 + 0.00</b>	140.50	119.55
	BHRO	<b>96.33</b>	95.08	96.11 + 0.45	577.60	96.42
	IBACO	<b>96.33</b>	95.08	95.21 + 0.38	<b>138.80</b>	80.40
	ACO	92.68	91.63	91.95 + 0.51	1797.80	<b>58.96</b>
GLIOMA	C-IBACO	<b>92.55</b>	<b>90.05</b>	91.02 + 1.00	<b>32.20</b>	103.66
	R-IBACO	<b>92.55</b>	<b>90.05</b>	<b>91.09 + 1.07</b>	44.00	97.84
	BHRO	90.05	84.41	87.81 + 1.42	617.30	78.20
	IBACO	91.86	90.05	90.23 + 0.55	40.70	71.13
	ACO	84.00	82.00	82.60 + 0.97	2024.70	<b>51.47</b>
Leukemia_1	C-IBACO	<b>100.00</b>	97.24	99.15 + 0.93	562.70	145.68
	R-IBACO	<b>100.00</b>	<b>98.67</b>	<b>99.60 + 0.61</b>	222.90	127.19
	BHRO	<b>100.00</b>	98.46	99.04 + 0.63	1771.10	109.93
	IBACO	97.24	95.81	96.96 + 0.55	<b>41.30</b>	83.20
	ACO	98.67	95.81	97.35 + 1.03	2507.20	<b>65.67</b>
DLBCL	C-IBACO	<b>100.00</b>	<b>100.00</b>	<b>100.00 + 0.00</b>	33.80	159.75
	R-IBACO	100.00	98.75	99.88 + 0.38	34.30	134.91
	BHRO	100.00	96.07	98.13 + 1.05	1272.70	135.82
	IBACO	97.50	97.50	97.50 + 0.00	<b>26.70</b>	87.01
	ACO	96.08	94.75	95.82 + 0.56	2510.80	<b>68.39</b>
Brain_Tumor_1	C-IBACO	<b>94.56</b>	<b>91.16</b>	<b>92.29 + 1.30</b>	631.60	143.3
	R-IBACO	93.44	<b>91.16</b>	92.08 + 0.88	<b>527.40</b>	136.09
	BHRO	91.11	88.89	90.00 + 0.91	1730.80	151.05
	IBACO	92.22	90.00	91.11 + 1.05	866.70	110.53
	ACO	88.89	87.78	88.56 + 0.54	2725.50	<b>81.58</b>
Prostate_GE	C-IBACO	<b>96.10</b>	<b>95.14</b>	<b>95.71 + 0.47</b>	<b>53.10</b>	190.07
	R-IBACO	<b>96.10</b>	<b>95.14</b>	95.52 + 0.47	54.30	181.44
	BHRO	91.24	89.33	90.47 + 0.71	1055.40	179.52
	IBACO	<b>96.10</b>	94.19	95.24 + 0.51	63.10	109.26
	ACO	87.29	83.29	85.95 + 1.06	2788.20	<b>101.51</b>
ALLAML	C-IBACO	<b>100.00</b>	<b>98.57</b>	<b>99.30 + 0.71</b>	<b>23.60</b>	194.03
	R-IBACO	<b>100.00</b>	<b>98.57</b>	98.86 + 0.57	23.90	172.54
	BHRO	92.86	88.67	90.67 + 1.28	1269.10	151.43
	IBACO	<b>100.00</b>	<b>98.57</b>	98.72 + 0.43	23.70	112.03
	ACO	86.00	83.33	85.29 + 0.94	3390.60	<b>94.55</b>
Brain_Tumor_2	C-IBACO	92.05	87.27	89.44 + 1.34	1122.70	224.97
	R-IBACO	<b>94.55</b>	<b>88.23</b>	<b>90.15 + 2.00</b>	711.90	234.40
	BHRO	90.05	86.41	88.05 + 0.98	2455.50	185.15
	IBACO	78.95	78.95	78.95 + 0.00	<b>40.10</b>	145.91
	ACO	84.00	82.00	83.20 + 1.03	4958.10	<b>106.05</b>
Prostate_Tumors	C-IBACO	<b>97.05</b>	<b>92.14</b>	<b>95.30 + 1.59</b>	<b>43.60</b>	293.19
	R-IBACO	96.10	<b>92.14</b>	94.71 + 1.62	52.60	283.53
	BHRO	91.14	87.29	89.65 + 1.21	3287.60	317.04
	IBACO	91.19	83.33	88.46 + 2.20	1932.10	243.51
	ACO	89.29	85.38	87.24 + 1.61	5060.70	<b>185.11</b>

TABLE 7: Continued.

Datasets	Algorithm	Max	Min	Mean + STD	AvgN	Time
Leukemia_2	C-IBACO	<b>97.32</b>	<b>94.56</b>	95.47 + 1.11	<b>715.60</b>	158.47
	R-IBACO	<b>97.32</b>	93.40	<b>95.60 + 1.25</b>	936.70	159.02
	BHRO	94.48	91.62	92.80 + 0.88	1839.20	170.61
	IBACO	94.48	88.86	91.83 + 2.00	1325.30	121.25
	ACO	93.05	90.19	90.96 + 0.98	3364.10	<b>81.99</b>
11_Tumor	C-IBACO	<b>94.99</b>	89.43	<b>91.96 + 1.16</b>	<b>2629.20</b>	382.04
	R-IBACO	94.27	89.17	91.18 + 1.00	3156.30	378.14
	BHRO	93.09	<b>90.79</b>	91.77 + 0.72	5342.90	448.75
	IBACO	91.95	87.92	89.69 + 1.32	2728.80	290.66
	ACO	91.93	89.04	90.15 + 0.96	6139.50	<b>275.12</b>

Bold numbers represent the best results on each dataset.

TABLE 8: Performance comparison with other basic metaheuristic-based methods.

Datasets	Algorithm	Max	Min	Mean + STD	AvgN	Time
Colon	C-IBACO	<b>98.33</b>	<b>91.92</b>	<b>95.27 + 1.94</b>	73.90	49.81
	R-IBACO	95.26	<b>91.92</b>	93.71 + 1.32	<b>57.00</b>	51.51
	BFPA	88.97	87.31	88.17 + 0.81	922.80	<b>44.63</b>
	BQPSO	88.97	87.31	88.15 + 0.82	756.20	64.28
	BABC	87.44	84.10	85.72 + 1.28	623.80	72.68
	IBSSA	87.31	84.10	85.58 + 0.87	964.70	48.25
	BGWO	90.51	87.31	88.60 + 0.95	827.00	84.15
WarpAR10P	C-IBACO	<b>84.67</b>	<b>81.33</b>	<b>83.07 + 0.98</b>	<b>143.60</b>	133.88
	R-IBACO	84.00	81.00	82.87 + 0.79	153.90	<b>87.36</b>
	BFPA	75.00	72.33	73.93 + 0.74	1177.30	120.47
	BQPSO	77.00	72.67	75.70 + 1.22	1055.80	135.75
	BABC	75.67	69.67	72.80 + 1.73	889.80	202.24
	IBSSA	72.67	69.33	71.20 + 1.03	1180.80	126.72
Lung	BGWO	77.00	73.67	75.70 + 0.86	1137.90	187.09
	C-IBACO	98.50	98.00	98.44 + 0.15	<b>323.00</b>	278.02
	R-IBACO	<b>98.99</b>	<b>98.02</b>	<b>98.54 + 0.26</b>	367.40	<b>276.07</b>
	BFPA	98.49	97.97	98.10 + 0.19	1569.80	383.08
	BQPSO	98.49	97.97	98.24 + 0.25	1346.40	393.66
	BABC	98.00	97.02	97.46 + 0.34	1217.70	665.02
Lymphoma	IBSSA	98.00	97.02	97.56 + 0.26	1608.70	385.90
	BGWO	98.49	97.97	98.23 + 0.25	1415.90	503.41
	C-IBACO	<b>96.33</b>	<b>96.33</b>	<b>96.33 + 0.00</b>	145.20	132.89
	R-IBACO	<b>96.33</b>	<b>96.33</b>	<b>96.33 + 0.00</b>	<b>140.50</b>	<b>119.55</b>
	BFPA	95.22	93.97	94.76 + 0.52	1816.80	146.87
	BQPSO	95.22	93.97	94.78 + 0.53	1568.90	177.84
GLIOMA	BABC	95.08	93.97	94.08 + 0.33	1522.70	226.52
	IBSSA	95.22	93.97	94.21 + 0.47	1932.70	145.15
	BGWO	95.22	93.97	94.76 + 0.52	1654.30	254.29
	C-IBACO	<b>92.55</b>	<b>90.05</b>	91.02 + 1.00	<b>32.20</b>	103.66
	R-IBACO	<b>92.55</b>	<b>90.05</b>	<b>91.09 + 1.07</b>	44.00	97.84
	BFPA	86.23	84.41	85.86 + 0.73	2009.90	<b>93.65</b>
Leukemia_1	BQPSO	90.05	84.41	86.08 + 1.78	1759.70	122.37
	BABC	84.41	82.59	82.77 + 0.55	1718.00	139.92
	IBSSA	84.41	82.59	83.50 + 0.91	2124.30	94.11
	BGWO	86.23	84.41	85.32 + 0.91	1840.40	218.09
	C-IBACO	<b>100.00</b>	97.24	99.15 + 0.93	562.70	145.68
	R-IBACO	<b>100.00</b>	<b>98.67</b>	<b>99.60 + 0.61</b>	<b>222.90</b>	<b>127.19</b>
Leukemia_2	BFPA	<b>100.00</b>	<b>98.67</b>	99.33 + 0.67	2589.90	135.48
	BQPSO	<b>100.00</b>	<b>98.67</b>	99.20 + 0.65	2261.50	186.38
	BABC	98.67	94.57	97.36 + 1.30	2205.10	210.64
	IBSSA	98.67	95.81	97.79 + 0.95	2623.70	133.27
	BGWO	<b>100.00</b>	<b>98.67</b>	99.47 + 0.65	2411.40	305.17

TABLE 8: Continued.

Datasets	Algorithm	Max	Min	Mean + STD	AvgN	Time
DLBCL	C-IBACO	<b>100.00</b>	<b>100.00</b>	<b>100.00 + 0.00</b>	<b>33.80</b>	159.75
	R-IBACO	<b>100.00</b>	98.75	99.88 + 0.38	34.30	<b>134.91</b>
	BFPA	97.32	96.07	97.20 + 0.38	2649.10	165.45
	BQPSO	97.32	96.07	96.70 + 0.63	2311.40	206.95
	BABC	96.07	93.57	94.82 + 0.97	2233.30	276.54
	IBSSA	97.32	93.49	95.30 + 1.02	2680.50	172.03
	BGWO	97.32	94.82	96.82 + 0.83	2436.80	357.81
Brain_Tumor_1	C-IBACO	<b>94.56</b>	<b>91.16</b>	<b>92.29 + 1.30</b>	631.60	143.30
	R-IBACO	93.44	<b>91.16</b>	92.08 + 0.88	<b>527.40</b>	136.09
	BFPA	90.10	87.88	89.21 + 0.70	2831.80	110.34
	BQPSO	90.10	87.88	89.21 + 0.70	2479.80	157.36
	BABC	88.89	86.67	88.00 + 0.88	2438.40	163.02
	IBSSA	88.99	86.70	87.95 + 0.64	2867.50	<b>105.20</b>
	BGWO	91.21	87.82	89.64 + 0.96	2625.80	236.60
Prostate_GE	C-IBACO	<b>96.10</b>	<b>95.14</b>	<b>95.71 + 0.47</b>	<b>53.10</b>	190.07
	R-IBACO	<b>96.10</b>	<b>95.14</b>	95.52 + 0.47	54.30	<b>181.44</b>
	BFPA	88.38	87.43	87.71 + 0.44	2861.80	269.91
	BQPSO	87.43	87.43	87.43 + 0.00	2532.50	308.30
	BABC	85.52	83.52	84.69 + 0.60	2474.40	434.64
	IBSSA	87.43	83.52	85.58 + 1.19	2892.30	266.88
	BGWO	89.33	86.43	87.51 + 1.01	2690.20	486.84
ALLAML	C-IBACO	<b>100.00</b>	<b>98.57</b>	<b>99.30 + 0.71</b>	<b>23.60</b>	194.03
	R-IBACO	<b>100.00</b>	<b>98.57</b>	98.86 + 0.57	23.90	<b>172.54</b>
	BFPA	91.43	87.24	88.55 + 1.15	3456.50	198.63
	BQPSO	90.00	87.24	88.68 + 0.86	3160.00	248.23
	BABC	88.67	81.81	84.68 + 1.79	3057.10	319.67
	IBSSA	86.00	83.14	84.82 + 1.03	3487.40	195.86
	BGWO	91.43	87.33	88.98 + 1.47	3293.30	483.79
Brain_Tumor_2	C-IBACO	92.05	87.27	89.44 + 1.34	1122.70	224.97
	R-IBACO	<b>94.55</b>	<b>88.23</b>	<b>90.15 + 2.00</b>	<b>711.90</b>	234.40
	BFPA	88.23	86.41	87.50 + 0.89	5019.70	213.69
	BQPSO	88.23	86.23	88.01 + 0.60	4620.00	318.11
	BABC	86.41	80.77	83.57 + 1.51	4620.10	309.57
	IBSSA	86.41	82.59	85.43 + 1.25	5125.80	<b>210.23</b>
	BGWO	88.23	86.41	87.68 + 0.83	4788.90	770.95
Prostate_Tumors	C-IBACO	<b>97.05</b>	<b>92.14</b>	<b>95.30 + 1.59</b>	<b>43.60</b>	293.19
	R-IBACO	96.10	<b>92.14</b>	94.71 + 1.62	52.60	283.53
	BFPA	90.29	87.38	89.00 + 0.91	5198.30	231.91
	BQPSO	91.14	88.33	89.76 + 1.01	4834.80	334.19
	BABC	88.29	84.43	85.66 + 1.12	4731.60	374.56
	IBSSA	88.29	84.43	86.66 + 1.29	5198.80	<b>229.40</b>
	BGWO	92.19	89.24	90.44 + 1.01	5018.40	613.25
Leukemia_2	C-IBACO	<b>97.32</b>	<b>94.56</b>	95.47 + 1.11	<b>715.60</b>	158.47
	R-IBACO	<b>97.32</b>	93.40	<b>95.60 + 1.25</b>	936.70	159.02
	BFPA	93.40	91.74	92.63 + 0.67	3488.80	115.89
	BQPSO	93.40	91.98	93.09 + 0.59	3182.70	175.82
	BABC	91.62	88.86	90.11 + 0.79	3046.00	171.42
	IBSSA	91.98	87.72	89.90 + 1.18	3527.70	<b>114.06</b>
	BGWO	94.65	91.65	93.17 + 1.08	3316.80	297.47
11_Tumor	C-IBACO	<b>94.99</b>	89.43	<b>91.96 + 1.16</b>	<b>2629.20</b>	382.04
	R-IBACO	94.27	89.17	91.18 + 1.00	3156.30	378.14
	BFPA	91.39	89.02	90.15 + 0.65	6422.00	345.36
	BQPSO	93.28	<b>90.23</b>	91.88 + 0.95	6176.00	448.97
	BABC	90.20	86.77	89.06 + 1.01	5771.90	568.54
	IBSSA	89.53	86.79	88.15 + 0.88	6315.00	<b>334.28</b>
	BGWO	92.52	89.33	91.73 + 0.93	6225.40	885.58

Bold numbers represent the best results on each dataset.

TABLE 9: Comparative study between the proposed hybrid methods and the state-of-the-art methods.

Datasets	Algorithm	Max	Min	Mean + STD	AvgN	Time
Colon	C-IBACO	<b>98.33</b>	91.92	<b>95.27 + 1.94</b>	73.90	49.81
	R-IBACO	95.26	91.92	93.71 + 1.32	57.00	51.51
	MBAO	95.00	<b>93.33</b>	93.76 + 0.66	<b>29.60</b>	53.28
	HFSIA	93.33	90.13	92.06 + 1.01	288.70	<b>23.36</b>
	SCHHO	93.46	88.72	91.08 + 1.34	641.10	99.91
	MSGWO	93.46	91.92	92.69 + 0.81	866.50	39.30
	CMSRSSMA	91.92	90.26	91.44 + 0.76	977.60	260.7
WarpAR10P	C-IBACO	<b>84.67</b>	<b>81.33</b>	<b>83.07 + 0.98</b>	143.60	133.88
	R-IBACO	84.00	81.00	82.87 + 0.79	153.90	87.36
	MBAO	82.33	80.00	81.10 + 0.80	<b>48.30</b>	69.28
	HFSIA	77.69	76.15	76.54 + 0.54	351.30	<b>33.32</b>
	SCHHO	71.00	63.67	67.37 + 2.53	190.20	136.61
	MSGWO	78.52	72.49	75.80 + 1.03	1098.90	53.70
	CMSRSSMA	77.02	73.82	75.04 + 0.57	1182.10	348.66
Lung	C-IBACO	98.50	98.00	98.44 + 0.15	323.00	278.02
	R-IBACO	<b>98.99</b>	<b>98.02</b>	<b>98.54 + 0.26</b>	367.40	276.07
	MBAO	94.58	93.58	93.96 + 0.33	<b>39.40</b>	89.15
	HFSIA	97.52	97.04	97.43 + 0.21	486.40	<b>50.71</b>
	SCHHO	97.51	96.52	97.06 + 0.28	1023.30	299.71
	MSGWO	97.52	96.54	96.88 + 0.41	1457.30	119.07
	CMSRSSMA	97.02	96.54	96.78 + 0.26	1632.00	688.81
Lymphoma	C-IBACO	<b>96.33</b>	<b>96.33</b>	<b>96.33 + 0.00</b>	145.20	132.89
	R-IBACO	<b>96.33</b>	<b>96.33</b>	<b>96.33 + 0.00</b>	140.50	119.55
	MBAO	<b>96.33</b>	<b>96.33</b>	<b>96.33 + 0.00</b>	<b>31.50</b>	55.85
	HFSIA	95.79	94.74	94.97 + 0.43	587.00	<b>37.61</b>
	SCHHO	94.39	93.56	94.22 + 0.35	1109.20	199.45
	MSGWO	92.68	91.63	92.47 + 0.44	1732.90	76.64
	CMSRSSMA	92.68	91.68	92.48 + 0.42	1984.50	588.36
GLIOMA	C-IBACO	<b>92.55</b>	<b>90.05</b>	91.02 + 1.00	<b>32.20</b>	103.66
	R-IBACO	<b>92.55</b>	<b>90.05</b>	<b>91.09 + 1.07</b>	44.00	97.84
	MBAO	84.41	80.95	82.63 + 0.82	46.80	47.35
	HFSIA	92.00	90.00	90.80 + 1.03	641.50	<b>35.72</b>
	SCHHO	86.91	84.23	85.10 + 1.11	514.10	192.22
	MSGWO	84.00	82.00	82.40 + 0.84	1921.50	69.39
	CMSRSSMA	84.00	82.00	83.00 + 1.05	2182.00	621.72
Leukemia_1	C-IBACO	<b>100.00</b>	97.24	99.15 + 0.93	562.70	145.68
	R-IBACO	<b>100.00</b>	<b>98.67</b>	<b>99.60 + 0.61</b>	222.90	127.19
	MBAO	87.30	83.41	85.26 + 1.20	<b>24.50</b>	50.76
	HFSIA	<b>100.00</b>	97.33	98.61 + 0.63	790.80	<b>42.72</b>
	SCHHO	98.46	92.62	96.39 + 1.59	2315.30	228.61
	MSGWO	<b>100.00</b>	95.81	97.77 + 1.64	2415.70	88.81
	CMSRSSMA	97.24	95.81	96.40 + 0.70	2656.30	803.21
DLBCL	C-IBACO	<b>100.00</b>	<b>100.00</b>	<b>100.00 + 0.00</b>	<b>33.80</b>	159.75
	R-IBACO	<b>100.00</b>	98.75	99.88 + 0.38	34.30	134.91
	MBAO	92.50	91.17	91.93 + 0.65	36.30	52.04
	HFSIA	98.67	97.42	97.79 + 0.60	806.20	<b>44.97</b>
	SCHHO	97.42	94.82	95.74 + 0.86	1407.10	238.84
	MSGWO	97.42	94.75	96.49 + 1.09	2473.00	94.55
	CMSRSSMA	97.42	96.08	96.25 + 0.41	2725.00	837.04
Brain_Tumor_1	C-IBACO	<b>94.56</b>	<b>91.16</b>	<b>92.29 + 1.30</b>	631.60	143.30
	R-IBACO	93.44	<b>91.16</b>	92.08 + 0.88	527.40	136.09
	MBAO	84.61	81.28	83.02 + 0.95	<b>34.00</b>	59.12
	HFSIA	91.11	90.00	90.22 + 0.47	874.40	<b>50.86</b>
	SCHHO	90.11	87.76	88.51 + 0.80	710.50	266.94
	MSGWO	90.10	87.82	88.72 + 0.89	2677.20	103.88
	CMSRSSMA	88.99	87.82	88.20 + 0.52	2926.80	949.01



TABLE 9: Continued.

Datasets	Algorithm	Max	Min	Mean + STD	AvgN	Time
Prostate_GE	C-IBACO	<b>96.10</b>	<b>95.14</b>	<b>95.71 + 0.47</b>	53.10	190.07
	R-IBACO	<b>96.10</b>	<b>95.14</b>	95.52 + 0.47	54.30	181.44
	MBAO	90.24	89.14	89.80 + 0.51	<b>30.10</b>	59.01
	HFSIA	89.24	88.24	88.73 + 0.50	879.40	<b>55.67</b>
	SCHHO	93.05	85.29	88.55 + 2.98	337.40	330.33
	MSGWO	87.29	85.29	86.86 + 0.68	2731.50	131.99
	CMSRSSMA	86.24	85.24	85.58 + 0.46	2947.30	1044.34
ALLAML	C-IBACO	<b>100.00</b>	<b>98.57</b>	<b>99.30 + 0.71</b>	<b>23.60</b>	194.03
	R-IBACO	<b>100.00</b>	<b>98.57</b>	98.86 + 0.57	23.90	172.54
	MBAO	92.95	90.10	91.30 + 1.08	34.90	<b>52.66</b>
	HFSIA	97.24	95.71	95.90 + 0.47	1065.10	56.47
	SCHHO	91.62	84.57	87.97 + 2.55	491.10	334.98
	MSGWO	88.76	84.67	86.80 + 1.15	3301.10	122.14
	CMSRSSMA	84.67	83.24	84.08 + 0.66	3543.00	1214.39
Brain_Tumor_2	C-IBACO	92.05	87.27	89.44 + 1.34	1122.70	224.97
	R-IBACO	<b>94.55</b>	<b>88.23</b>	<b>90.15 + 2.00</b>	711.90	234.40
	MBAO	75.23	71.14	72.44 + 1.20	<b>31.70</b>	<b>47.45</b>
	HFSIA	86.00	84.00	84.60 + 0.97	1531.10	83.04
	SCHHO	89.57	83.30	86.46 + 1.35	2263.40	368.85
	MSGWO	92.48	83.43	88.45 + 1.84	4831.60	146.46
	CMSRSSMA	90.51	85.19	87.80 + 1.23	5190.10	1980.1
Prostate_Tumors	C-IBACO	<b>97.05</b>	<b>92.14</b>	<b>95.30 + 1.59</b>	43.60	293.19
	R-IBACO	96.10	<b>92.14</b>	94.71 + 1.62	52.60	283.53
	MBAO	89.24	87.14	88.13 + 0.74	<b>38.20</b>	<b>64.53</b>
	HFSIA	87.24	85.24	86.00 + 0.66	1566.30	101.11
	SCHHO	93.10	86.24	88.46 + 2.38	377.00	575.53
	MSGWO	90.24	87.29	88.49 + 0.90	5020.80	235.13
	CMSRSSMA	88.29	84.43	85.77 + 1.03	5241.90	2348.15
Leukemia_2	C-IBACO	<b>97.32</b>	<b>94.56</b>	95.47 + 1.11	715.60	158.47
	R-IBACO	<b>97.32</b>	93.40	<b>95.60 + 1.25</b>	936.70	159.02
	MBAO	87.46	85.06	85.95 + 0.84	<b>42.90</b>	<b>55.32</b>
	HFSIA	93.05	91.62	91.77 + 0.45	1061.10	57.22
	SCHHO	90.74	87.73	89.71 + 0.91	2425.50	293.48
	MSGWO	93.40	90.31	91.79 + 1.11	3361.90	112.56
	CMSRSSMA	90.40	89.06	89.78 + 0.62	3574.80	1173.43
11_Tumor	C-IBACO	<b>94.99</b>	<b>89.43</b>	<b>91.96 + 1.16</b>	2629.20	382.04
	R-IBACO	94.27	89.17	91.18 + 1.00	3156.30	378.14
	MBAO	66.36	64.41	65.40 + 0.65	<b>59.60</b>	<b>83.30</b>
	HFSIA	90.24	89.09	89.86 + 0.39	1872.40	149.68
	SCHHO	89.07	84.15	87.12 + 1.70	5557.20	889.44
	MSGWO	90.96	88.90	90.06 + 0.72	6087.90	341.85
	CMSRSSMA	90.22	87.90	88.71 + 0.78	6258.60	3278.5

Bold numbers represent the best results on each dataset.

MBAO. Similar trends were also observed in the Brain\_Tumor\_2 and warpAR10P datasets, where R-IBACO and C-IBACO were 24.45% and 23.30% higher than MBAO and SCHHO, respectively, which might be due to filters eliminating features potentially useful for classification.

While emphasizing classification accuracy, the hybrid filter approaches managed to select the minimum number of features across all datasets, with MBAO achieving thirteen and HFSIA securing one. This outcome can be ascribed to the ability of filter to constrain the search space to a lower dimension, thereby significantly shrinking the size of the feature subset to be searched. Regarding computational time, the hybrid filter methods demonstrate superior search efficiency as they conduct the search within the filtered, low-dimensional

feature space. Specifically, HFSIA registers the shortest computational time on the initial nine datasets, whereas MBAO exhibits exceptional search efficiency on the final five datasets characterized by higher dimensionality. In the context of search efficiency across the full feature space, both C-IBACO and R-IBACO perform slightly less effectively than MSGWO, which does not utilize a hybrid strategy. However, C-IBACO and R-IBACO prove to be more robust than SCHHO and CMSRSSMA, particularly as the feature dimension increases.

Although the C-IBACO and R-IBACO may not match the computational efficiency and selected feature subset size exhibited by MBAO and HFSIA, they still demonstrated acceptable performance and robustness in classification accuracy and stability, which is worth a slight sacrifice in time.

**5.3.5. Graphical Analysis.** Figure 4 provides a visual representation of the comparative superiority and efficacy of the proposed methods against the single standard algorithms and other metaheuristic-based FS approaches. This figure encapsulates the overall maximum and average classification accuracy, along with the number of selected features for all algorithms. Specifically, Figures 4(a) and 4(b) depict the overall maximum and average classification accuracies, respectively. As the figures suggest, the feature subsets selected by C-IBACO and R-IBACO deliver markedly superior classification performance compared to other methods. Higher overall classification accuracy underscores the superior search performance of the proposed methods and a heightened likelihood of discovering more promising candidate solutions. Conversely, MBAO did not exhibit optimal performance in overall classification accuracy, as discussed in the preceding section. This discrepancy could be due to the classifier used, which differs from those documented in the literature. Figure 4(c) displays the overall number of features selected by each algorithm. MBAO selected the fewest features, followed closely by R-IBACO and C-IBACO. It is worth noting that the number of features chosen by the proposed methods was less than that of HFSIA, even though the latter also employs a filtering method. This result reaffirms the effectiveness of the proposed methods in reducing the number of selected features.

The convergence results, illustrated in Figure 5, demonstrate that both C-IBACO and R-IBACO converge quickly to the vicinity of the global optimum before the termination of iterations. The hybrid algorithms possess an advantage in achieving a superior initial solution by properly selecting features with high correlation based on the modified heuristic factor in the first stage. As the convergence curves indicate, the proposed methods maintain a certain exploratory capability in the final stage of iterations, reducing the risk of getting trapped in the local optima. In contrast, the convergence curves of MBAO reveal its relatively underwhelming overall performance. Specifically, on datasets such as lung, Brain\_Tumor\_2, Leukemia\_2, and 11\_Tumor, MBAO prematurely converges to a local optimum and fails to further search for a better feature subset. One possible explanation for this is that mRMR filters out informative feature combinations, resulting in a search space with poor performance.

To provide a more intuitive understanding of the computational complexity of different algorithms, Figure 6 displays the average running time of all algorithms for each dataset. Evidently, HFSIA attains the shortest running time for the first nine datasets, whereas MBAO demonstrates superior search efficiency for the last five datasets with higher dimensionality. It is noticeable that IBACO achieves a shorter average CPU running time than HRO and their hybrid counterparts across all datasets, which can be attributed to its ability to quickly select features with high correlation using the modified HF, thereby accelerating convergence. However, this advantage is accompanied by lower classification accuracy and unsatisfactory average fitness values compared to the hybrid algorithms. Nevertheless, C-IBACO and R-IBACO display exceptional

robustness by generating acceptable candidate solutions without incurring a significant computational cost. Overall, the experimental results affirm the feasibility and potential of deploying the proposed methods for practical high-dimensional FS tasks.

**5.3.6. Experimental Results for Nonparametric Test.** To determine the significance of the difference between the proposed hybrid algorithms and compared metaheuristic-based FS approaches, the experimental results were analyzed using the Wilcoxon signed-rank test and the Friedman test [55]. The results of the Wilcoxon signed-rank test are presented in Table 10. In this table,  $W^+$  denotes the sum of ranks where the proposed hybrid algorithm outperforms the comparative method, whereas  $W^-$  symbolizes the reverse. The  $p$  value signifies the level of significance, with  $p < 0.05$  indicating a significant difference between the two algorithms under comparison. The outcomes reveal that C-IBACO outperforms R-IBACO, suggesting that C-IBACO maintains a lower individual fitness value and exhibits superior performance in the FS process. Apart from the null hypothesis being rejected between C-IBACO and R-IBACO, both are significantly superior to other basic and advanced methods, demonstrating significant differences.

Table 11 presents the results of the Friedman test, including average accuracy rankings and final ranks for each algorithm across all datasets. As demonstrated in Table 11, the proposed methods claim the top two positions in the final ranking. The Friedman test produces a  $p$  value of  $2.61E-18$ , which is less than the preset significance level of 0.01, indicating a statistical difference among the algorithms. Table 12 provides the results of the Holm test, which serves as a post hoc test to determine significant differences between the control method and other algorithms in pairwise comparisons. The Holm test rejected the null hypothesis at a significance level of 0.05, indicating significant differences between R-IBACO and other competing methods except for C-IBACO, which is consistent with the findings of the Wilcoxon signed-rank test.

Based on the statistical test results, it can be inferred that the hybrid algorithms incorporating IBACO and HRO exhibit superior optimization performance compared to their standalone counterparts. This emphasizes the effective combination of the optimization strategies of every single algorithm, leading to an enhancement in their competence in high-dimensional FS tasks.

**5.4. Discussion and Biological Interpretation.** Based on the results of three comparative experiments and statistical analysis, it can be inferred that the hybrid models of HRO and IBACO integrate the advantages of both single algorithms, allowing them to converge more quickly to the most promising region in the search space while maintaining a certain level of global exploration capability in later iterations. The problem-oriented HF enhances the ability of ACO to select more important features while discarding irrelevant ones. Moreover, the optimal solution obtained by HRO guides the update of pheromones. This collaborative

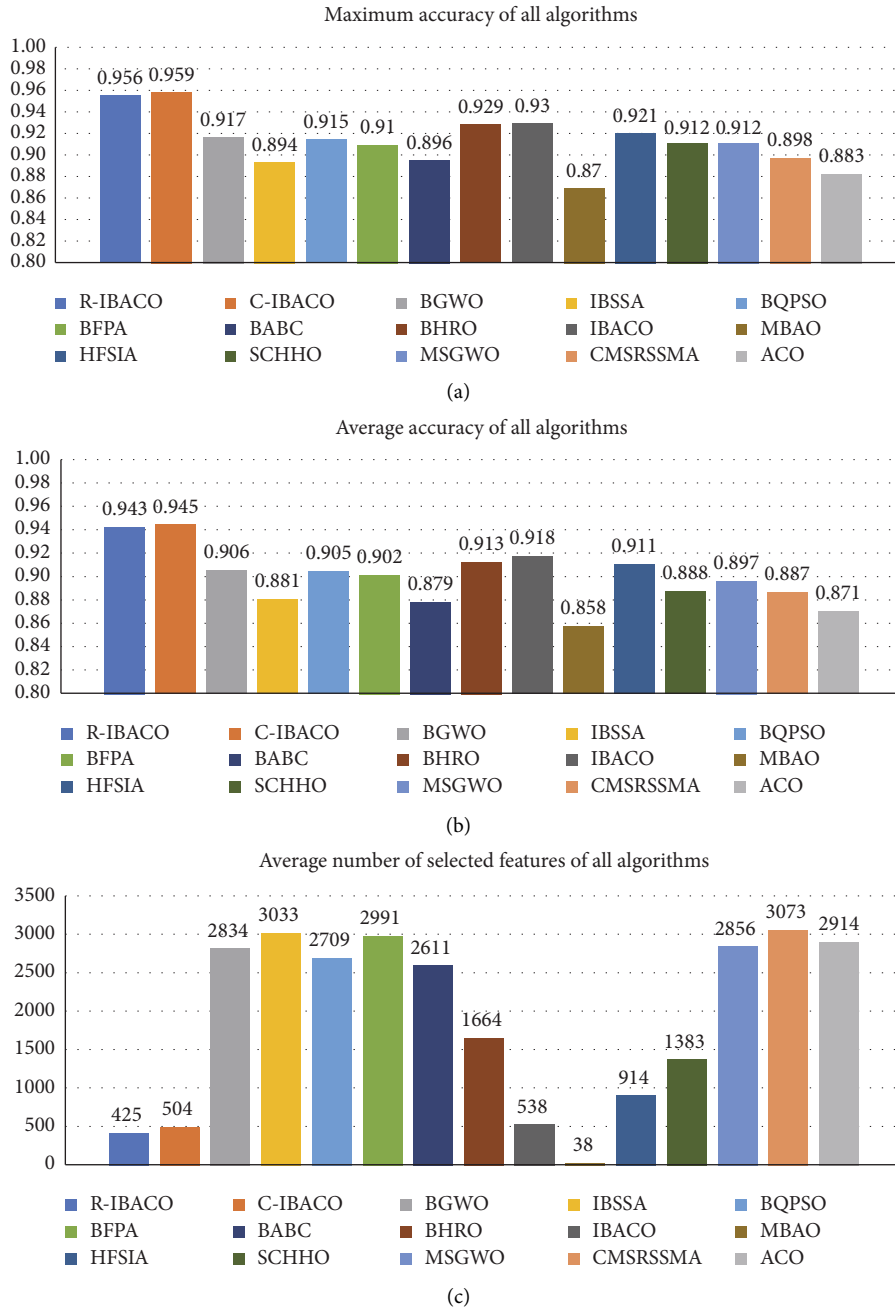
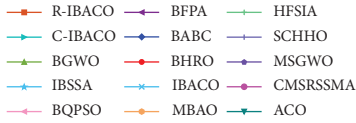
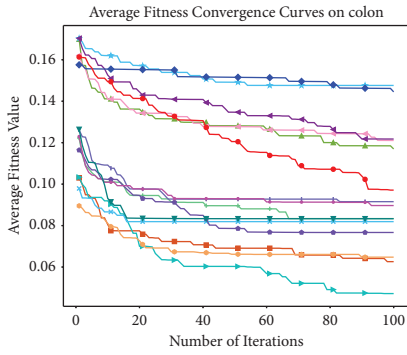


FIGURE 4: The statistics of key indicators for all algorithms. (a) The maximum classification accuracy for all compared algorithms. (b) The average classification accuracy for all compared algorithms. (c) The average number of selected features for all algorithms.

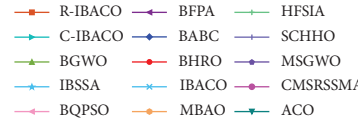
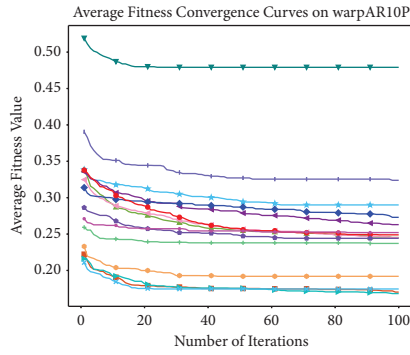
update process continues until the termination condition is achieved. By combining HRO and IBACO, the population benefits from greater diversity while reducing the probability of a single algorithm getting stuck into the local optima.

Numerous metaheuristics have been developed to solve low-dimensional continuous optimization problems, but these often struggle when addressing high-dimensional discrete optimization issues. This study introduces two hybrid wrapper FS methods, and the experimental results on fourteen high-dimensional datasets demonstrate that the feature subset obtained by the hybrid algorithms had higher

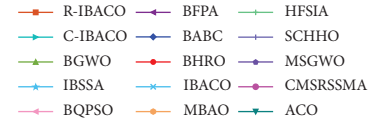
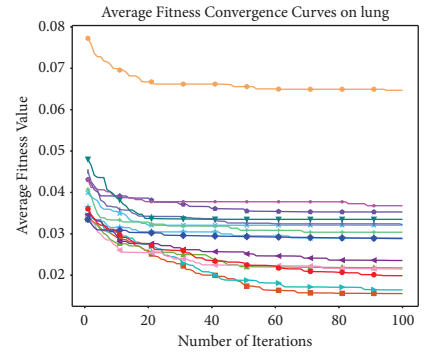
classification accuracy and smaller size. Nevertheless, their optimization performance still needs to improve in some instances, a shortcoming attributable to the inherent traits of metaheuristics. To enhance the potential of locating the optimal solution, metaheuristics often adopt a global search strategy that depends on certain randomness. This approach inevitably raises the risk of the algorithm getting stuck in the local optima. For instance, the algorithms only managed to achieve an average classification accuracy of 83% on the warpARIOP dataset, a rate that may not satisfy the stringent standards demanded in practical medical scenarios.



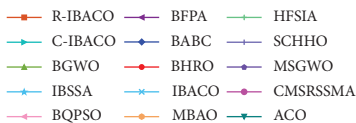
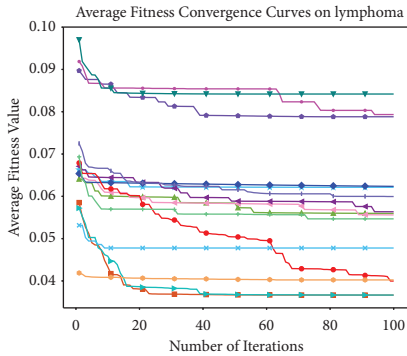
(a)



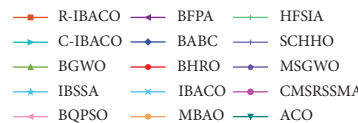
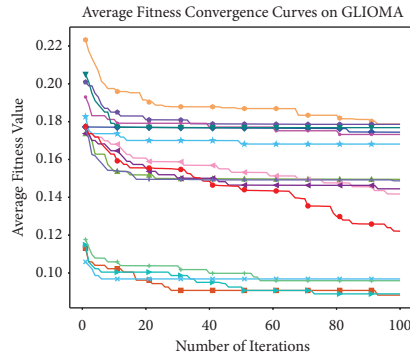
(b)



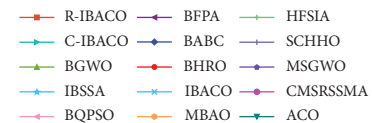
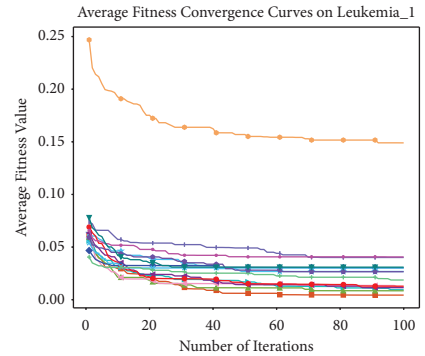
(c)



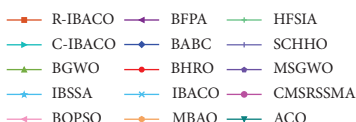
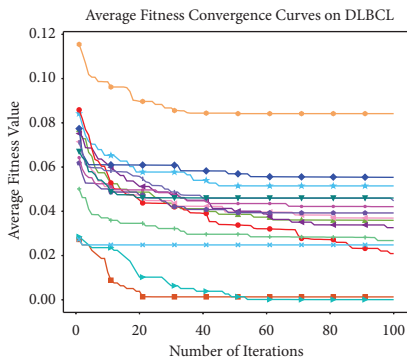
(d)



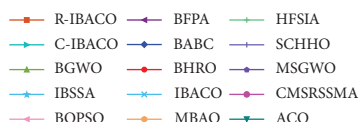
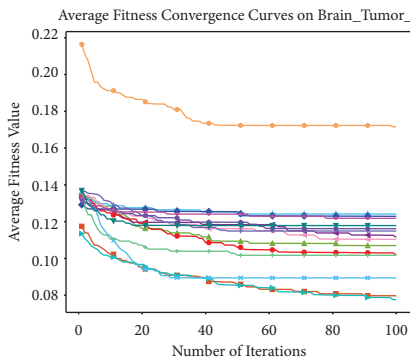
(e)



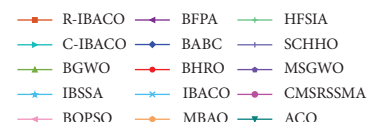
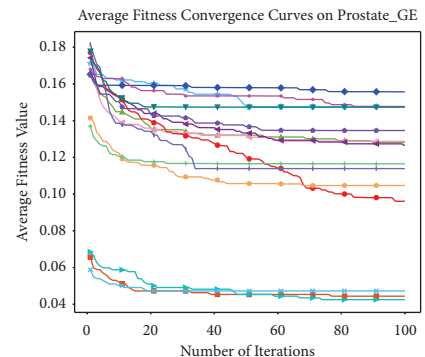
(f)



(g)



(h)



(i)

FIGURE 5: Continued.

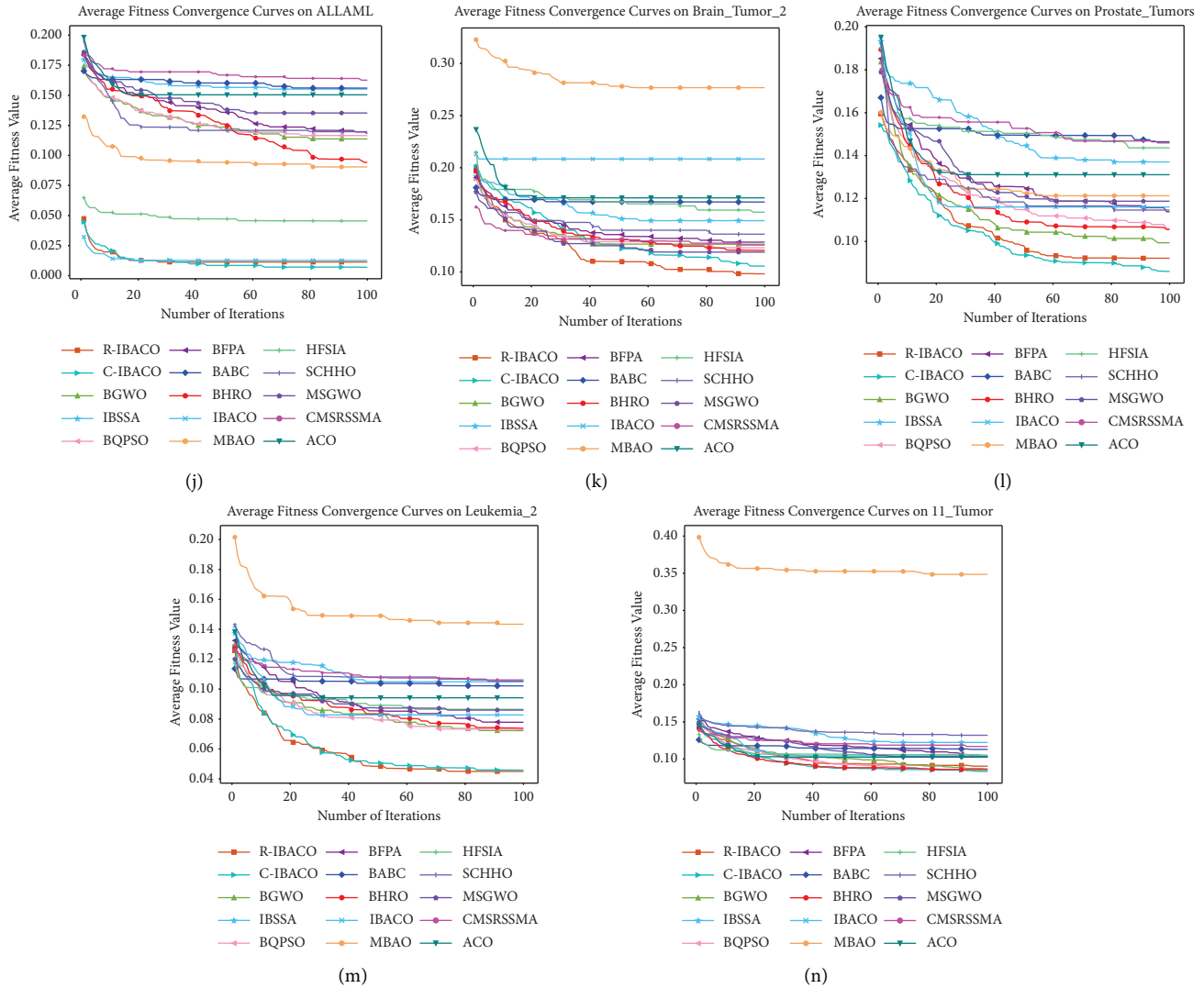


FIGURE 5: Convergence curves of different algorithms on high-dimensional FS datasets. (a) Colon. (b) WarpAR10P. (c) Lung. (d) Lymphoma. (e) GLIOMA. (f) Leukemia\_1. (g) DLBCL. (h) Brain\_Tumor\_1. (i) Prostate\_GE. (j) ALLAML. (k) Brain\_Tumor\_2. (l) Prostate\_Tumors. (m) Leukemia\_2. (n) 11\_Tumor.

Furthermore, the introduction of advanced collaborative hybrid strategies can reduce the optimization efficiency of the algorithm, given that updating the maintainer line and independent subpopulations amplifies the complexity of the model.

The performance improvement of the hybrid algorithms can primarily be attributed to the fact that they leverage the outstanding traits of individual algorithms. As illustrated in Table 7, BHRO exhibited superior classification accuracy, while IBACO surpassed in terms of optimization efficiency and robustness. Another crucial factor contributing to performance enhancement is the hybrid strategy and the improved HF. These elements ensure adequate coverage of the search area, providing the proposed methods with a better opportunity to locate the global optimal solution within the entire search space. This results in the leading performance regarding maximum classification accuracy across all datasets. However,

it is essential to note that the proposed methods still have limitations. For instance, setting the selfing upper limit parameter SC\_Max for HRO can be challenging since it is crucial for the transition between the selfing and renewal stages. Additionally, the inherent randomness in meta-heuristics cannot guarantee the acquisition of the optimal feature subset in a single run.

The index of the most frequently selected features (chosen in > rbin7 out of 10 runs) by C-IBACO is presented in Table 13. The results demonstrate that C-IBACO is capable of selecting a small number of highly discriminative features on most datasets except for warpAR10P, Brain\_Tumor\_1, and Brain\_Tumor\_2, which suggests its potential applications in disease diagnosis and gene expression problems. Notably, the feature subsets composed of high-frequency features on the Colon and Prostate\_Tumors datasets exhibit superior classification performance compared to the results in Table 7. Moreover, the size of these

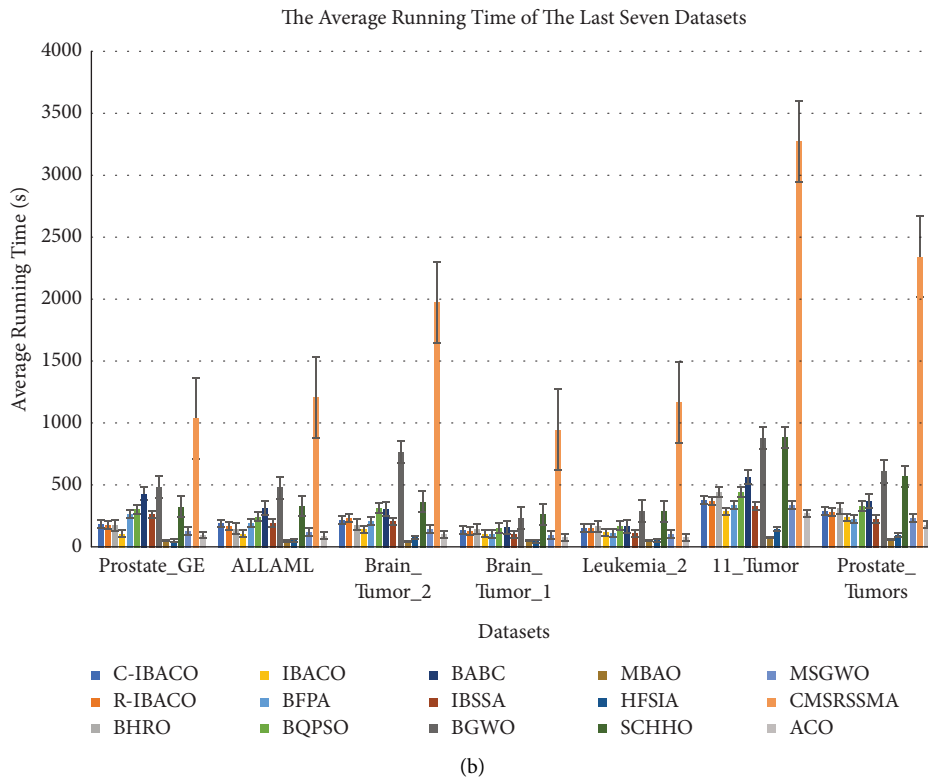
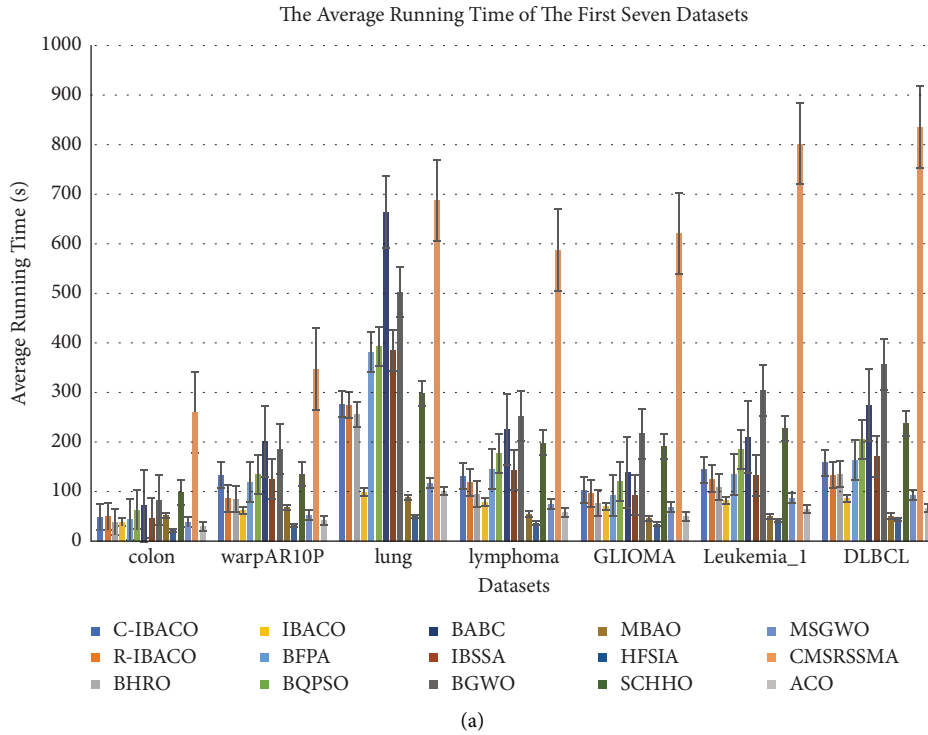


FIGURE 6: Average running time of different algorithms on high-dimension FS datasets. (a) Average running time of different algorithms on the first seven datasets. (b) Average running time of different algorithms on the last seven datasets.

subsets is only 31.12% and 20.5% of the original average selected features, respectively, which suggests that the proposed methods still have the potential for enhanced removal of irrelevant features. With more appropriate and

reasonable parameter settings in the future, the suggested approaches can be applied to other practical problems such as fault detection [56], scheduling problems [57, 58], text classification [59], and sentiment analysis [60].

TABLE 10: Experimental results of Wilcoxon signed-rank test.

Algorithms	C – IBACO <sub>vs</sub>			R – IBACO <sub>vs</sub>		
	$W^+$	$W^-$	$p$	$W^+$	$W^-$	$p$
R-IBACO	72.5	32.5	0.104206	—	—	—
C-IBACO	—	—	—	32.5	72.5	0.907882
BGWO	102	3	0.001049	102	3	0.001049
IBSSA	105	0	0.000549	105	0	0.000549
BQPSO	102	3	0.001049	104	1	0.000683
BFPA	105	0	0.000549	104	1	0.000683
BABC	105	0	0.000549	105	0	0.000549
BHRO	101	4	0.001292	105	0	0.000549
IBACO	105	0	0.000549	105	0	0.000549
MBAO	102.5	2.5	0.001053	104.5	0.5	0.000831
HFSIA	105	0	0.000549	105	0	0.000549
SCHHO	105	0	0.000549	105	0	0.000549
MSGWO	105	0	0.000549	105	0	0.000549
CMSRSSMA	105	0	0.000549	105	0	0.000549
ACO	105	0	0.000549	105	0	0.000549

TABLE 11: Average rankings in terms of classification accuracy using Friedman test.

Algorithm	Average ranking	Final rank	$p$ value
C-IBACO	1.64	1	$2.61E-18 < 0.01$
R-IBACO	1.93	2	
BHRO	5.14	3	
BGWO	6.29	4	
BQPSO	6.39	5	
HFSIA	6.79	6	
IBACO	6.93	7	
BFPA	7.46	8	
MSGWO	9.00	9	
MBAO	10.43	10	
SCHHO	10.93	11	
ACO	11.21	12	
IBSSA	11.71	13	
CMSRSSMA	11.79	14	
BABC	12.36	15	

TABLE 12: Post hoc Holm test (0.05) for R-IBACO and other algorithms.

R – IBACO <sub>vs</sub>	$p$ value	$\alpha$ /rank	Hypothesis
C-IBACO	0.196051	0.050000	Not rejected
BGWO	0.003052	0.012500	Rejected
IBSSA	0.001709	0.003846	Rejected
BQPSO	0.003052	0.010000	Rejected
BFPA	0.001709	0.006250	Rejected
BABC	0.001709	0.003333	Rejected
BHRO	0.003052	0.016667	Rejected
IBACO	0.001709	0.007143	Rejected
MBAO	0.003743	0.005000	Rejected
HFSIA	0.001709	0.008333	Rejected
SCHHO	0.001709	0.004545	Rejected
MSGWO	0.001709	0.005556	Rejected
CMSRSSMA	0.001709	0.003571	Rejected
ACO	0.001709	0.004167	Rejected

TABLE 13: The most frequent features selected by the C-IBACO for each dataset.

Datasets	Index of features	Accuracy
Colon	1507, 44, 1205, 1472, 1347, 95, 764, 1441, 1643, 1561 124, 1581, 1770, 770, 376, 755, 469, 494, 1953, 1670 484, 1052, 1439	98.33
WarpAR10P	720, 2059, 2170, 615, 2167, 2112, 677, 1328, 1423, 1920, 1078, 1912, 1022, 1021, 1506, 1517, 1748, 1556, 1566, 2220, 2235, 308, 43, 57, 59	72.00
Lung	1715, 2232, 2721, 2556, 2185	95.79
Lymphoma	1621, 3752, 2850, 2844, 308, 129, 3749, 1234	92.54
GLIOMA	554, 1913, 312, 4280, 3943, 128, 4157, 1789, 4331	89.27
Leukemia_1	4306, 821, 4199, 2072, 1982, 1380, 2861, 2782, 317, 3246	98.23
DLBCL	5451, 3126, 3463, 4069, 3289, 1105, 2811, 3459	98.75
Brain_Tumor_1	3955, 2508, 1868	74.22
Prostate_GE	125, 4282, 2585, 3380, 716, 4959, 5901, 4283, 497, 173	93.19
ALLAML	3983, 1673, 460, 4365, 6918, 1684, 2640, 206	99.23
Brain_Tumor_2	6995, 7961, 3471, 1810	56.77
Prostate_Tumors	4822, 10132, 6488, 8037, 3199, 8353, 2717, 6837, 7651	97.14
Leukemia_2	4679, 6224, 1867, 1778	90.13
11_Tumor	3945, 9629, 3623, 8016, 10063, 5408, 2796, 3604, 7688, 7057, 33, 5816, 6625, 9431, 10339	90.39

## 6. Conclusion

In this research, two innovative hybrid wrapper-based FS methods that integrate HRO and IBACO are proposed to identify the most informative features in high-dimensional disease diagnosis and gene expression data. The primary objective of hybridization is to enhance the performance of IBACO by harnessing the power of HRO to facilitate the exploration and exploitation of the high-dimensional search space. By combining the superior search efficiency and robustness of IBACO and the excellent performance of HRO in searching the global optima, the proposed hybrid methods manifest enhanced FS capabilities. Moreover, IBACO attempts to boost performance through a problem-oriented assignment strategy that employs the correlation of the knee point feature, enabling the algorithm to exploit valuable latent information in the features. This strategy is also integrated into HRO to compensate for the absence of update to the maintainer line.

Two distinct forms of hybridization are presented in this study: R-IBACO and C-IBACO. In R-IBACO, IBACO plays a critical role in updating the maintainer line of HRO, with the best solution derived from HRO subsequently used to update PD at each iteration, while in C-IBACO, the subpopulations of HRO and IBACO perform the update process independently, and the local search results are shared to update PD and maintainer line after each iteration. In the methods proposed, the KNN algorithm functions as the classifier, and RF is employed to calculate the feature importance required for assigning HF. The proposed methods were evaluated on fourteen well-known biomedical datasets. Their performance was benchmarked against thirteen other algorithms, including single standard algorithms that comprise them, as well as basic and advanced metaheuristic-

based wrapper FS methods. The experimental results indicate that the proposed methods outperform the other techniques in terms of both the number of selected features and classification accuracy on most datasets. Furthermore, the statistical results of the Wilcoxon signed-rank test and Friedman test reveal that the proposed methods achieved the top rank in terms of classification accuracy, which corroborates their effectiveness as practical strategies for selecting the most representative features related to diseases.

Although the suggested approaches effectively enhance the exploration and exploitation capabilities of a single algorithm, it is important to note that optimization efficiency might decline as the model complexity increases. Moreover, optimal performance may not be achievable in certain specific scenarios. Consequently, future research should prioritize the consideration of more efficient hybrid strategies that can ensure the generalization capabilities of the algorithms across problems of varying scales and dimensions. One direction for further research involves exploring more efficient parameter settings and integrated optimization of algorithm and classifier parameters, which will help determine synergistic settings that maximize the overall performance. Additionally, it would be beneficial to explore different transfer functions and advanced classifiers, as these components have the potential to further bolster the performance of the algorithms in a variety of optimization scenarios.

## Data Availability

The datasets analyzed during the current study are available in the following websites: <https://jundongli.github.io/scikit-feature/datasets.html> and <https://ckzixf.github.io/dataset.html>.



## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant no. 42201464 and the Wuhan Science and Technology Bureau Knowledge Innovation Dawning under grant no. 2022010801020270.

## References

- [1] A. Chaudhuri and T. P. Sahu, "A hybrid feature selection method based on binary jaya algorithm for micro-array data classification," *Computers and Electrical Engineering*, vol. 90, Article ID 106963, 2021.
- [2] E. Pashaei and E. Pashaei, "Gene selection using hybrid dragonfly black hole algorithm: a case study on rna-seq covid-19 data," *Analytical Biochemistry*, vol. 627, Article ID 114242, 2021.
- [3] J. T. Pintas, L. A. F. Fernandes, and A. C. B. Garcia, "Feature selection methods for text classification: a systematic literature review," *Artificial Intelligence Review*, vol. 54, no. 8, pp. 6149–6200, 2021.
- [4] S. Chakraborty, A. K. Saha, S. Nama, and S. Debnath, "Covid-19 x-ray image segmentation by modified whale optimization algorithm with population reduction," *Computers in Biology and Medicine*, vol. 139, Article ID 104984, 2021.
- [5] M. Canayaz, "Mh-covidnet: diagnosis of covid-19 using deep neural networks and meta-heuristic-based feature selection on x-ray images," *Biomedical Signal Processing and Control*, vol. 64, Article ID 102257, 2021.
- [6] S. K. S. Tyagi and Q. Boyang, "An intelligent internet-of-things-aided financial crisis prediction model in fintech," *IEEE Internet of Things Journal*, vol. 10, no. 3, pp. 2183–2193, 2023.
- [7] M.-F. Leung and J. Wang, "Cardinality-constrained portfolio selection based on collaborative neurodynamic optimization," *Neural Networks*, vol. 145, no. 68–79, pp. 68–79, 2022.
- [8] H. Chan, M. Yang, H. Wang et al., "Assessing gait patterns of healthy adults climbing stairs employing machine learning techniques," *International Journal of Intelligent Systems*, vol. 28, no. 3, pp. 257–270, 2013.
- [9] P. Agrawal, H. F. Abutarboush, T. Ganesh, and A. W. Mohamed, "Metaheuristic algorithms on feature selection: a survey of one decade of research (2009-2019)," *IEEE Access*, vol. 9, pp. 26766–26791, 2021.
- [10] L. Y. Yab, N. Wahid, and R. A. Hamid, "A meta-analysis survey on the usage of meta-heuristic algorithms for feature selection on high-dimensional datasets," *IEEE Access*, vol. 10, pp. 122832–122856, 2022.
- [11] H. Jia, W. Zhang, R. Zheng, S. Wang, X. Leng, and N. Cao, "Ensemble mutation slime mould algorithm with restart mechanism for feature selection," *International Journal of Intelligent Systems*, vol. 37, no. 3, pp. 2335–2370, 2022.
- [12] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, "Slime mould algorithm: a new method for stochastic optimization," *Future Generation Computer Systems*, vol. 111, pp. 300–323, 2020.
- [13] C. Qu, L. Zhang, J. Li et al., "Improving feature selection performance for classification of gene expression data using Harris Hawks optimizer with variable neighborhood learning," *Briefings in Bioinformatics*, vol. 22, no. 5, Article ID bbab097, 2021.
- [14] K. K. Ghosh, R. Guha, S. K. Bera, N. Kumar, and R. Sarkar, "S-shaped versus v-shaped transfer functions for binary manta ray foraging optimization in feature selection problem," *Neural Computing and Applications*, vol. 33, no. 17, pp. 11027–11041, 2021.
- [15] W. Zhao, Z. Zhang, and L. Wang, "Manta ray foraging optimization: an effective bio-inspired optimizer for engineering applications," *Engineering Applications of Artificial Intelligence*, vol. 87, Article ID 103300, 2020.
- [16] A. A. Ewees, R. R. Mostafa, R. M. Ghoniem, and M. A. Gaheen, "Improved seagull optimization algorithm using lévy flight and mutation operator for feature selection," *Neural Computing and Applications*, vol. 34, no. 10, pp. 7437–7472, 2022.
- [17] E. Pashaei, "Mutation-based binary aquila optimizer for gene selection in cancer classification," *Computational Biology and Chemistry*, vol. 101, Article ID 107767, 2022.
- [18] M. A. Awadallah, M. A. Al-Betar, M. S. Braik, A. I. Hammouri, I. A. Doush, and R. A. Zitar, "An enhanced binary rat swarm optimizer based on local-best concepts of pso and collaborative crossover operators for feature selection," *Computers in Biology and Medicine*, vol. 147, Article ID 105675, 2022.
- [19] M. Mafarja, T. Thaher, J. Too et al., "An efficient high-dimensional feature selection approach driven by enhanced multi-strategy grey wolf optimizer for biological data classification," *Neural Computing and Applications*, vol. 35, no. 2, pp. 1749–1775, 2023.
- [20] E. Pashaei and E. Pashaei, "Hybrid binary arithmetic optimization algorithm with simulated annealing for feature selection in high-dimensional biomedical data," *The Journal of Supercomputing*, vol. 78, no. 13, pp. 15598–15637, 2022.
- [21] P. Stephan, T. Stephan, R. Kannan, and A. Abraham, "A hybrid artificial bee colony with whale optimization algorithm for improved breast cancer diagnosis," *Neural Computing and Applications*, vol. 33, no. 20, pp. 13667–13691, 2021.
- [22] R. Al-Wajih, S. J. Abdulkadir, N. Aziz, Q. Al-Tashi, and N. Talpur, "Hybrid binary grey wolf with Harris hawks optimizer for feature selection," *IEEE Access*, vol. 9, pp. 31662–31677, 2021.
- [23] R. Bandyopadhyay, A. Basu, E. Cuevas, and R. Sarkar, "Harris hawks optimisation with simulated annealing as a deep feature selection method for screening of covid-19 ct-scans," *Applied Soft Computing*, vol. 111, Article ID 107698, 2021.
- [24] K. Hussain, N. Neggaz, W. Zhu, and E. H. Houssein, "An efficient hybrid sine-cosine Harris hawks optimization for low and high-dimensional feature selection," *Expert Systems with Applications*, vol. 176, Article ID 114778, 2021.
- [25] Y. Zhu, W. Li, and T. Li, "A hybrid artificial immune optimization for high-dimensional feature selection," *Knowledge-Based Systems*, vol. 260, no. 110, Article ID 110111, 2023.
- [26] E. Pashaei and E. Pashaei, "Hybrid binary coot algorithm with simulated annealing for feature selection in high-dimensional microarray data," *Neural Computing and Applications*, vol. 35, no. 1, pp. 353–374, 2023.
- [27] Z. Ye, L. Ma, and H. Chen, "A hybrid rice optimization algorithm," in *Proceedings of the 2016 11th International Conference on Computer Science and Education (ICCSE)*, pp. 169–174, Nagoya, Japan, August 2016.
- [28] Z. Shu, Z. Ye, X. Zong et al., "A modified hybrid rice optimization algorithm for solving 0-1 knapsack problem," *Applied Intelligence*, vol. 52, no. 5, pp. 5751–5769, 2022.

- [29] Z. Ye, W. Cai, S. Liu, K. Liu, M. Wang, and W. Zhou, "A band selection approach for hyperspectral image based on a modified hybrid rice optimization algorithm," *Symmetry*, vol. 14, no. 7, p. 1293, 2022.
- [30] Z. Wang, S. Gao, Y. Zhang, and L. Guo, "Symmetric uncertainty-incorporated probabilistic sequence-based ant colony optimization for feature selection in classification," *Knowledge-Based Systems*, vol. 256, p. 2022, Article ID 109874, 2022.
- [31] M. Paniri, M. B. Dowlatshahi, and H. Nezamabadi-pour, "Mlaco: a multi-label feature selection algorithm based on ant colony optimization," *Knowledge-Based Systems*, vol. 192, Article ID 105285, 2020.
- [32] F. Karimi, M. B. Dowlatshahi, and A. Hashemi, "Semiaco: a semi-supervised feature selection based on ant colony optimization," *Expert Systems with Applications*, vol. 214, Article ID 119130, 2023.
- [33] E.-G. Talbi, "A taxonomy of hybrid metaheuristics," *Journal of Heuristics*, vol. 8, no. 5, pp. 541–564, 2002.
- [34] P. Hu, J.-S. Pan, and S.-C. Chu, "Improved binary grey wolf optimizer and its application for feature selection," *Knowledge-Based Systems*, vol. 195, Article ID 105746, 2020.
- [35] M. Abdel-Basset, K. M. Sallam, R. Mohamed, I. Elgendi, K. Munasinghe, and O. M. Elkomy, "An improved binary grey-wolf optimizer with simulated annealing for feature selection," *IEEE Access*, vol. 9, pp. 139792–139822, 2021.
- [36] X.-F. Song, Y. Zhang, D.-W. Gong, and X.-Z. Gao, "A fast hybrid feature selection based on correlation-guided clustering and particle swarm optimization for high-dimensional data," *IEEE Transactions on Cybernetics*, vol. 52, no. 9, pp. 9573–9586, 2022.
- [37] A. W. Mohamed, A. A. Hadi, and A. K. Mohamed, "Gaining-sharing knowledge based algorithm for solving optimization problems: a novel nature-inspired algorithm," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 7, pp. 1501–1529, 2020.
- [38] P. Agrawal, T. Ganesh, D. Oliva, and A. W. Mohamed, "S-shaped and v-shaped gaining-sharing knowledge-based algorithm for feature selection," *Applied Intelligence*, vol. 52, no. 1, pp. 81–112, 2022.
- [39] P. Agrawal, T. Ganesh, and A. W. Mohamed, "Chaotic gaining sharing knowledge-based optimization algorithm: an improved metaheuristic algorithm for feature selection," *Soft Computing*, vol. 25, no. 14, pp. 9505–9528, 2021.
- [40] K. Hanbay, "A new standard error based artificial bee colony algorithm and its applications in feature selection," *Journal of King Saud University Computer and Information Sciences*, vol. 34, no. 7, pp. 4554–4567, 2022.
- [41] M. Tubishat, N. Idris, L. Shuib, M. A. Abushariah, S. Mirjalili, and S. Mirjalili, "Improved salp swarm algorithm based on opposition based learning and novel local search algorithm for feature selection," *Expert Systems with Applications*, vol. 145, Article ID 113122, 2020.
- [42] D. O. Owuor, T. Runkler, A. Laurent, J. O. Orero, and E. O. Menya, "Ant colony optimization for mining gradual patterns," *International Journal of Machine Learning and Cybernetics*, vol. 12, no. 10, pp. 2989–3009, 2021.
- [43] X. Zhang, Z.-H. Zhan, W. Fang, P. Qian, and J. Zhang, "Multipopulation ant colony system with knowledge-based local searches for multiobjective supply chain configuration," *IEEE Transactions on Evolutionary Computation*, vol. 26, no. 3, pp. 512–526, 2022.
- [44] D. Zhao, L. Liu, F. Yu et al., "Ant colony optimization with horizontal and vertical crossover search: fundamental visions for multi-threshold image segmentation," *Expert Systems with Applications*, vol. 167, Article ID 114122, 2021.
- [45] M. Rostami, K. Berahmand, E. Nasiri, and S. Forouzandeh, "Review of swarm intelligence-based feature selection methods," *Engineering Applications of Artificial Intelligence*, vol. 100, Article ID 104210, 2021.
- [46] Z. Manbari, F. Akhlaghian Tab, and C. Salavati, "Fast unsupervised feature selection based on the improved binary ant system and mutation strategy," *Neural Computing and Applications*, vol. 31, no. 9, pp. 4963–4982, 2019.
- [47] M. Ghosh, R. Guha, R. Sarker, and A. Abraham, "A wrapper-filter feature selection technique based on ant colony optimization," *Neural Computing and Applications*, vol. 32, no. 12, pp. 7839–7857, 2020.
- [48] M. Paniri, M. B. Dowlatshahi, and H. Nezamabadi-pour, "Anttd: ant colony optimization plus temporal difference reinforcement learning for multi-label feature selection," *Swarm and Evolutionary Computation*, vol. 64, Article ID 100892, 2021.
- [49] A. Hashemi, M. Joodaki, N. Z. Joodaki, and M. B. Dowlatshahi, "Ant colony optimization equipped with an ensemble of heuristics through multi-criteria decision making: a case study in ensemble feature selection," *Applied Soft Computing*, vol. 124, Article ID 109046, 2022.
- [50] Y. Wan, M. Wang, Z. Ye, and X. Lai, "A feature selection method based on modified binary coded ant colony optimization algorithm," *Applied Soft Computing*, vol. 49, pp. 248–258, 2016.
- [51] M. Li, X. Ren, Y. Wang, W. Qin, and Y. Liu, "Advanced antlion optimizer with discrete ant behavior for feature selection," *IEICE Transactions on Info and Systems*, vol. 103, no. 12, pp. 2717–2720, 2020.
- [52] W. Ma, X. Zhou, H. Zhu, L. Li, and L. Jiao, "A two-stage hybrid ant colony optimization for high-dimensional feature selection," *Pattern Recognition*, vol. 116, Article ID 107933, 2021.
- [53] A. Thakkar and R. Lohiya, "Role of swarm and evolutionary algorithms for intrusion detection system: a survey," *Swarm and Evolutionary Computation*, vol. 53, Article ID 100631, 2020.
- [54] K. Chen, B. Xue, M. Zhang, and F. Zhou, "An evolutionary multitasking-based feature selection method for high-dimensional classification," *IEEE Transactions on Cybernetics*, vol. 52, no. 7, pp. 7172–7186, 2022.
- [55] J. Derrac, S. García, D. Molina, and F. Herrera, "A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms," *Swarm and Evolutionary Computation*, vol. 1, no. 1, pp. 3–18, 2011.
- [56] E. S. M. El-kenawy, F. Albalawi, S. A. Ward et al., "Feature selection and classification of transformer faults based on novel meta-heuristic algorithm," *Mathematics*, vol. 10, no. 17, p. 3144, 2022.

- [57] J. Mou, P. Duan, L. Gao, X. Liu, and J. Li, "An effective hybrid collaborative algorithm for energy-efficient distributed permutation flow-shop inverse scheduling," *Future Generation Computer Systems*, vol. 128, pp. 521–537, 2022.
- [58] F. Zhao, X. He, and L. Wang, "A two-stage cooperative evolutionary algorithm with problem-specific knowledge for energy-efficient scheduling of no-wait flow-shop problem," *IEEE Transactions on Cybernetics*, vol. 51, no. 11, pp. 5291–5303, 2021.
- [59] K. Thirumoorthy and K. Muneeswaran, "Feature selection using hybrid poor and rich optimization algorithm for text classification," *Pattern Recognition Letters*, vol. 147, pp. 63–70, 2021.
- [60] D. K. Jain, P. Boyapati, J. Venkatesh, and M. Prakash, "An intelligent cognitive-inspired computing with big data analytics framework for sentiment analysis and classification," *Information Processing and Management*, vol. 59, no. 1, Article ID 102758, 2022.