Bulletin of Electrical Engineering and Informatics
Vol. 15, No. 3, June 2026, pp. 2345~2359
ISSN: 2302-9285, DOI: 10.11591/eei.v15i3.10991

Hybrid metaheuristic algorithms for feature selection in classification: a systematic literature review

Manal Othman[1,2], Ku Ruhana Ku-Mahamud[1]
[1]School of Computing, College of Arts and Sciences, Universiti Utara Malaysia, Sintok, Malaysia
[2]Department of Computer Science, Faculty of Applied Science, Taiz University, Taiz, Yemen

Article Info

Article history:

Received Jul 9, 2025
Revised Mar 14, 2026
Accepted Mar 31, 2026

Keywords:

Binary
Classification
Feature selection
Hybridization
Metaheuristic

ABSTRACT

Feature selection (FS) is a popular technique for improving machine learning (ML) model's effectiveness by eliminating irrelevant and redundant features. It is challenging because of the intricate relationship between features and large search space. Recent studies have focused on using hybrid metaheuristics to solve FS problem. This systematic literature review (SLR) is performed on three significant databases that explores recent studies from 2019 to 2024 that used hybrid metaheuristics for FS in classification. This paper aims to understand the existing hybrid algorithms, hybridization goal, hybridization type, and application domains. Moreover, crucial parameters, fitness and transfer functions, initial population method, traditional FS approach, classification algorithm, evaluation criteria, and statistical test are investigated in this paper. The qualitative findings derived from the systematic review encompassed 646 publications, systematically categorized based on predefined inclusion and exclusion criteria. Consequently, 35 papers were analyzed to develop new insights in the domain of FS in classification, focusing on single-objective metaheuristics. Hybrid metaheuristics surpass the efficacy of their individual components in enhancing algorithmic performance to attain optimal or near-optimal solutions. The limitations of hybrid metaheuristics and research gaps are identified for scholars interested in developing metaheuristic algorithms for FS.

Corresponding Author:

Manal Othman
School of Computing, College of Arts and Sciences, Universiti Utara Malaysia
Sintok, Kedah, Malaysia
Email: manal_m_othman@ahsgs.uum.edu.my

1. INTRODUCTION

Classification is a fundamental task in machine learning (ML), wherein ML algorithms are utilized to extract knowledge from a set of given data [1]. It is a supervised learning that aims to develop a decision model that searches for patterns in training set data to predict the category for new unknown data. Classification is a highly studied subject in data mining and ML, with diverse applications across various fields. Classification techniques have been extensively researched and evolved over the years to enhance accuracy, efficiency, and interpretability. Nonetheless, all these algorithms reflect problems with various levels of complexity [2]. The growing number of datasets presents challenges in managing classifier implementation, necessitating the separation of desirable features from irrelevant and redundant ones. The procedure of choosing the relevant features is referred to as feature selection (FS).

FS aims to reduce features by removing redundant and irrelevant attributes from certain datasets [3]. Real-world applications commonly utilize numerous features for data representation, some of which could be redundant and can be eliminated. FS has been employed in many ML and data mining problems, especially with high-dimensional data. It is essential for four main reasons. The first is to reduce overfitting. By choosing the most relevant features, the model can improve generalization to new data. Secondly, reducing the complexity of a model in terms of computing will increase its speed and efficiency [4]. The third reason is to improve the interpretability of the model [5]. Finally, FS helps to mitigate the “curse of dimensionality” by reducing the number of features used in datasets [6]. Overall, FS is a powerful classification tool that can enhance the classification model's efficiency, interpretability and effectiveness.

FS can be reviewed using three primary techniques: filter, wrapper, and embedded. The filter technique is independent of the ML algorithm that is being utilized. It has the merit of being fast, computationally inexpensive, and able to scale up to large datasets [7]. Despite these advantages, it provides a poor-quality solution. The wrapper technique selects a subset of features based on the model's predictive power. This technique selects features by iterating through multiple combinations of features and evaluating their performance using a chosen ML algorithm [8]. Embedded techniques select features during training rather than features before training the model. The benefits of embedded techniques are that they can be computationally effective since FS is incorporated into the training process. Additionally, they can be used to build more interpretable models since they only include the most relevant features. However, they may not perform similarly to wrapper methods in some cases, mainly when many features exist.

Finding the most practical features can be complex and computationally demanding. Despite various suggested approaches, most of them still encounter difficulties, such as becoming stuck in local optimal solutions and high computational expense problems mainly because of the large search area. As a result, it is necessary to possess a global search mechanism that effectively manages FS tasks. The most effective techniques for handling FS are metaheuristic algorithms due to their robustness, intelligibility, and efficiency in resolving complex optimization issues [9]. Several metaheuristic methods have previously been employed to solve the FS problems. Scholars are currently encountering a challenge in effectively applying and providing accurate recommendations for modern metaheuristics because every metaheuristic algorithm has pros and cons [10]. Most recent studies have focused on using hybrid metaheuristics to enhance the algorithm's performance, adaptability, and efficiency by integrating strengths from different metaheuristic algorithms. These hybridizations increase the probability of finding high-quality feature subsets. The selection of hybridization depends on the characteristics of the optimization issue and the capabilities of the individual algorithms involved. While metaheuristic algorithms have been widely used for FS, the application of hybrid metaheuristics remains insufficiently investigated. Specifically, there is a lack of comprehensive evaluation of these hybrid methods and limited discussion on how hybridization is done. This paper intends to provide insightful analysis and direction for further research and development in the field of hybrid metaheuristics for FS.

A comprehensive survey on using metaheuristic algorithms for the FS problem was conducted by Dokeroglu *et al.* [11]. The authors focus only on the most remarkable recent 22 metaheuristics that have been developed in the last two decades, from harmony search (HS) in 2001 to Harris Hawks optimization (HHO) and Bayesian optimization algorithm (BOA) in 2019. Agrawal *et al.* [12] reviewed the binary metaheuristics developed for FS from 2009 to 2019. The discussion focused on the performance of algorithms and analyzed the classifiers, datasets, transfer function, and evaluation metrics, as well as identified the challenges and emphasized areas where further research is needed. Furthermore, a case study on UCI datasets was provided. This review paper offers a comprehensive analysis of studies that employ metaheuristics for FS across various domains. While it examines different metaheuristics and their effectiveness in FS, it does not specifically focus on the hybridization of multiple metaheuristic techniques.

A systematic literature review (SLR) on FS-based metaheuristics in text classification has been conducted by Abiodun *et al.* [13] and Alyasiri *et al.* [14]. The review addressed several key questions, including datasets, classifiers, sub-field of metaheuristics, and its impact on the accuracy of text classification during the period from 2015 to 2021 [13]. On the other hand, a total of 37 papers from 2016 to 2021 that employed metaheuristic algorithms to solve FS problem were selected by Alyasiri *et al*. [14]. It significantly contributes to the current body of knowledge to comprehensively understanding the methods and their advantages and disadvantages in terms of decreasing the computational resources needed to execute text classification. However, this SLR examines FS techniques-based metaheuristic algorithms for exclusively text classification in the English language text.

A comprehensive analysis of metaheuristics used for FS issues and their categorical list was presented by Sharma and Kaur [15]. The authors highlighted the performance and role of binary and chaotic variants of different nature-inspired metaheuristic algorithms in 10 major research domains. Metaheuristic algorithms for different disease predictions are also illustrated. However, the study did not analyse the transfer function, fitness employed or the parameter values that provide an important basis for the improvement of metaheuristics. A review of swarm intelligence algorithms for FS and their categorization is performed in Rostami *et al.* [16]. The authors focused on the strengths and drawbacks of 11 swarm intelligence-based FS algorithms studies that were evaluated. The algorithms are ant colony optimization (ACO), particle swarm optimization (PSO), differential evolution (DE), artificial bee colony (ABC), salp swarm algorithm (SSA), whale optimization algorithm (WOA), grey wolf optimuzation (GWO), bat algorithm (BA), gravitational search algorithm (GSA), firefly algorithm (FA), and coati optimization algorithm (COA) with six medical datasets and three classifiers. The review scope is considered narrow and lacks hybridization analysis.

Akinola *et al.* [17] provide a review of the literature (from 2000 to 2022) on applying metaheuristics to multi-class FS problems, which enables classifiers to choose optimal or nearly optimal features with incredible speed and accuracy. The authors emphasize the review of metaheuristics and details of their six categories (evolutionary, swarm, human, physics, system, and bio)-based algorithms, as well as the categorization of the FS approaches. Pham and Raahemi [18] conducted a SLR of studies published between 2020 and 2023 that employed bio-inspired algorithms for FS across various domains, including healthcare, text classification, image processing, and cybersecurity. It detailed the categories of bio-inspired FS algorithms, bio-inspired algorithms with frequency of occurrence and improvement techniques. Piri *et al.* [10] provide a critical assessment of the current hybrid FS based metaheuristics considering 35 studies published between 2009 and 2022. The review presents a thorough picture of the metaheuristic algorithms utilized in hybridization, along with their associated classifiers and datasets. Table 1 provides a summary of the details and limitations of these studies. These reviews serve as good resources for FS-based metaheuristics and offer valuable information and discussion about the challenges and future directions for future research. However, a significant limitation in the reviews is that most reviews fail to provide a comprehensive examination of hybridization, which is one of the most significant recent advancements in the field of optimization, and none of them distinguish between low-level (cooperative) and high-level (integrative) hybrids, where the performance and computational efficiency is directly affected by the structural design of hybridization. Accordingly, this SLR focuses specifically on hybrid metaheuristics and the specifics of their hybridization methodologies.

Table 1. Summary of previous review papers on metaheuristics for FS

Reference	Period	Databases/scope	Specific focus	Limitation
Dokeroglu *et al.* [11]	2001-2021	Google Scholar. Multiple domains.	Survey of 22 metaheuristics for FS.	Lacks hybridization analysis. Focus on specific metaheuristic algorithms.
Agrawal *et al.* [12]	2009-2019	- Multiple domains.	Survey of metaheuristics for FS.	Lacks hybridization analysis.
Abiodun *et al.* [13]	2015–2021	EEE Xplore, Science Direct, ACM Digital Library, Scopus, Elsevier, Springer, EBSCO Host, Taylor and Francis, Research Gate, and Google Scholar. Text classification.	SLR for metaheuristics for FS in text classification.	Lacks hybridization analysis. Narrow domain.
Alyasiri *et al.* [14]	2016-2021	IEEE, Scopus, Research Gate, Springer, Google Scholar, Science Direct, and Taylor & Francis. Text classification.	SLR for metaheuristics for FS in text classification.	Lacks hybridization analysis. Narrow domain.
Sharma and Kaur [15]	2000-2019	- Multiple domains.	SLR for nature-inspired metaheuristics for FS.	Lacks hybridization analysis.
Rostami *et al.* [16]	2008-2020	- Multiple domains.	Review of metaheuristics for FS.	Lacks hybridization analysis. Narrow discussion.
Akinola *et al.* [17]	2000-2022	Scopus, Elsevier, IEEE Xplore, Springer link, Research gate, Google Scholar, and Web of Science. Multiple domains.	Systematic survey on metaheuristics for multiclass FS.	Lacks hybridization analysis. Limited to multiclass FS.
Pham and Raahemi [18]	2020-2023	ACM, IEEE Xplore, Scopus, Web of Science, and ScienceDirect. Multiple domains.	SLR on bio-inspired metaheuristics for FS.	Lacks hybridization analysis Limited to bio-inspired (no focus on bio and non-bio inspired hybrid).
Piri *et al.* [10]	2009-2022	Springer, IEEE Xplore, ScienceDirect, Web of Science, and ACM. Multiple domains.	Review of hybrid metaheuristics for FS.	No deep analysis of hybrid metaheuristics and their hybridization structures.

This SLR presents a complete analysis and synthesis of six years (2019–2024) of studies that use hybrid metaheuristic algorithms to solve the FS problem in classification tasks. This attempt will yield advantages for forthcoming investigations in this domain and help determine a suitable modelling methodology for classification. The essential contributions of this review are as follows:

a. Provide a critical review of 35 papers that have used hybrid metaheuristic algorithms to solve FS in classification. In addition, the result of an analysis of the fundamental information regarding the purpose of hybridization, datasets utilized, and application domains in the selected studies are provided.

b. Identify studies that have used statistical tests and highlight the most common evaluation metrics used to assess the performance of algorithms, analyse and summarize the type of hybridization used.
c. Conduct an analysis and summarize of the FS approach, ML techniques used to create fitness functions, and methods used to initialize the population.
d. Investigate the employed transfer function, fitness function and values of crucial parameters. Address the challenges and potential future directions of the metaheuristics-FS field, which might serve as a reference framework for future studies.

2. METHOD

This section presents the methodology used to choose and examine the literature on hybrid metaheuristic algorithms for the FS problem. This SLR was conducted using the established guidelines in Kitchenham *et al.* [19].

2.1. Scope of discussion

This SLR focuses mostly on using single objective hybrid metaheuristic algorithms for FS in classification. The aim is to investigate algorithms, datasets, application domains, transfer function, fitness function, classifiers, FS approaches, population initialization method, and values of crucial parameters (population size, iterations, and run times). In addition, important details of hybridization include the type (parallel, sequential, low-level, and high-level), research trends and gaps.

2.2. Research questions

The following questions are attempted to be addressed by this review.
a. Which hybridized metaheuristic algorithms are used for the FS problem? What is the purpose of hybridization? What are their application domains?
b. What type of hybridization is used? What are the evaluation metrics? What statistical tests are used?
c. What are the FS techniques applied with metaheuristic to achieve good classification accuracy and minimum number of features? What classifiers are used? What initial population methods are used?
d. What are the crucial/optimal parameters values? Which transfer and fitness functions are used?

2.3. Search strategy and data source

One of the most important tasks in creating an SLR is figuring out what sufficient and correct keywords to use. Initially, we formulated the query using basic keywords and searched to find the most important publications in this field. These queries were designed to extract the primary keywords and important databases from those publications. Subsequently, the primary keywords from the articles are chosen, and the queries were initially conducted in December 2023 and updated in November 2024 on the three databases in the fields of science and engineering including Scopus, Web of Science, and ACM digital library databases at the University Utara Malaysia library (electronic resources). These sources were chosen for their comprehensive coverage of leading journals and conferences in computer science, artificial intelligence (AI), and optimization. The results studies from the query exported as .CSV file to examine the duplication. The search is within the article's title, abstract, and keywords to locate papers that are relevant to this subject. The first author identified 35 studies that fully met inclusion criteria, while the second author independently reviewed and verified all selections, achieving a 100% agreement rate. Full-text papers were exported to the Endnote library for deep analysis. Data from the included studies were systematically extracted, including details on hybridized metaheuristics, application domains, type and aim of hybridization, FS approach, classifiers, evaluation metrics, statistical tests, crucial parameters values, initial population methods, transfer and fitness functions and reported results. The second author double-checked the extracted data to ensure accuracy and consistency. We manually checked the references and citations of the included studies to find more relevant literature (snowballing), and only those that met our inclusion criteria were added to the review. Studies beyond this scope was excluded, and snowballing was not conducted on these studies. The search strategy, including databases, queries, period (2019-2024), inclusion/exclusion criteria, and preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow diagram were documented to guarantee transparency and reproducibility.

2.4. Search query

The search query was carefully designed to ensure all relevant studies were included. Multiple variants of the keyword “metaheuristic” (e.g., metaheuristic, meta-heuristic, metaheuristics, meta-heuristics) were included to capture several spellings, explicit linking of “feature selection” with “classification” to limit the scope to the target domain, while Boolean operators (AND/OR) were employed to achieve a balance between inclusion and specificity. Furthermore, the constraint for the term "hybrid" guaranteed that only studies specifically focused on hybrid metaheuristics for FS were retrieved. ("Hybrid" AND ("metaheuristic" OR "meta-heuristic" OR "metaheuristics" OR "meta-heuristics") AND "for" AND "feature selection" AND ("on" OR "in") AND "classification").

2.5. Inclusion and exclusion criteria

To find solutions to the research questions, inclusion criteria were established to categorize papers retrieved from scientific databases. We evaluate the quality of the primary and final studies for this SLR based on the predefined inclusion and exclusion criteria, as shown in Table 2.

Table 2. Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
1- Papers on single objective hybrid metaheuristics for FS for classification are included.	Papers that do not satisfy the inclusion criteria.
	1- Hybrid metaheuristics for non-classification tasks.
2- Only related peer-reviewed works utilized the hybridization for metaheuristics.	2- Hybridization was used in FS approaches (e.g., filter-wrapper) rather than metaheuristics.
3- Papers that were published in the period (2019-2024)	3- Papers prior to 2019 are not included.
4- Only English-language papers were included.	4- Paper that is inaccessible
	5- Papers written in other languages (non-English).

The PRISMA diagram [20], presented in Figure 1, illustrates the procedures involved in choosing papers. The primary search included 646 papers (431 research papers retrieved from Scopus, 61 research papers retrieved from ACM digital library, and 154 research papers retrieved from Web of Science). There were 455 studies remaining to screen after removing 191 duplicated papers. There were still 432 studies that needed to be screened after removing 23 review papers. Out of those, 47 papers were excluded because they do not reflect current state-of-the-art advancements (published before 2019), and 350 papers were eliminated as they did not meet the inclusion criteria. These included studies with unavailable full texts, non-metaheuristic hybridization (hybridization used was in FS or ML approaches), and non-classification applications, along with 8 papers addressing multi-objective optimization. A total of 35 publications were ultimately chosen in-depth examinations.

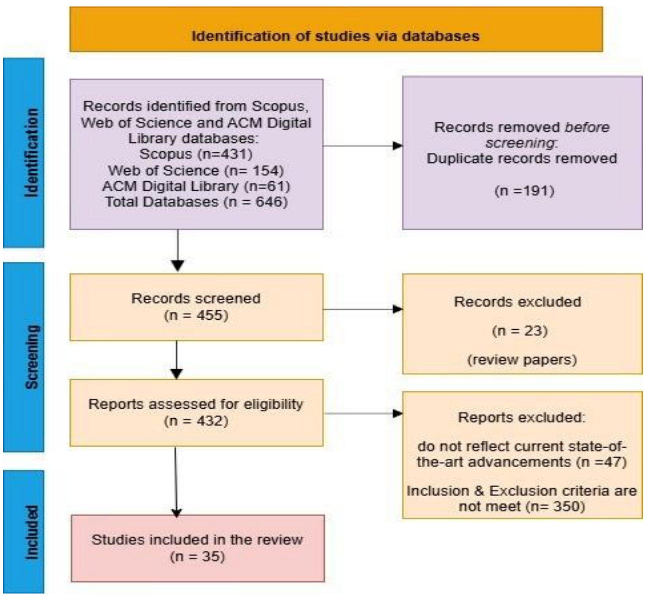


Figure 1. PRISMA flow diagram

2.6. Data extraction

This stage involves extracting pertinent data and information from the chosen papers. The data that has been extracted is subsequently analyzed to address our study questions. We conducted an analysis of this data to categorize studies that were comparable to one another. Extracting the essential information from each study enables identification of significant advancements and patterns in the research, assessing any existing gaps, and establishing directions for future work. As a result, the eligibility of studies was evaluated based on the information required in research questions.

3. DATA SYNTHESIS AND RESULTS

This section examines selected papers Table 3 (in Appendix) [3], [4], [7], [21]–[49], on the application of hybrid metaheuristic algorithms for FS in classification.

RQ1: Which hybridized metaheuristic algorithms are used for the FS problem? What is the purpose of hybridization? What are their application domains?

Hybrid metaheuristics enable efficient exploration of the feature space and ultimately aid in identifying an optimal subset of features and improve the predictive. We can conclude the benefits of hybridization in studies reviewed in this paper as follows:

- It assists in enhancing the effectiveness of the original algorithm.
- It assists in resolving the problems of premature convergence and the local optimal trap.
- It assists in striking a balance between exploration and exploitation processes.

Table 3 shows the studies selected for this review that used different hybridization algorithms and Figure 2 shows the application domains of the reviewed studies. In the context of datasets, some studies have used real-time datasets [27] that reflect the problems of unbalanced, noisy, or incomplete data. This enables hybrid metaheuristics to exhibit its robustness. Several studies have used UCI ML and NSL-KDD benchmark datasets to evaluate the performance of hybrid metaheuristics; this increased transparency and illustrated the rigorousness of the evaluation process. Some studies use high-dimensional datasets to demonstrate the algorithm's ability to manage the "curse of dimensionality," with limited samples but a large number of features per sample, such as the Leukemia dataset [43] and several benchmark cancer microarray gene expression datasets [49]. The succsssessful selection of significant features from these datasets illustrates the robustness of the metaheuristic algorithms.

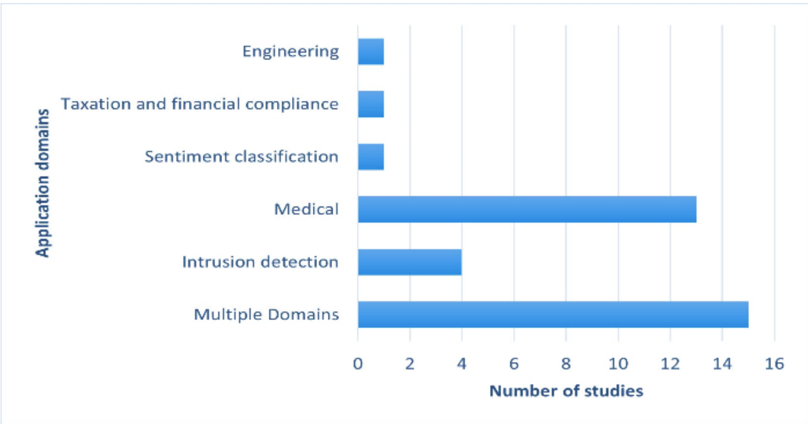


Figure 2. Number of studies by application domains

One successful method to improve metaheuristic performance is the incorporation of adaptive GA operators. By dynamically tuning crossover and mutation probabilities based on population feedback, these adaptive mechanisms enhance search efficiency, prevent being stuck in local optima and premature convergence. Hence, scholars has interested to integrate the adaptive GA operators with metaheuristic like FA [44] and PSO [45]. The suggested hybrid algorithm has a heightened capacity for exploration to prevent becoming stuck in local optimal. Moreover, the hybridization of GA with arithmetic optimization algorithm (AOA) was proposed by Ewees *et al.* [27]. The suggested technique addressed the primary limitations of the traditional AOA by avoiding the difficulty of local search. In comparison to AOA, the hybrid AOAGA attained approximately 4% greater accuracy, decreased the number of selected features by 15%, and lowered runtime by roughly 36%, confirming the efficacy of hybridization in improving performance.

The GWO, a modern bio-inspired algorithm, is gaining interest in hybrid metaheuristic algorithms for FS due to its adaptability and simplicity. Moreover, it exhibits a lower number of parameters, rendering it more cost-effective in terms of computational expenses and facilitating faster convergence when compared with other metaheuristics [50]. Each solution in GWO is updated based on the positions of the three optimal solutions within the population, making the algorithm more oriented to the exploitation than to the exploration and causing premature convergence because search agents do not explore the search space efficiently [23]. Within studies reviewed in this SLR, GWO-based hybrids were the most prevalent reported in 10 studies as shown in Figure 3.

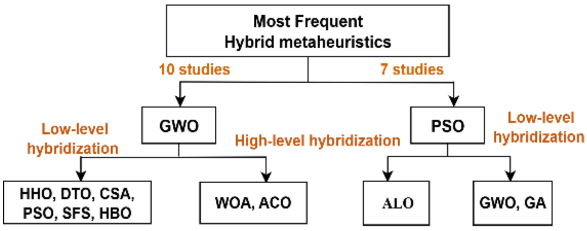


Figure 3. Most frequent hybrid metaheuristics for FS in classification

Several studies have hybridized GWO with other metaheuristics at the low-level (cooperative), combining operators from different metaheuristics within the same search process to improve its performance such as using operators from PSO in [3] and crow search algorithm (CSA) in [23]. Other studies have also used complementary strategies such as using HHO to explore the search space while GWO refined solutions (exploitation) in Al-Wajih *et al.* [21]. Another study [7] separated the population into two groups, one directed by PSO and the other by GWO, thereby enhancing the search process. Similarly, [22], [33] introduced a control parameter that dictates whether agents follow dipper throated optimization (DTO) or GWO principles. The stochastic fractal search (SFS) algorithm enhances GWO by implementing a diffusion procedure on the optimal wolf for random walk in a growth process and preventing premature convergence [25]. High-level hybridization of GWO are relatively rare, reported in only two studies [24], [42] demonstrating that GWO is mostly utilized as a cooperative component for exploitation, with its exploration weaknesses being addressed through complementary strategies. Most of these hybrid algorithms fail to incorporate adaptive techniques that adjust the balance parameters depending on feedback from search process. The adaptive balance techniques can effectively regulate the transition between exploration and exploitation, enabling the algorithm to accelerate early convergence while maintaining solution quality in later stages such as adaptive balance probability in [23].

Among the reviewed studies, the hybrid of PSO and GA was noted three times [29], [38], [45] indicating its popularity as a low-level hybridization, where GA contributes global exploration and PSO improves convergence. Furthermore, PSO-based hybrids appeared in seven studies, five of them were low-level hybridization. It used to employed to improve the performance of other metaheuristics such as increasing the computational efficiency of antlion optimization (ALO) [43], accelerating the convergence of GA [29], enhancing exploration of GWO [3]. In contrast, other studies concentrate on enhancing PSO by hybrid it with other algorithm like, utilizing CS to enhance PSO exploitation [51], GA to prevent its immature convergence [45]. Overall, hybridization enables algorithms to share each other's strengths to enhance their performance.

RQ2: What type of hybridization is used? What are the evaluation metrics? What statistical tests are used?

Reviewed studies adopt high-level hybridization that integrate classifiers with metaheuristic algorithms, where binary metaheuristics generate feature subsets and classifiers assess their efficacy. Nevertheless, only six studies conducted hybridization between two metaheuristic algorithms at the high level, two employed a parallel [24], [46], where algorithms run simultaneously and they can operate independently and periodically to exchange information, while the remaining employed a sequential structure (the algorithms run one after another). In contrast, a smaller number of studies implement low-level hybridization [3], [23], [27], which involves merging the functionality of the two algorithms within a single framework by adopting or transferring a specific operator or control parameter from one algorithm to another to improve overall performance. In contrast, other studies employ the same low-level hybridization concept in a more integrated way, that employing most operators from both algorithms within the same iterations, for example population split, one group follows one algorithm, the other follows second algorithm [7], one algorithm for exploration and another one for exploitation [21], or alternates between the two algorithms' position update formulas based on the random parameter [22]. Low-level hybridization generally exhibits less time complexity due to the integration of operators from both traditional metaheuristics occurs within a single optimization framework, allowing shared fitness evaluations. However, high-level hybrids involve sequential execution tend to increase computational time due to multiple fitness assessments [24].

An essential step in the classification process is evaluating a predictive model. The evaluation of FS-based metaheuristics performance involved the use of various standard measures, including mean, best, worst fitness, classification accuracy/error, computational time, sensitivity, specificity, F-score and standard deviation, as shown in Figure 4. Nevertheless, the criteria that were most frequently employed in reported studies included classification accuracy/error rate, mean fitness, recall, and the size of selected features because they represent the main goals of FS in classification tasks, where accuracy and recall reflect the classification quality, mean fitness and selected feature size reflect the optimization performance, and reduction efficiency respectfully.

Statistical tests play an essential role in assessing the quality of the model. The common statistical tests used in FS with metaheuristics comprise Wilcoxon, Friedman, ANOVA and T-test. Only 14 papers were rated (40%) as having used statistical tests to evaluate their models, this considered a limitation in the literature because statistical testing is crucial to analysis the significance of the results and ensure the superiority of hybrid algorithms [41]. The most popular test is the Wilcoxon test, which was used in 31% of publications, followed by the Friedman test, which was employed in 6 papers [23], [27], [31], [32], [41], [42]. Additionally, the ANOVA was used in 4 papers [4], [26], [33], [39].

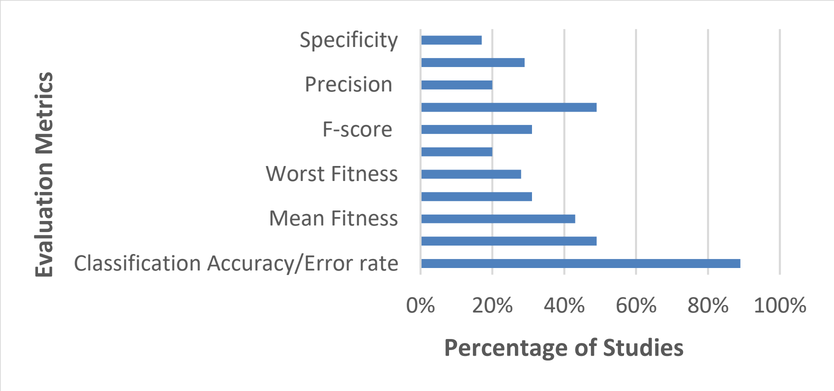


Figure 4. Evaluation metrics in the analysed FS studies

RQ3: What are the FS techniques applied with metahumans to achieve good classification accuracy and minimum number of features? What classifiers are used? What initial population methods are used?

Traditional FS techniques are essentially categorized into three main groups: filter, wrapper, and embedded. The filter technique assesses individual features in the dataset according to their theoretical or statistical characteristics without using any classification algorithms [52]. It is less expensive computationally and has faster execution time than wrapper methods due to working independently of classifier. However, it has the drawback of ignoring the performance of the selected features [53]. Only 14% of the studies reviewed [37], [42], [47]–[49] employed a filter-based approach to solve the FS problem. In these studies, the FS process has occurred in two sequential phases that integrate the qualities of both filter and wrapper approaches. First, a filter-based approach is utilized to eliminate irrelevant and redundant features, subsequently the wrapper approach-based hybrid metaheuristics is applied for the final choice of the relevant features. This sequential process is particularly useful in high-dimensional datasets [43], [48] to diminish the solution space through a filtering phase and enhance the subset using wrapper-based metaheuristics.

Wrapper techniques require a learning algorithm and evaluate its performance. This dependence criterion requires the predefined learning strategy in FS and relies on its efficacy to choose features [8]. Many articles have employed wrapper technique-based metaheuristics to decrease the number of features due to it giving a more superior performance than the filter method. Several classification approaches are available, each with strengths and drawbacks that are appropriate for various datasets and problem domains. Among the most popular approaches are K-nearest neighbors (KNN), support vector machine (SVM), arificial neural network (ANN), Naive Bayes (NB), deep learning (DL), and random forest (RF).

The KNN method, a simple and widely used method that integrates with metaheuristics, is used for improving FS in classification tasks due to its efficacy and stability. Only two study [45], [47] used KNN with other classifiers when building the fitness function, while 18 papers reported that KNN classification accuracy is only used to construct fitness functions. In most of these studies (9 papers), a common practice is to assign a value of 5 for the parameter "k" as a suitable value to achieve high accuracy. Additionally, SVM attracted attention from 6 studies due to SVM’s advantages, including its efficacy in space with high dimensions, adaptability, and memory efficiency, regarding other classifiers used in a few studies, as displayed in Figure 5.

Metaheuristics are exploration search algorithms with a wide diversity of initial populations, as insufficient diversity can cause stagnation in local optima [54]. Usually, the initial population is generated at random, as reported in the reviewed studies. While Almazini *et al.* [42] initialized the wolves' population in GWO by using a heuristic-based ACO aiming to generate solutions by choosing features that optimize the classification accuracy. Houssein *et al.* [40] introduced the utilization of chaotic maps for initializing solutions for updating the control energy parameters in HHO to prevent the occurrence of local optima and premature convergence.

RQ4: What are the crucial/optimal parameters values? Which transfer and fitness functions are used?

Research on metaheuristics emphasizes fine-tuning parameters like population size and iteration number for efficient and high-quality computations [54]. The algorithm's repeated execution broadens its search space, minimizing randomness's impact and increasing the likelihood of discovering optimal solutions. Among the studies analyzed in this SLR, in most studies 43%, the iteration parameter is assigned to a value of 100, and 33% of studies used 10 agents as a population size. Nevertheless, it has been noted that the metaheuristic-based FS algorithms described in [33], [42], [46] lack the essential statistical analysis to illustrate the significance and superiority of these parameters, which is a crucial component of empirical research.

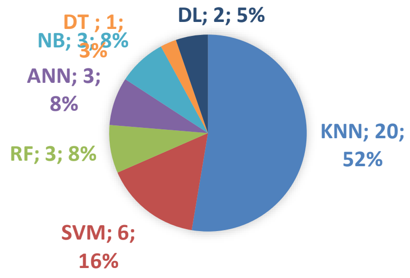


Figure 5. Number of studies by classification algorithm

Metaheuristics are typically used to solve continuous problems, while FS is a binary problem with d-dimensional vectors, requiring transfer functions to convert continuous values into binary. The dominance of S-shaped transfer functions in 15 studies indicates a preference for smoother transition behavior. Nonetheless, the absence of reporting in 12 research indicates limited methodological transparency. Furthermore, V-shaped function was only examined in one study [30], indicating that alternative transfer mechanisms are still underexplored. The fitness function determines the quality of the solution based on the features that have been chosen. The formulation of an efficient fitness function is necessary for the success of the process. The common fitness function used by recent studies reviewed in this SLR focuses on classification error and the number of selected features.

The primary goals of the FS problem are to reduce the total number of selected features and maximize performance accuracy. Consequently, we focused investigation exclusively on these two metrics. In some studies, error rates were provided; accuracy was recalculated as (1-Error Rate), and it was normalized to a percentage format. The chosen features were represented as a percentage of the total features in the original dataset. Additionally, we computed the average results using the results available in articles for studies that did not report the average results across all datasets.

The comparison between the hybrid and native metaheuristics results in terms of average accuracy, selected feature size and computational time presents in Figures 6 and 7. Figure 6 provides a comparative analysis of the 11 reviewed studies that reported accuracy for both single and hybrid metaheuristics, the average accuracy of the hybrid metaheuristics surpassed that of the two single baselines.

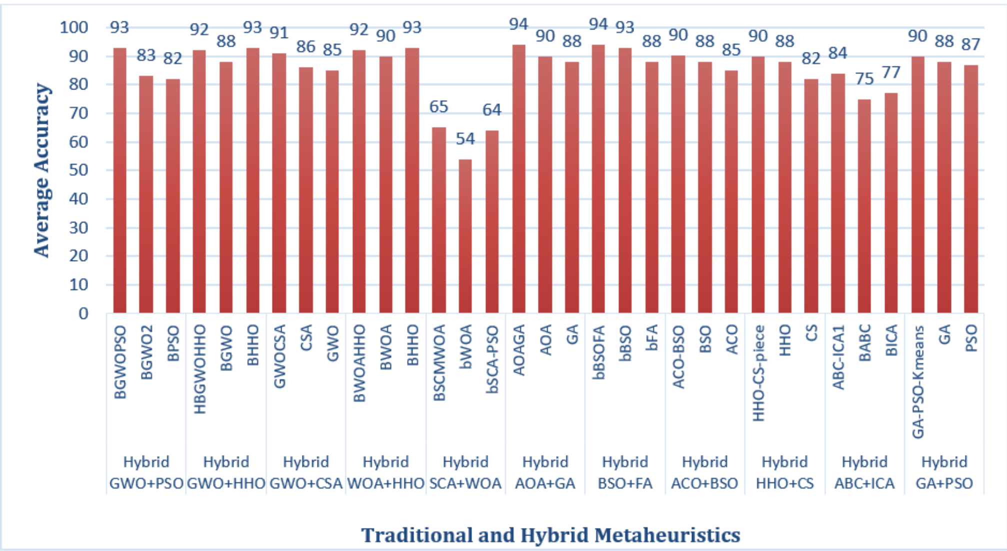


Figure 6. Average accuracy comparison between traditional and hybrid metaheuristics

Figure 7 presents comparison boxplots that include the published results from a total of 20 studies, all of which reported accuracy results, while 15 of these studies provided the selected feature size, and 7 reported computational time. The boxplots show that the hybrid metaheuristics achieve superior and more stable accuracy compared to the single version. Moreover, the findings indicate that hybrid metaheuristics surpassed the native algorithms in reduced dimensionality. The optimal feature subset selected by hybrid metaheuristics improved the model performance, as demonstrated by the high classification accuracy achieved by these algorithms. Although both single and hybrid methods exhibit variation in computation time, the hybrid methods outperform the single approach in terms of median time, indicating its effectiveness in converging faster towards optimal or near-optimal solutions. This improvement is due to the capability of hybrid metaheuristics to balance between exploration and exploitation effectively, avoid getting stuck in local optima, and eliminate the irrelevant and redundant features that generally improve performance of the model.

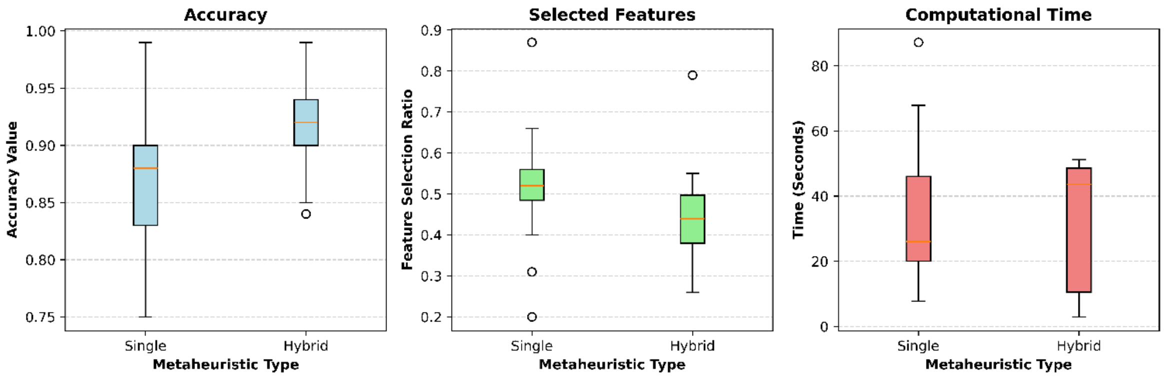


Figure 7. Boxplot comparison of traditional and hybrid metaheuristics

4. LIMITATIONS AND FUTURE DIRECTIONS

Although the selected studies prove effective in solving the FS problem, they still suffer from shortcomings that indicate the need for more investigation in future studies.

a. Integration with deep learning, the dominance of wrapper-based hybrid metaheuristics in solving FS problem a promising future direction is the combination of hybrid metaheuristics-based FS techniques with DL models to reduce dimensionality while preserving classification accuracy, thus, enhancing the scalability and efficiency of DL systems.
b. Most metaheuristic algorithms have a lot of parameters [17] that need to be carefully tuned. Finding the best parameter tuning is a challenge.
c. Expanding applications to emerging domains: Most existing studies focus on benchmark or biomedical datasets. Future research could extend hybrid metaheuristics to emerging domains such as smart cities, climate modeling, and intelligent systems. Moreover, incorporating recent trends such as hybridization with transformers, vision transformers (ViTs), and reinforcement learning may further enhance adaptability and decision-making capabilities, offering more robust and intelligent solutions to complex real-world challenges.
d. Most selected studies use the hybrid metaheuristics and wrapper approach as traditional FS. Hence, creating adaptive hybrid approaches that optimize performance by dynamically adjusting the filter wrapper balance during selection is an open research topic.
e. Adaptively adjusting the crucial parameters based on the algorithm's feedback results, and hybrid advanced method with metaheuristics to use it for FS.
f. There is a limited amount of research on multi-objective hybrid metaheuristics-based FS. Therefore, the use of hybridization of metaheuristics in multi-objective FS is an open research topic for scholars.
g. The computational overhead of hybrid metaheuristic algorithms is a significant challenge that arises from the complexity of integrating many techniques and fine-tuning additional parameters for enhanced performance, some reported studies have demonstrated improvement but have issues related to computational burden and time complexity. It is advisable to provide more suitable methods to reduce it.
h. Future research should investigate hybrid integrations of bio-inspired, evolutionary, and predator-prey metaheuristics, which remain significantly underexplored in the existing literature.

5. CONCLUSION

The growing interest in metaheuristics for FS is due to their robust capacity to manage dimensionality reduction tasks and attain superior performance compared to traditional algorithms. This SLR analyzed and synthesized 35 papers from a total of 646 papers obtained from well-known search databases aimed to provide a comprehensive understanding of how hybrid metaheuristics are applied in classification. This SLR outlines the statistical analysis conducted to ensure that relevant studies are carefully chosen. The most common algorithm used is GWO, followed by PSO and GA with KNN and SVM serving as dominant classifiers for fitness evaluation. The results indicate that metaheuristic algorithms hybridized at a low level are generally more computationally efficient than those hybridized at a high level. This efficiency results from algorithmic operators being directly integrated, which reduces redundant evaluations and communication overhead across separate components. S-shaped transfer function is the approach that is most used in the reviewed studies. The random initialize population is common in most reported studies. Hybrid metaheuristics surpassed their individual counterparts by attaining a better balance between exploration and exploitation, alleviating premature convergence, and enhancing overall optimization performance. Future research should prioritize the developing adaptive and multi-objective hybrid frameworks that use DL or clustering techniques in creating fitness functions for FS problems and focus on the interpretability and ethical dimensions by including explainable methodologies and guaranteeing transparency and fairness in model outcomes. Implementing these hybrid algorithms in practical fields like finance, cybersecurity, engineering, and smart city can enhance classification accuracy, diminish computational expenses, and facilitate more effective decision-making in high-dimensional datasets.

FUNDING INFORMATION

The authors state that no funding was involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Manal Othman	✓	✓	✓		✓	✓	✓	✓	✓		✓		✓	
Ku Ruhana Ku-Mahamud	✓			✓						✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data and materials included and referenced in the manuscript.

REFERENCES

[1] M. M. O. F. Asaad, J. Wahid, and A. R. Rahmat, "Employing artificial bee colony algorithm to optimize the artificial neural network in heart disease prediction," in *AIP Conference Proceedings*, 2024, doi: 10.1063/5.0192144.

[2] K. Kowsari, K. J. Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, "Text classification algorithms: A survey," *Information*, vol. 10, no. 4, Apr. 2019, doi: 10.3390/info10040150.

[3] Q. Al-Tashi, S. J. A. Kadir, H. M. Rais, S. Mirjalili, and H. Alhussian, "Binary optimization using hybrid grey wolf optimization for feature selection," *IEEE Access*, vol. 7, pp. 39496–39508, 2019, doi: 10.1109/ACCESS.2019.2906757.

[4] R. Alwajih *et al.*, "Hybrid binary whale with Harris Hawks for feature selection," *Neural Computing and Applications*, vol. 34, no. 21, pp. 19377–19395, Nov. 2022, doi: 10.1007/s00521-022-07522-9.

[5] Z. S. Mahdi, R. M. Zaki, A. K. Farhan, and N. Majma, "Development of a hybrid methodology of deep learning and machine learning for lung nodule detection in medical computed tomography images," *Journal of Soft Computing and Computer Applications*, vol. 1, no. 2, Dec. 2024, doi: 10.70403/3008-1084.1011.

[6] Z. M. Radeef, S. H. Hashem, and E. K. Gbashi, "New feature selection using principal component analysis," *Journal of Soft Computing and Computer Applications*, vol. 1, no. 2, Dec. 2024, doi: 10.70403/3008-1084.1012.
[7] E.-S. M. El-kenawy and M. Eid, "Hybrid gray wolf and particle swarm optimization for feature selection," *International Journal of Innovative Computing, Information & Control (IJICIC)*, vol. 16, no. 3, pp. 831–844, 2020.
[8] J. Zhang, Y. Xiong, and S. Min, "A new hybrid filter/wrapper algorithm for feature selection in classification," *Analytica Chimica Acta*, vol. 1080, pp. 43–54, Nov. 2019, doi: 10.1016/j.aca.2019.06.054.
[9] K. Chen, F.-Y. Zhou, and X.-F. Yuan, "Hybrid particle swarm optimization with spiral-shaped mechanism for feature selection," *Expert Systems with Applications*, vol. 128, pp. 140–156, Aug. 2019, doi: 10.1016/j.eswa.2019.03.039.
[10] J. Piri, P. Mohapatra, R. Dey, B. Acharya, V. C. Gerogiannis, and A. Kanavos, "Literature review on hybrid evolutionary approaches for feature selection," *Algorithms*, vol. 16, no. 3, Mar. 2023, doi: 10.3390/a16030167.
[11] T. Dokeroglu, A. Deniz, and H. E. Kiziloz, "A comprehensive survey on recent metaheuristics for feature selection," *Neurocomputing*, vol. 494, pp. 269–296, Jul. 2022, doi: 10.1016/j.neucom.2022.04.083.
[12] 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, doi: 10.1109/ACCESS.2021.3056407.
[13] E. O. Abiodun, A. Alabdulatif, O. I. Abiodun, M. Alawida, A. Alabdulatif, and R. S. Alkhawaldeh, "A systematic review of emerging feature selection optimization methods for optimal text classification: The present state and prospective opportunities," *Neural Computing and Applications*, vol. 33, no. 22, pp. 15091–15118, Nov. 2021, doi: 10.1007/s00521-021-06406-8.
[14] O. M. Alyasiri, Y.-N. Cheah, A. K. Abasi, and O. M. Al-Janabi, "Wrapper and hybrid feature selection methods using metaheuristic algorithms for English text classification: A systematic review," *IEEE Access*, vol. 10, pp. 39833–39852, 2022, doi: 10.1109/ACCESS.2022.3165814.
[15] M. Sharma and P. Kaur, "A comprehensive analysis of nature-inspired meta-heuristic techniques for feature selection problem," *Archives of Computational Methods in Engineering*, vol. 28, no. 3, pp. 1103–1127, May 2021, doi: 10.1007/s11831-020-09412-6.
[16] M. Rostami, K. Berahmand, E. Nasiri, and S. Forouzandeh, "Review of swarm intelligence-based feature selection methods," *Engineering Applications of Artificial Intelligence*, vol. 100, Apr. 2021, doi: 10.1016/j.engappai.2021.104210.
[17] O. O. Akinola, A. E. Ezugwu, J. O. Agushaka, R. A. Zitar, and L. Abualigah, "Multiclass feature selection with metaheuristic optimization algorithms: A review," *Neural Computing and Applications*, vol. 34, no. 22, pp. 19751–19790, Nov. 2022, doi: 10.1007/s00521-022-07705-4.
[18] T. H. Pham and B. Raahemi, "Bio-inspired feature selection algorithms with their applications: A systematic literature review," *IEEE Access*, vol. 11, pp. 43733–43758, 2023, doi: 10.1109/ACCESS.2023.3272556.
[19] B. Kitchenham *et al.*, "Systematic literature reviews in software engineering – A tertiary study," *Information and Software Technology*, vol. 52, no. 8, pp. 792–805, Aug. 2010, doi: 10.1016/j.infsof.2010.03.006.
[20] M. J. Page *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, Mar. 2021, doi: 10.1136/bmj.n71.
[21] 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, doi: 10.1109/ACCESS.2021.3060096.
[22] D. S. Khafaga *et al.*, "Hybrid dipper throated and grey wolf optimization for feature selection applied to life benchmark datasets," *Computers, Materials & Continua*, vol. 74, no. 2, pp. 4531–4545, 2023, doi: 10.32604/cmc.2023.033042.
[23] S. Arora, H. Singh, M. Sharma, S. Sharma, and P. Anand, "A new hybrid algorithm based on grey wolf optimization and crow search algorithm for unconstrained function optimization and feature selection," *IEEE Access*, vol. 7, pp. 26343–26361, 2019, doi: 10.1109/ACCESS.2019.2897325.
[24] M. Mafarja, A. Qasem, A. A. Heidari, I. Aljarah, H. Faris, and S. Mirjalili, "Efficient hybrid nature-inspired binary optimizers for feature selection," *Cognitive Computation*, vol. 12, no. 1, pp. 150–175, Jan. 2020, doi: 10.1007/s12559-019-09668-6.
[25] E.-S. M. El-Kenawy, M. M. Eid, M. Saber, and A. Ibrahim, "MbGWO-SFS: Modified binary grey wolf optimizer based on stochastic fractal search for feature selection," *IEEE Access*, vol. 8, pp. 107635–107649, 2020, doi: 10.1109/ACCESS.2020.3001151.
[26] E.-S. M. El-Kenawy *et al.*, "Novel meta-heuristic algorithm for feature selection, unconstrained functions and engineering problems," *IEEE Access*, vol. 10, pp. 40536–40555, 2022, doi: 10.1109/ACCESS.2022.3166901.
[27] A. A. Ewees *et al.*, "Boosting arithmetic optimization algorithm with genetic algorithm operators for feature selection: Case study on cox proportional hazards model," *Mathematics*, vol. 9, no. 18, Sep. 2021, doi: 10.3390/math9182321.
[28] H. Liang, Z. Wang, and Y. Liu, "A new hybrid ant colony optimization based on brain storm optimization for feature selection," *IEICE Transactions on Information and Systems*, vol. E102.D, no. 7, pp. 1396–1399, Jul. 2019, doi: 10.1587/transinf.2019EDL8001.
[29] X. Li, J. Zhang, and F. Safara, "Improving the accuracy of diabetes diagnosis applications through a hybrid feature selection algorithm," *Neural Processing Letters*, vol. 55, no. 1, pp. 153–169, Feb. 2023, doi: 10.1007/s11063-021-10491-0.
[30] R. Hans and H. Kaur, "Hybrid binary sine cosine algorithm and ant lion optimization (SCALO) approaches for feature selection problem," *International Journal of Computational Materials Science and Engineering*, vol. 09, no. 01, Mar. 2020, doi: 10.1142/S2047684119500210.
[31] T. Bezdan, M. Zivkovic, N. Bacanin, A. Chhabra, and M. Suresh, "Feature selection by hybrid brain storm optimization algorithm for Covid-19 classification," *Journal of Computational Biology*, vol. 29, no. 6, pp. 515–529, Jun. 2022, doi: 10.1089/cmb.2021.0256.
[32] O. A. Akinola, A. E. Ezugwu, O. N. Oyelade, and J. O. Agushaka, "A hybrid binary dwarf mongoose optimization algorithm with simulated annealing for feature selection on high dimensional multi-class datasets," *Scientific Reports*, vol. 12, no. 1, Sep. 2022, doi: 10.1038/s41598-022-18993-0.
[33] R. Alkanhel *et al.*, "Network intrusion detection based on feature selection and hybrid metaheuristic optimization," *Computers, Materials & Continua*, vol. 74, no. 2, pp. 2677–2693, 2023, doi: 10.32604/cmc.2023.033273.
[34] M. Alweshah *et al.*, "Hybrid black widow optimization with iterated greedy algorithm for gene selection problems," *Heliyon*, vol. 9, no. 9, Sep. 2023, doi: 10.1016/j.heliyon.2023.e20133.
[35] C.-Y. Lee, T.-A. Le, and Y.-T. Lin, "A feature selection approach hybrid grey wolf and heap-based optimizer applied in bearing fault diagnosis," *IEEE Access*, vol. 10, pp. 56691–56705, 2022, doi: 10.1109/ACCESS.2022.3177735.
[36] S. Thawkar, "A hybrid model using teaching–learning-based optimization and salp swarm algorithm for feature selection and classification in digital mammography," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 9, pp. 8793–8808, Sep. 2021, doi: 10.1007/s12652-020-02662-z.
[37] D. Kumar and M. Phogat, "A hybrid metaheuristics based technique for mutation based disease classification," *International Journal of Electrical and Computer Engineering Systems*, vol. 14, no. 6, pp. 635–646, Jul. 2023, doi: 10.32985/ijeces.14.6.3.

[38] M. G. El-Shafiey, A. Hagag, E.-S. A. El-Dahshan, and M. A. Ismail, "A hybrid GA and PSO optimized approach for heart-disease prediction based on random forest," *Multimedia Tools and Applications*, vol. 81, no. 13, pp. 18155–18179, May 2022, doi: 10.1007/s11042-022-12425-x.

[39] A. A. Alhussan *et al.*, "Classification of diabetes using feature selection and hybrid Al-Biruni earth radius and dipper throated optimization," *Diagnostics*, vol. 13, no. 12, Jun. 2023, doi: 10.3390/diagnostics13122038.

[40] E. H. Houssein, M. E. Hosney, M. Elhoseny, D. Oliva, W. M. Mohamed, and M. Hassaballah, "Hybrid Harris Hawks optimization with cuckoo search for drug design and discovery in chemoinformatics," *Scientific Reports*, vol. 10, no. 1, Sep. 2020, doi: 10.1038/s41598-020-71502-z.

[41] A. Osmani, J. B. Mohasefi, and F. S. Gharehchopogh, "Sentiment classification using two effective optimization methods derived from the artificial bee colony optimization and imperialist competitive algorithm," *The Computer Journal*, vol. 65, no. 1, pp. 18–66, Jan. 2022, doi: 10.1093/comjnl/bxz163.

[42] H. Almazini, K. R. Ku-Mahamud, and H. Almazini, "Heuristic initialization using grey wolf optimizer algorithm for feature selection in intrusion detection," *International Journal of Intelligent Engineering and Systems*, vol. 16, no. 1, pp. 410–418, Feb. 2023, doi: 10.22266/ijies2023.0228.36.

[43] T. R. Mahesh, D. Santhakumar, A. Balajee, H. S. Shreenidhi, V. V. Kumar, and J. R. Annand, "Hybrid ant lion mutated ant colony optimizer technique with particle swarm optimization for leukemia prediction using microarray gene data," *IEEE Access*, vol. 12, pp. 10910–10919, 2024, doi: 10.1109/ACCESS.2024.3351871.

[44] K. Alwan, A. AbuEl-Atta, and H. Zayed, "Feature selection models based on hybrid firefly algorithm with mutation operator for network intrusion detection," *International Journal of Intelligent Engineering and Systems*, vol. 14, no. 1, pp. 192–202, Feb. 2021, doi: 10.22266/ijies2021.0228.19.

[45] S. Masrom, R. A. Rahman, M. Mohamad, A. S. A. Rahman, and N. Baharun, "Machine learning of tax avoidance detection based on hybrid metaheuristics algorithms," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 3, pp. 1153–1163, Sep. 2022, doi: 10.11591/ijai.v11.i3.pp1153-1163.

[46] S. Shanthi, V. S. Akshaya, J. A. Smitha, and M. Bommy, "Hybrid TABU search with SDS based feature selection for lung cancer prediction," *International Journal of Intelligent Networks*, vol. 3, pp. 143–149, 2022, doi: 10.1016/j.ijin.2022.09.002.

[47] A. Yaqoob, N. K. Verma, and R. M. Aziz, "Optimizing gene selection and cancer classification with hybrid sine cosine and cuckoo search algorithm," *Research Square,* Sep. 25, 2023, doi: 10.21203/rs.3.rs-3357558/v1.

[48] R. Mahto *et al.*, "A novel and innovative cancer classification framework through a consecutive utilization of hybrid feature selection," *BMC Bioinformatics*, vol. 24, no. 1, Dec. 2023, doi: 10.1186/s12859-023-05605-5.

[49] A. A. Joshi and R. M. Aziz, "A two-phase cuckoo search based approach for gene selection and deep learning classification of cancer disease using gene expression data with a novel fitness function," *Multimedia Tools and Applications*, vol. 83, no. 28, pp. 71721–71752, Feb. 2024, doi: 10.1007/s11042-024-18327-4.

[50] Q. Al-Tashi *et al.*, "Binary multi-objective grey wolf optimizer for feature selection in classification," *IEEE Access*, vol. 8, pp. 106247–106263, 2020, doi: 10.1109/ACCESS.2020.3000040.

[51] A. A. Joshi and R. M. Aziz, "Deep learning approach for brain tumor classification using metaheuristic optimization with gene expression data," *International Journal of Imaging Systems and Technology*, vol. 34, no. 2, Mar. 2024, doi: 10.1002/ima.23007.

[52] B. H. Nguyen, B. Xue, and M. Zhang, "A survey on swarm intelligence approaches to feature selection in data mining," *Swarm and Evolutionary Computation*, vol. 54, May 2020, doi: 10.1016/j.swevo.2020.100663.

[53] A. Chaudhuri and T. P. Sahu, "A hybrid feature selection method based on Binary Jaya algorithm for micro-array data classification," *Computers & Electrical Engineering*, vol. 90, Mar. 2021, doi: 10.1016/j.compeleceng.2020.106963.

[54] U. Mumtahina, S. Alahakoon, and P. Wolfs, "Hyperparameter tuning of load-forecasting models using metaheuristic optimization algorithms—a systematic review," *Mathematics*, vol. 12, no. 21, Oct. 2024, doi: 10.3390/math12213353.

APPENDIX

Table 3. Summary of hybrid metaheuristics for FS

Ref.	Hybrid alg	Hybrid aim	Hybrid strategy	Type		Classifier	Wrapper or filter
				Low	High		
Al-Tashi *et al.* [3]	GWO+PSO	Integrate PSO exploitation with GWO exploration for enhanced search capability.	coevolutionary integration with inertia-based control of agent updates.	√		KNN	Wrapper
Al-Wajih *et al.* [21]	GWO+HHO	Avoid premature convergence. Balance of exploration-exploitation.	HHO for exploration, GWO for exploitation.	√			Wrapper
El-Kenawy and Eid [7]	GWO+PSO	Balance of exploration-exploitation.	Population split, one group follows GWO, the other PSO.	√			Wrapper
Khafaga *et al.* [22]	GWO+DTO		Alternates between GWO/DTO position update based on the random parameter.	√			Wrapper
Arora *et al.* [23]	GssWO+CSA	Enhance exploration–exploitation to reach global optima.	Combines position update rules of GWO with flight intelligence mechanism of CSA.	√			Wrapper
Mafarja *et al.* [24]	GWO+WOA	Prevent immature convergence and local optima stagnation.	Serial, random, and adaptive switcher for (GWO-WOA).		√		Wrapper
Alwajih *et al.* [4]	WOA+HHO	Increase randomness and improve search capability.	Embed HHO exploration in WOA.	√			Wrapper

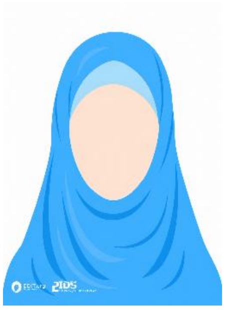
Table 3. Summary of hybrid metaheuristics for FS *(continued)*

Ref.	Hybrid alg	Hybrid aim	Hybrid strategy	Type Low	Type High	Classifier	Wrapper or filter
El-Kenawy *et al.* [25]	GWO+SFS	Balance exploration-exploitation and improve diversity.	Apply SFS diffusion on the best solution of Modified GWO using Gaussian random walk.	√		KNN	Wrapper
El-Kenawy *et al.* [26]	Sine cosine algorithm (SCA)+WOA	Enhance WOA by avoiding local optima and improving convergence rate.	Balance agent position updates between SCA and modified WOA during iterations with added random agents for stronger global search.	√			Wrapper
Ewees *et al.* [27]	AOA+GA	Improve search strategies of AOA and balance exploration–exploitation by overcoming its weak local search and trade-off limitations.	Incorporate GA operators into AOA.	√			Wrapper
Liang *et al.* [28]	ACO+ brain storm optimization (BSO)	Overcome ACO's stagnation in local optima and premature convergence.	Integrates ACO with BSO by modelling ants as individuals and colonies as clusters in brainstorming process.	√			Wrapper
Li *et al.* [29]	PSO+GA	Avoid local optima and accelerate convergence.	PSO improves GA convergence.		√		Wrapper
Hans and Kaur [30]	ALO+SCA	Improve exploration–exploitation balance and maintain diversity.	Split population: SCA updates first half, ALO random walks update second half.	√			Wrapper
Bezdan *et al.* [31]	BSO+FA	Achieve a better balance between exploration and exploitation.	Integrates FA operators into BSO to enhance search balance.	√			Wrapper
Akinola *et al.* [32]	DMO+ simulated annealing (SA)	To improve the limited exploitative process of the DMO.	BDMO for global search and SA for local refinement.	√			Wrapper
Alkanhel *et al.* [33]	GWO+DTO	Overcome GWO stagnation and improve exploration–exploitation.	DTO is used to modify the search operators of GWO.	√			Wrapper
Alweshah *et al.* [34]	Black widow optimization algorithm (BWO)+ iterated greedy algorithm (IG)	To enhance the local search capabilities of the BWO algorithm.	IG embedded into BWO to improve local search.	√			Wrapper
Lee *et al.* [35]	GWO+heap-based optimizer (HBO)	Improve exploration–exploitation balance and avoid local optima.	Hybrid GWO–HBO with crossover and mutation updates.	√			Wrapper
Thawkar [36]	TLBO+SSA	Improve convergence speed, efficiency, and population diversity.	Integrates SSA update procedure into TLBO structure.	√		ANN	Wrapper
Kumar and Phogat [37]	Improved binary competitive swarm optimization (IBCSO)+WOA	Balance of exploration-exploitation.	IBCSO for feature reduction, WOA for parameter tuning.	√			Wrapper and filter
El-Shafiey *et al.* [38]	PSO+GA	Improving accuracy through global–local search balance.	GA for global and PSO for local refinement.		√	RF	Wrapper
Alhussan *et al.* [39]	Dynamic al-Biruni earth radius (DBER)+DTO	Improve exploration and exploitation of the search space.	Dynamic hybrid BER and DTO operators.	√			Wrapper
Houssein *et al.* [40]	HHO+CS	Improve exploration, avoid local optima and premature convergence.	CS operators refine HHO positions.	√		SVM	Wrapper
Osmani *et al.* [41]	ABC+CA	Enhance ABC exploitation and balance exploration-exploitation.	Embed imperialist competitive algorithm (ICA) exploitation in ABC phases; solutions guided toward best bee or neighbours.	√			Wrapper

Table 3. Summary of hybrid metaheuristics for FS *(continued)*

Ref.	Hybrid alg	Hybrid aim	Hybrid strategy	Type Low	Type High	Classifier	Wrapper or filter
Almazini *et al.* [42]	GWO+ACO	Enhance wolf population initialization through the concept of ACO.	Heuristic ACO for MBGWO initialization.		√	SVM	Wrapper and filter
Mahesh *et al.* [43]	ALO+PSO	To enhance the computational efficiency of both algorithms.	PSO initializes and refines the population, ALO diversifies the search with roulette selection and random walks.		√		Wrapper
Alwan *et al.* [44]	FA+GA	Prevent being stuck in local optima through improving the exploration abilities of the standard FA.	GA mutation operator updates the best firefly's position.	√		NB	Wrapper
Masrom *et al.* [45]	PSO+GA	To resolve the problem of immature convergence in PSO.	Integrates adaptive GA operators into PSO.	√		KNN, SVM, and RF	Wrapper
Shanthi *et al.* [46]	Stochastic diffusion search (SDS)+tabu search (TS)	SDS injects diversity into TS if there are no good solutions.	Replace old solutions in the tabu list with new ones from SDS.		√	ANN, decision tree (DT), and NB	Wrapper
Yaqoob *et al.* [47]	SCA+CS	Improve convergence and solution quality.	CS expands search space, SCA refines within dimensions.			KNN, SVM, and NB	Wrapper and filter
Mahto *et al.* [48]	Spider Monkey Optimization (SMO)+CS	Improve accuracy and balance exploration–exploitation.	CS modifies SMO fitness.	√		DL	Wrapper and filter
Joshi and Aziz [49]			Combines SMO global exploration with CS local search for exploitation.	√			Wrapper and filter

BIOGRAPHIES OF AUTHORS



Manal Othman holds a B.Sc. degree in computer science from Taiz University, Yemen, and the M.S. degree in information technology (IT) from the School of Computing, Universiti Utara Malaysia (UUM), Malaysia, in 2022. She is currently pursuing a Ph.D. degree in in computer science (AI) from the School of Computing, Universiti Utara Malaysia (UUM), Malaysia. Her research interests include feature selection, optimization, and machine learning. She can be contacted at email: manal_m_othman@ahsgs.uum.edu.my or manal.oshari@taiz.edu.ye.

Ku Ruhana Ku-Mahamud holds a Bachelor in Mathematical Sciences and a Master degree in Computing, both from Bradford University, United Kingdom in 1983 and 1986 respectively. Her Ph.D. in Computer Science was obtained from Universiti Pertanian Malaysia in 1994. As an academic, her research interests include ant colony optimization, pattern classification and vehicle routing problems. She can be contacted at email: ruhana@uum.edu.my.