IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 12 August 2023, accepted 28 August 2023, date of publication 30 August 2023, date of current version 7 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3310429

RESEARCH ARTICLE

BAOA: Binary Arithmetic Optimization Algorithm With K-Nearest Neighbor Classifier for Feature Selection

NIMA KHODADADI[1], (Member, IEEE), EHSAN KHODADADI[2], QASEM AL-TASHI[3,4], EL-SAYED M. EL-KENAWY[5], (Senior Member, IEEE), LAITH ABUALIGAH[6,7,8,9,10], (Senior Member, IEEE), SAID JADID ABDULKADIR[11], (Senior Member, IEEE), ALAWI ALQUSHAIBI[7], AND SEYEDALI MIRJALILI[12,13], (Senior Member, IEEE)

[1]Department of Civil and Architectural Engineering, University of Miami, Coral Gables, FL 33146, USA
[2]Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701, USA
[3]Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, Houston, TX 77094, USA
[4]Faculty of Administrative and Computer Sciences, Albaydha University, Albaydha, Yemen
[5]Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura 35111, Egypt
[6]Computer Science Department, Prince Hussein Bin Abdullah Faculty for Information Technology, Al al-Bayt University, Mafraq 25113, Jordan
[7]Department of Electrical and Computer Engineering, Lebanese American University, Byblos 1102-2801, Lebanon
[8]Hourani Center for Applied Scientific Research, Al-Ahliyya Amman University, Amman 19328, Jordan
[9]MEU Research Unit, Middle East University, Amman 11831, Jordan
[10]Applied science research center, Applied science private university, Amman 11931, Jordan
[11]Department of Computer and Information Sciences, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Malaysia
[12]Centre for Artificial Intelligence Research and Optimisation, Torrens University Australia, Brisbane, SA 4006, Australia
[13]University Research and Innovation Center, Obuda University, 1034 Budapest, Hungary

Corresponding authors: El-Sayed M. El-Kenawy (skenawy@ieee.org) and Nima Khodadadi (nima.khodadadi@miami.edu)

ABSTRACT The Arithmetic Optimization Algorithm (AOA) is a recently proposed metaheuristic algorithm that has been shown to perform well in several benchmark tests. The AOA is a metaheuristic that uses the main arithmetic operators' distribution behavior, such as multiplication, division, subtraction, and addition. This paper proposes a binary version of the Arithmetic Optimization Algorithm (BAOA) to tackle the feature selection problem in classification. The algorithm's search space is converted from a continuous to a binary one using the sigmoid transfer function to meet the nature of the feature selection task. The classifier uses a method known as the wrapper-based approach K-Nearest Neighbors (KNN), to find the best possible solutions. This study uses 18 benchmark datasets from the University of California, Irvine (UCI) repository to evaluate the suggested binary algorithm's performance. The results demonstrate that BAOA outperformed the Binary Dragonfly Algorithm (BDF), Binary Particle Swarm Optimization (BPSO), Binary Genetic Algorithm (BGA), and Binary Cat Swarm Optimization (BCAT) when various performance metrics were used, including classification accuracy, selected features as well as the best and worst optimum fitness values.

INDEX TERMS Feature selection, binary optimization, arithmetic optimization algorithm, classification.

I. INTRODUCTION

The volume of currently available data on the various resources has been overgrowing in current years [1]. Data mining is an analysis and knowledge-based process including different methods dealing with many applications such as machine learning, statistics, and databases [2]. These applications typically trade with a massive amount of data and complicated processes. Data mining aims to obtain knowledge from data represented in an understandable structure [3]. Many researchers have been interested in metaheuristic algorithms in recent years because of their potential to solve a variety of issues in various fields [4], [5], [6], [7], [8], [9], [10], [11]. High dimensional issues are also becoming more prevalent as the number of variables and complexity of the problems increases, and recently some

The associate editor coordinating the review of this manuscript and approving it for publication was Dominik Strzalka.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

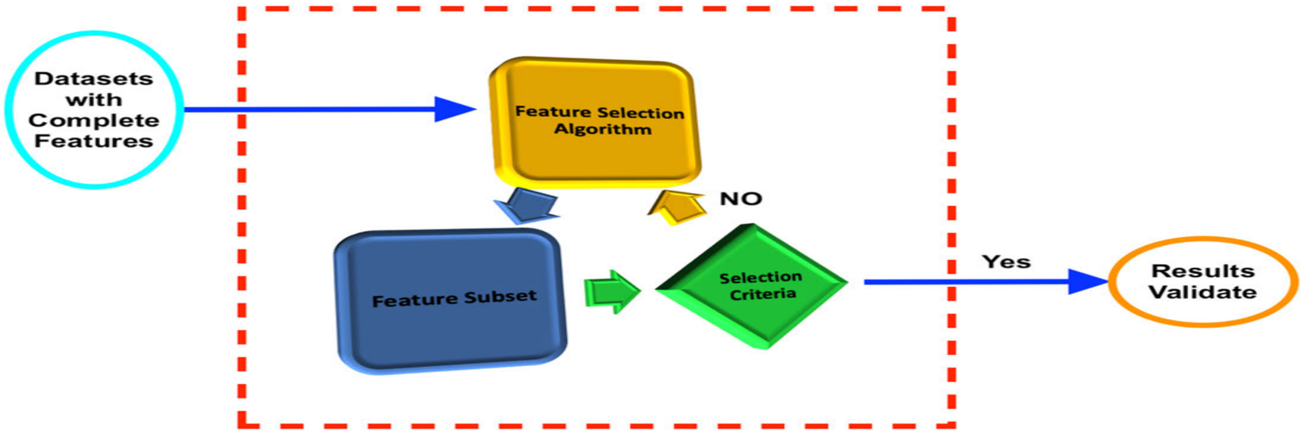


FIGURE 1. The process of feature selection.

metaheuristic such as Stochastic Paint Optimizer (SPO) [12], Battle Royale Optimization (BRO) [13], Mountain Gazelle Optimizer (MGO) [14], Dynamic Group-based Cooperative Optimization (DCGO) Algorithm [15], Advanced Neural Network Algorithm (ANNA) [16], Flow Direction Algorithm (FDA) [17], Dipper Throated Optimization (DTO) [18], and Waterwheel Plant Algorithm (WWPA) [19] have been presented for the solution of large-scale optimization issues.

In machine learning and computation, dimension reduction degrades the number of available features and creates dense copies of high-dimensional data [20]. Dimension reduction can be classified into two main methods; feature extraction and selection [21]. Moreover, high dimensional data have several problems, such as extensive computational time for model development, irrelevant data, and low performance; building data analysis is very challenging [22], [23]. Finding and eliminating unnecessary features is one of the basic data mining approaches that may be used to build a new set of features that correctly and effectively describe the problem (data) [24], [25]. This rule performs a critical task in many various intelligent systems and applications. It increases the classification effectiveness and decreases the computational expense of data science and analysis. Consequently, the feature selection process is essential in addressing the complicated classification problems [26], [27].

Pattern recognition, big data, classification, summarization, image mining, text categorization, clustering, regression tasks, machine learning, and other tasks involving complex, high-dimensional data all make use of Feature Selection (FS), which is frequently used in these and other contexts [28], [29]. High-dimensional datasets include an enormous collection of features, making the learning process more complex and diminishing the overall performance of the learned methods [30], [31], [32]. Regularly, feature selection removes unreliable features from the first dataset without losing the overall performance.

Feature selection techniques in machine learning involve two main actions: the search procedure and the subset evaluation. The search procedure navigates through the possible feature subsets, using methods like forward selection, backward elimination, and recursive feature elimination to optimize computational efficiency. Subset evaluation assesses the quality of candidate feature subsets using performance metrics or statistical tests while guarding against overfitting through cross-validation or separate validation sets. These methods need to be adapted and experimented with depending on the specific properties of the dataset and the problem at hand [33], [34], [35]. In the first part, the search approach uses various search techniques to choose subsets of new features. The method through which the feature selection process operates is depicted in Figure 1. According to this figure, the feature selection process consists of five main elements: the original dataset, discovery of the feature subset, feature subset evaluation, selection criteria, and validation. The latter part uses a classifier to assess the goodness of the newly generated subsets by the used search technique.

Feature selection methods utilized in machine learning include filters, wrappers, and embedded approaches [36], [37]. Figure 2 depicts the categorizations of feature selection. Filter Methods work without the aid of any learning algorithms. They provide a score that you may use to decide whether or not to keep a feature by basing your decision on the score, which is derived from ranking the features based on particular statistical indicators. Chi-square test, information gain, correlation coefficient scores, etc., are examples of common procedures. Although these techniques are frequently computationally effective, they do not consider the link between the model and the chosen features, which may result in less-than-ideal choices [38]. These techniques incorporate a machine learning model into the selection procedure. Based on the model's level of prediction accuracy, they assess a subset of attributes and decide how valuable they are. Examples include genetic algorithms, sequential

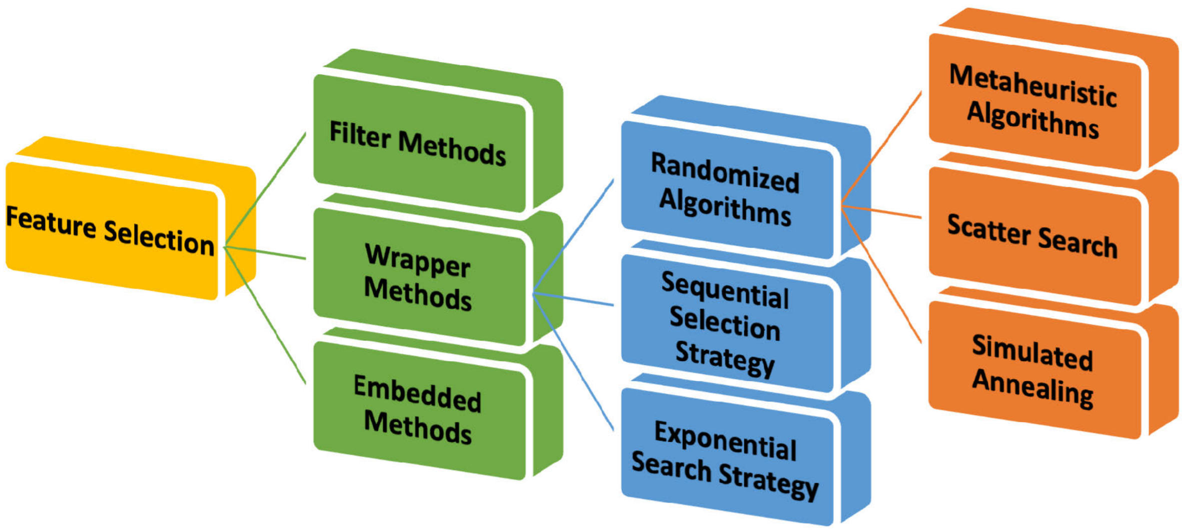


FIGURE 2. The categorization of feature selection.

feature selection, and recursive feature elimination. Wrapper methods can be expensive computationally even if they often perform better than filter approaches, especially when working with a large number of features [39]. Embedded Methods combine the advantages of filter and wrapper techniques. They are often more effective than wrapper techniques and include feature selection in the model training process. They take into consideration the dynamic relationship that exists between the models and the characteristics. Decision trees, LASSO regression, and Ridge regression are examples of algorithms that use embedded methods. Other examples of algorithms that use embedded methods include algorithms that derive from decision trees, such as Random Forest [40], [41]. Each of these approaches has pros and cons, and the best approach to adopt will frequently rely on the particular issue at hand, the dataset, and the available computational resources [42].

The No Free Lunch (NFL) [43] theorem highlights the intrinsic constraints of any algorithm's ability to address a comprehensive suite of optimization problems. This theory is notably relevant when formulating new binary metaheuristic algorithms because it signifies that no single algorithm can ensure exceptional performance on all potential binary optimization problems. Thus, scientists might conceive new metaheuristic algorithms custom-built for binary optimization's distinct features and needs. These specialized algorithms strive for superior results in binary problems, recognizing that they may not be the optimal solution for other optimization issues.

This paper focuses on a new optimizer, the Arithmetic Optimization Algorithm (AOA), in machine learning-based feature selection applications. The AOA was proposed recently by [44] inspired by the mathematical operators in science. The specified algorithm demonstrates satisfactory efficacy in resolving high-dimensional, complex problems. As pointed out in reference [44], this rapid, sturdy algorithm is highly competent in maintaining a balance between exploration and exploitation phases, and it exhibits accelerated convergence rates along with successful evasion of local optima. This leads to accurate solutions for both unimodal and multimodal benchmark functions. Furthermore, the AOA is capable of identifying the optimal solution swiftly, rendering it an attractive choice for researchers aiming to tackle FS problems. The AOA operates on continuous values within a continuous search space in its standard form. However, FS is a binary optimization problem with decision variables having a different structure, limited to values "0" and "1". A binary version of the AOA has been proposed to adapt the algorithm for FS problems to address this. It demonstrated promising results in solving other problems, which is the primary motivation behind applying it in this domain. As mentioned above, the best methods used to solve the selection problems are the heuristics, and the current results request a deep investigation to find a new approach that can find the best subset of features while increasing the performance of the underlying process, smaller number of features and reduce the execution time. Thus, the following demonstrates the main contribution and novelty of this study:

- A binary version of the AOA is developed to solve various feature selection problems, including medical, artificial, games, and biological data.
- The KNN classifier is adopted to evaluate the goodness of the selected features.
- To ensure that the selection methods work as intended, a total of eighteen public datasets are examined.
- Three standard evaluation measures aid the proposed method's performance.
- The proposed binary AOA was evaluated against several well-known existing methods and was found to perform better.

The following are the remaining sections of this paper: The related works are discussed in more detail in Section II. Section III present the proposed feature selection method-based Arithmetic Optimization Algorithm. The experiment setup is shown in Section IV, followed by results and conclusions in Section V. Finally, in Section VI, the conclusions and future work are discussed.

II. LITERATURE REVIEW

Metaheuristic optimizers are the most used methods of addressing feature selection problems, so the most related studies are presented. Common feature selection optimizers are Gradient-based Optimizer [45], Grey Wolf Optimizer [46], [47], [48], Multi-verse Optimizer [49], Competitive Swarm Optimizer [50], Genetic Algorithm [51], [52], Stochastic Fractal Search [53], Dipper Throated Optimization [54], [55], Cat Swarm Optimization Algorithm [56], Firefly Algorithm [57], Dragonfly Algorithm [58], Krill Herd Optimizer [59], Sine Cosine Optimizer [60], Moth-Flame Optimization Algorithm [61], Whale Optimization Algorithm and Harris Hawks Optimization [62] and Bat Optimizer [63]. Recently, Zakeri and Hokmabadi [40] presented a new feature selection approach based on an analytical cooperation model between grasshoppers in discovering food sources. The significance of the proposed method was demonstrated by comparing it to other well-known selection methods. A binary version of monarch butterfly optimization was proposed [64]; the KNN is used to evaluate the goodness of the selected features and evaluated on nineteen small-size and seven large datasets. The results demonstrated better classification accuracy compared to other existing technologies. Reference [65] introduced a binary version of the artificial algae algorithm, which was evaluated on 25 datasets collected from UCI. The results show good performance compared to other existing approaches. Reference [66] proposed a hyperlearning binary dragonfly algorithm to tackle the feature selection problem in coronavirus disease; this method was also tested on 21 UCI datasets. The finding shows an improvement in the classification accuracy and a reduction in the feature number. Reference [67] introduced a feature selection approach based on a binary crow search algorithm and evaluated it against 20 public datasets. The outcomes show a better performance comparing the state-of-the-art methods. A review of the modification strategies of the nature-inspired algorithms for feature selection problems is recently published by [68]. In 2020, [69] presented a new dimension reduction method based on the Henry gas solubility optimizer for choosing meaningful features to improve the classification rate. The experimental study recommended that the proposed method is significantly proper on various high-dimensional data. Reference [70] introduced a multi-objective search method for addressing the feature selection problems based on continuous and binary forest optimizers applying the archive and region-based selection theories. Compared to the binary strategy, a continuous approach has been demonstrated to be preferable, and it reduced the classification rate by choosing fewer features than other methods. Reference [71] proposed a three-optimizer parallel mixed-group feature selection. Based on the findings, it appears that the suggested technique has the potential to improve forecast accuracy and execution time.

Improved PSO for tackling complex feature selection problems has been proposed by [72]. In addition, a new exploration method has been incorporated with the basic optimizer to improve its capability to handle the given problems. The findings confirmed that the suggested method is the best compared to other methods in terms of text classification and the confidence of various dimensions. Reference [73] presented a catfish binary particle swarm optimizer to tackle the weaknesses of the original optimizer and solve the feature selection problems. The results revealed that the proposed method simplifies the procedure efficiently and achieved higher classification precision compared to other methods. Reference [74] presented a hybrid feature selection technique based on using an enhanced whale optimizer. The results demonstrated that the proposed method's classification precision outperforms other well-known selection methods. The problem of feature selection was addressed in [75] by proposing a new wrapper-based selection method based on the harris hawks optimizer, simulated annealing, and a Bitwise mechanism. The proposed method obtained better results and provided a better ability to solve the given problems compared to other selection methods. The following section discusses the methodology, and the proposed binary version of AOA is discussed in detail.

III. METHODOLOGY

The original AOA and the proposed BAOA are both briefly discussed in this section.

A. ARITHMETIC OPTIMIZATION ALGORITHM (AOA)

In 2021, [44] devised this algorithm by combining several mathematical equations and operators. The AOA algorithm, like other metaheuristics, begins with a random sample of solutions. The objective value of each solution is calculated after each iteration. There are two controlling parameters in this algorithm called Math Optimizer Accelerated (MOA) and Math Optimizer probability (MOP) that should be updated prior to updating the position of solutions as follows:

$$MOA(t) = \mathrm{Min}_{MOA} + t \times \left(\frac{Max_{MOA} - \mathrm{Min}_{MOA}}{T} \right) \quad (1)$$

$$\mathrm{MOP}(t) = 1 - \left(\frac{t^{\frac{1}{\alpha}}}{T^{\frac{1}{\alpha}}} \right) \quad (2)$$

where t represents the current iteration, T defines the maximum iteration, Max/Min denote the maximum and minimum values used to constrain the MOA, and *alpha* indicates a controlling parameter.

After updating MOA and MOP, a random number is generated called r_1 to switch between exploration and exploitation, which actually identify the feasible solution. For exploration, the following equation is used:

$$x_{i,j}(t+1) = \begin{cases} \dfrac{\text{best}\left(x_j\right)}{MOP+\epsilon} \\ \times\left(UB_j - LB_j\right)\times\mu + LB_j & \text{if } r_2 < 0.5 \\ \text{best}\left(x_j\right)\times MOP \\ \times\left(UB_j - LB_j\right)\times\mu + LB_j & \text{if } r_2 \geq 0.5 \end{cases} \tag{3}$$

where t is the current iteration, μ is a controlling parameter, ϵ is a small number to avoid division by 0, and r_2 is a random number in [0, 1].

For exploitation, the following equation is used:

$$x_{i,j}(t+1) = \begin{cases} \text{best}\left(x_j\right) - MOP \\ \times\left(UB_j - LB_j\right)\times\mu + LB_j & \text{if } r_3 < 0.5 \\ \text{best}\left(x_j\right) + MOP \\ \times\left(UB_j - LB_j\right)\times\mu + LB_j & \text{if } r_3 \geq 0.5 \end{cases} \tag{4}$$

Complete details of the AOA algorithm's main inspiration and mathematical model can be found in [44].

B. BINARY ARITHMETIC OPTIMIZATION ALGORITHM (BAOA)

Feature selection is naturally known to be a discrete (Binary) problem. As a result, the original AOA introduced in Section III-A cannot simply be utilized to handle such problems. This paper proposes the binary variant of the AOA to be fit for the nature of the feature selection task. Candidates in the original AOA [44] can roam the search space continuously because their position vectors have a continuous real value. It is necessary to adjust Eq. (3) in order for candidates to work with binary data, as shown in the following equation:

$$B_1\left(x_{i,j}(t+1)\right) = \begin{cases} 1 & \text{if sigmoid}\left(x_{i,j}(t+1)\right) \geq \text{ rand} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $x_{i,j}\,(t+1)$ represents the original exploration stage presented in Eq. (3), $B_1(x_{i,j}\,(t+1))$ is an updated binary position of the exploration phase at iteration t, a rand is a number generated at random from the distributions uniform $\in$ [1, 0], and *sigmoid* (x) Binary transfer function converts continuous search space to binary. Using the sigmoid function, any input value can be translated mathematically into an output between 0 and 1. Its S-shaped curve makes it a suitable activation function for neural networks and machine learning. The output approaches 0 when the input is extremely negative and 1 when the input is extremely positive. Because of this characteristic, the sigmoid function is useful for tasks like binary classification, in which we want to forecast the probability of an event occurring or not occurring. In mathematics, sigmoid can be described as follows:

$$\text{sigmoid}(x) = \frac{1}{1+e^{-10(x-0.5)}} \tag{6}$$

Similarly, the position updating of the exploitation phase in Eq. (4) can be amended into the following equation:

$$B_2\left(x_{i,j}(t+1)\right) = \begin{cases} 1 & \text{if sigmoid}\left(x_{i,j}(t+1)\right) \geq \text{ rand} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $x_{i,j}\,(t+1)$ represents the original exploitation stage presented in Eq. (4) and $B_2(x_{i,j}\,(t+1))$ is a modified binary position of the exploitation phase at iteration t, rand is a number generated at random from the distributions uniform $\in$ $[1, 0]$, and $sigmoid(x)$ is calculated as in Eq. (6).

Algorithm 1 Pseudo-Code of the Proposed BAOA

Load dataset, divided into training and testing sets
AOA' parameters initialization α, μ
Randomly initialize the positions of solutions (i=1, …, n) *LOOP Process*
while $(t < T)$ **do**
  Calculate the objective function (fitness) as in Eq. (8)
  Get the finest solution (obtained finest so far).
  Modify MOA values utilizing Eq. (1).
  Modify MOP values using Eq. (2).
  for i = 1 to Solutions **do**
    for j = 1 to Solutions **do**
      Produce values randomly within the range [0, 1] (r1, r2, and r3)
      if $r1 > MOA$ **then**
        Exploration stage
        if $r2 > 0.5$ **then**
          a) The division operation is utilized
          Modify the positions of the ith solutions utilizing the 1st rule in Eq. (3).
          Binarize the positions of the 1st rule of the ith solutions utilizing Eq. (5).
        else
          b) The multiplication operation is utilized
          Modify the positions of the ith solutions utilizing the 2nd rule in Eq. (3).
          Binarize the positions of the 2nd rule of the ith solutions utilizing Eq. (5).
        end if
      else
        Exploration stage
        if $r3 > 0.5$ **then**
          a) The subtraction operation is utilized
          Modify the positions of the ith solutions utilizing the 1st rule in Eq. (4).
          Binarize the positions of the 1st rule of the ith solutions utilizing Eq. (7).
        else
          b) The addition operation is utilized
          Modify the positions of the ith solutions utilizing the 2nd rule in Eq. (4).
          Binarize the positions of the 2nd rule of the ith solutions utilizing Eq. (7).
        end if
      end if
    end for
  end for
  t=t+1
end while
Return the finest solution (minimum features, maximum accuracy).

It should be noted that the solution representation in this study is demonstrated in a vector with one dimension. The length of this vector donates the number of features. This binary vector's 0 and 1 values are as follows: Nothing has been selected when you see a 0 value. 1: a feature has been selected. By definition, the feature selection issue is a bi-objective problem with two goals, one is to obtain the smallest number of features possible, and the second goal is to improve classification accuracy. In order to consider both goals, the following fitness function is utilized where

TABLE 1. The details of datasets.

No.	Dataset	Instances	No. Features	No. Classes
1	Breastcancer	699	9	2
2	BreastEW	569	30	2
3	CongressEW	435	16	2
4	Exactly	1000	13	2
5	Exactly2	1000	13	2
6	HeartEW	270	13	2
7	IonosphereEW	351	34	2
8	KrvskpEW	3196	36	2
9	Lymphography	148	18	2
10	M-of-n	1000	13	2
11	PenglungEW	73	325	2
12	SonarEW	208	60	2
13	SpectEW	267	22	2
14	Tic-tac-toe	958	9	2
15	Vote	300	16	2
16	WaveformEW	5000	40	3
17	WineEW	178	13	3
18	Zoo	101	16	6

the K-nearest neighbor classifier (KNN) is used in this study:

$$\text{fitness } = \alpha\rho_R(Y) + \beta\frac{|K|}{|N|} \tag{8}$$

where $\rho_R(Y)$ denotes the error rate of the KNN classifier, $\alpha = [0, 1]$ and $\beta = (1-\alpha)$ they are parameters, Besides, $|K|$ is the number of features selected and $|N|$ is the original number of features in the dataset.

The Pseudo-code of the proposed BAOA is outlined in Algorithm 1. Figure 3 illustrates BAOA's general methodology.

IV. EXPERIMENTATION

In this section, the description of the datasets used, the parameter settings, and the evaluation measures are clearly presented.

A. DATASETS

The proposed Binary Arithmetic Optimization Algorithm is validated using 18 benchmark datasets from the UC Irvine Machine Learning Repository [76]. They are the same data sets used by many researchers to evaluate various feature selection approaches. A total number of instances ranging from 73 to 5000 is included in this study for datasets.Table 1 lists the datasets in detail.

B. PARAMETER SETTINGS

The performance of BAOA is compared against four well-known such as BDF, BPSO, BGA, and BCAT. For assessment in this work, each dataset is separated using cross-validation similar to the one used in [77]. $K-1$ folds are applied for validation and training in cross-approval K-overlap, whereas the overlay's remainder is applied for testing. Note that, N times, the suggested method has been repeated. The proposed method is then tested $k \times M$ times across all datasets. Data for both validation and training are approximated in the same way. The setting of parameters utilized in BDF, BPSO, BGA, and BCAT are similar to their own parameter settings. MATLAB R2021a is used to code the algorithms, and all datasets are run on a computer with a 2.3 GHz 8-Core Intel Core i9 processor, 16 GB of DDR4 RAM, and MacOS Big Sur as the operating system for 20 independent runs with a population of 10 and 100 iterations were set for each algorithm to have a fair comparison. Remarkably, following the execution of multiple experiments, we ascertained that the optimal values for iterations and runs were 100 and 20, respectively. It was observed that augmenting the number of iterations and runs did not yield discernible enhancements in outcomes. The following are the parameters of BAOA for this study:

- Number of populations is equal to 10.
- Number of iterations is equal to 100.
- $Min_{MOA} = 0.2$, $Max_{MOA} = 1$
- μ=0.49, α =5

C. EVALUATION MEASURES

Validation, training, and testing data are randomly divided into three equal parts. Separating the data is repeated several times to ensure strong and measurable results. In each run, the validation data is used to test the following statistical measures:

1) MEAN OF CLASSIFICATION ACCURACY (MCA)

When the algorithm is run N times, this metric shows how accurate a classifier is given a set of characteristics and the following formula is used to determine it:

$$MCA = \frac{1}{N}\sum_{k=1}^{N} AvgAcc^k \tag{9}$$

where $AvgAcc^k$ is the average gain in accuracy at k run.

2) AVERAGE OF SELECTED FEATURE (ASF)

when the algorithm is performed N times, this measure reveals the ratio of the average picked features to the total features and it is computed as follows:

$$ASF = \frac{1}{N}\sum_{k=1}^{N} \frac{Avg\text{ Selection}^{\,k}}{M} \tag{10}$$

where $AvgSelection^k$ denotes the average features chosen at run k, and M denotes the total number of features in the dataset.

3) MEAN OF FITNESS FUNCTION (MFF)

After N iterations, this metric represents the average fitness function value, and The following formula is used to determine it:

$$MFF = \frac{1}{N}\sum_{k=1}^{N} g_k^* \tag{11}$$

g_k^* is the average amount of fitness gained during run K.

FIGURE 3. The methodology of the proposed BAOA.

4) BEST FITNESS FUNCTION (BFF)

To determine the fitness function's minimum value after N iterations, this indicator is calculated as follows:

$$BFF = \text{Min}_k\ g_k^* \tag{12}$$

where g_k^* denotes the minimum fitness value obtained during run k.

5) WORST FITNESS FUNCTION (WFF)

It is a numerical representation of the fitness function's maximum value obtained after performing the operation N times and is determined using the formula:

$$WFF = \text{Max}_k\ g_K^* \tag{13}$$

where g_k^* denotes the worst fitness value obtained during run k.

6) AVERAGE OF COMPUTATIONAL TIME (ACT)

It's a measure of the computation time gained in seconds after N iterations of the operation have been completed and are

TABLE 2. The outcomes of the proposed BAOA method.

Datasets No.	Average Accuracy	Mean Fitness	Feature Selected	Computational time(s)
1	0.98	0.02	3.60	5.08
2	0.98	0.03	9.65	5.11
3	0.97	0.03	1.26	3.80
4	0.84	0.01	3.95	3.84
5	0.78	0.22	1.15	3.52
6	0.85	0.18	5.00	4.59
7	0.92	0.10	5.30	4.58
8	0.96	0.02	7.50	11.24
9	0.86	0.16	4.80	4.84
10	0.92	0.03	5.92	5.85
11	0.96	0.08	15.20	5.52
12	0.83	0.12	8.90	5.24
13	0.86	0.17	3.60	3.45
14	0.82	0.19	5.09	5.80
15	0.98	0.02	2.63	4.53
16	0.86	0.18	11.40	18.04
17	0.98	0.01	4.35	5.08
18	0.96	0.02	5.05	4.90
Average	0.91	0.09	5.80	5.83

TABLE 3. The Classification accuracy of the proposed BAOA in comparison to different methods.

Datasets No.	BAOA	BDF	BPSO	BGA	BCAT
1	**0.98**	0.96	0.95	0.96	0.97
2	**0.98**	0.96	0.94	0.94	0.94
3	**0.97**	0.96	0.94	0.94	0.95
4	0.84	**0.98**	0.68	0.67	0.96
5	**0.78**	0.74	0.75	0.76	0.75
6	**0.85**	0.83	0.78	0.82	0.81
7	0.92	**0.93**	0.84	0.83	0.92
8	**0.96**	0.95	0.94	0.92	0.95
9	0.86	0.87	0.69	0.71	**0.89**
10	0.92	**0.99**	0.86	0.93	**0.99**
11	**0.96**	0.89	0.72	0.70	0.91
12	0.83	**0.91**	0.74	0.73	0.90
13	**0.86**	0.85	0.77	0.78	0.83
14	**0.82**	0.78	0.73	0.71	0.80
15	**0.98**	0.95	0.89	0.89	0.95
16	**0.86**	0.75	0.76	0.77	0.78
17	**0.98**	0.92	0.95	0.93	0.83
18	0.96	0.98	0.83	0.88	**0.97**
Average	**0.91**	0.90	0.82	0.83	0.89

TABLE 4. The average selected features of the proposed BAOA in comparison to different methods.

Datasets No.	BAOA	BDF	BPSO	BGA	BCAT
1	**3.60**	4.95	5.7	5.09	5.85
2	**9.65**	14.85	16.6	16.35	14.35
3	**1.26**	4.60	6.8	6.62	6.15
4	**3.95**	6.09	9.8	10.82	6.5
5	**1.15**	2.70	6.2	6.18	5.5
6	**5.00**	6.85	7.9	9.49	6.5
7	**5.30**	11.49	19.2	17.31	14.05
8	**7.50**	17.74	20.8	22.43	20.15
9	**4.80**	8.15	9.0	11.05	9.1
10	**5.92**	6.04	9.1	6.83	6.6
11	**15.20**	123.5	178.8	177.13	152.9
12	**8.90**	27.48	31.2	33.30	27.95
13	**3.60**	7.94	12.5	11.75	10.6
14	**5.09**	5.94	6.6	6.85	6.3
15	**2.63**	4.14	8.8	6.62	5.95
16	**11.40**	20.96	22.7	25.28	14.05
17	**4.35**	6.30	8.4	8.63	6.40
18	**5.05**	5.69	9.7	10.11	7.25
Total	**104.35**	285.47	389.8	391.84	326.15

TABLE 5. The mean fitness function of the proposed BAOA in comparison to different methods.

Datasets No.	BAOA	BDF	BPSO	BGA	BCAT
1	**0.02**	0.04	0.03	0.03	0.03
2	**0.03**	0.04	**0.03**	0.04	0.06
3	**0.03**	0.04	0.04	0.04	0.04
4	0.02	**0.01**	0.28	0.28	0.04
5	**0.22**	0.25	0.25	0.25	0.25
6	0.18	0.17	0.15	**0.14**	0.18
7	0.10	**0.07**	0.14	0.13	0.08
8	**0.02**	0.03	0.05	0.07	0.05
9	0.16	**0.13**	0.19	0.17	0.15
10	0.03	**0.01**	0.11	0.08	**0.01**
11	**0.08**	0.09	0.22	0.22	0.09
12	0.12	**0.07**	0.13	0.13	0.10
13	0.17	0.14	**0.13**	0.14	0.17
14	**0.19**	0.21	0.24	0.24	0.20
15	**0.02**	0.04	0.05	0.05	0.05
16	**0.18**	0.24	0.22	0.20	0.22
17	**0.01**	0.02	0.02	**0.01**	0.02
18	**0.02**	0.04	0.10	0.08	**0.02**
Total	**1.59**	1.62	2.38	2.3	1.76

determined using the formula:

$$ACT = \frac{1}{N}\sum_{k=1}^{N} AvgCT^{k} \tag{14}$$

where $AvgCT^k$ is the value of computing time acquired during run k.

V. RESULTS AND DISCUSSION

This section presents the outcomes of the comparison approaches in accordance with the metrics employed. The proposed BAOA's results in terms of Average Accuracy, Mean Fitness, Selected Features, and Computational Time are shown in Table 2. This table can help in making fair comparisons according to the given results. For instance, the proposed BAOA, in dataset 1, got 0.98 average accuracies, 0.02 mean fitness, 3.60 selected features, and 5.08 time. Another example, in dataset 2, the proposed BAOA got 0.98 average accuracies, 0.03 mean fitness, 9.65 chosen features, and 5.11 time.

Table 3 shows the classification accuracy values for the comparative methods (BDF, BPSO, BGA, and BCAT) using 18 datasets. The proposed BAOA got superior results by getting 12 best out of 18 cases, followed by BDF getting four best results out of 18, and BCAT getting three best out of 18 cases. It is obvious that the proposed technique (BAOA) is capable of discovering new best results, and the results demonstrated the proposed method's ability to solve feature section problems. Moreover, the results illustrated a clear difference between the comparative methods, especially the proposed BAOA.Table 3 shows that the suggested BAOA had the best overall average classification accuracy result (0.91), followed by BDF (0.90), BCAT (0.89), BGA (0.83), and finally, BPSO (0.82).

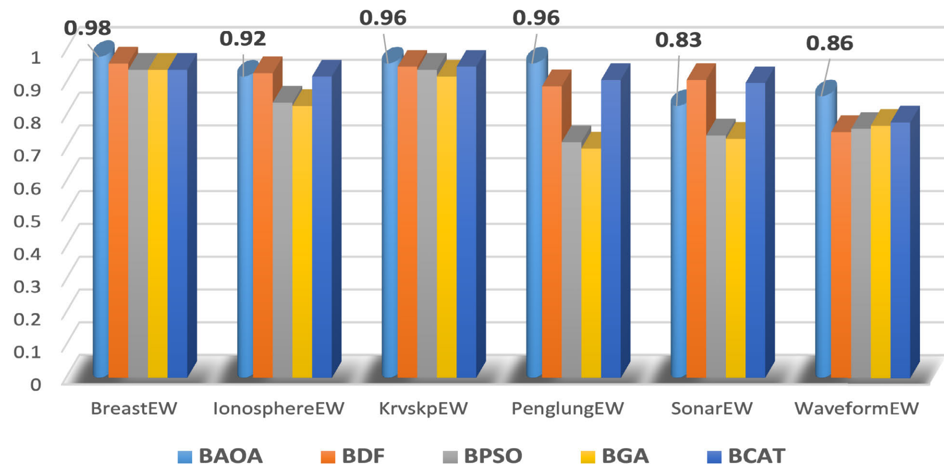


FIGURE 4. Comparison of average accuracy in large data.

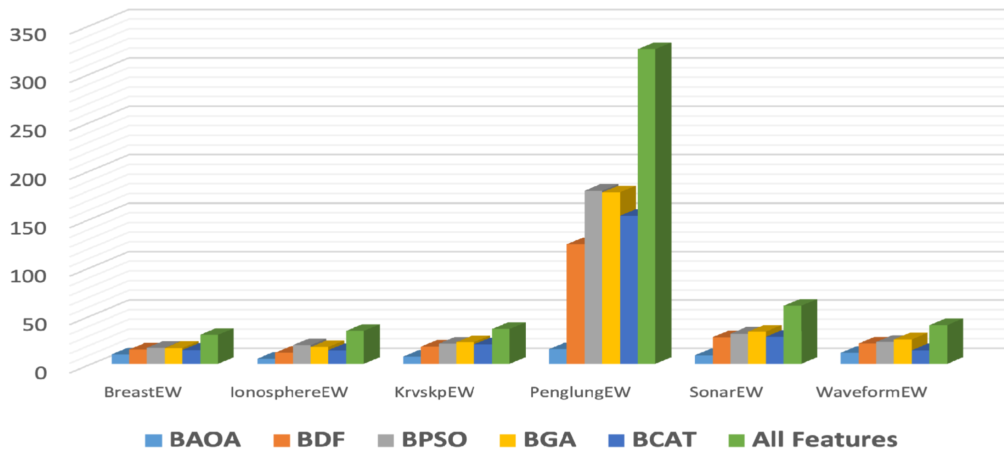


FIGURE 5. Comparison of selected features in large data.

Table 4 shows the average selected features by the comparative methods (BDF, BPSO, BGA, and BCAT) using 18 datasets. The proposed BAOA got superior results by getting the smallest features in all the tested cases. These are promising results, and it is clear that the proposed BAOA got the number of selected features in these problems efficiently. For all examined instances, the proposed method (BAOA) was found to find new best solutions, proving its effectiveness in solving feature section problems. Furthermore, the findings revealed a notable distinction between the comparison techniques, particularly the suggested BAOA.Table 4 shows that the suggested BAOA had the best overall average selected features result of (104.35), followed by BDF (285.47), BCAT (326.15), BGA (389.8), and finally, BPSO (391.84).

Table 5 shows the mean fitness function values obtained by the comparative methods (BDF, BPSO, BGA, and BCAT) using 18 datasets. The proposed BAOA got excellent results by getting the best fitness function values in 11 out of 18 test cases, followed, BDF got 5 best results out of 18 test cases, BPSO got 2 best results out of 18 test cases, BGA got 2 best results out of 18 test cases, and BCAT got 1 best result out of 18 test cases. This is a good outcome, and it is apparent that the suggested BAOA successfully obtained the desired number of chosen features in these cases. It is evident that the presented technique (BAOA) can find new best outcomes in all of the tested situations, and the results demonstrated the proposed BAOA's capacity to solve feature selection issues. Furthermore, the results indicated a significant difference between the comparison approaches, notably the proposed BAOA. Table 5shows that the suggested BAOA had the best overall average Mean fitness function result (1.59), followed by BDF (1.62), BCAT (1.76), BGA (2.3), and finally, BPSO (2.38).

Table 6 shows the best fitness function values obtained by the comparative methods (BDF, BPSO, BGA, and BCAT) using 18 datasets. Again, the suggested BAOA had the best fitness function values in 10 of the 18 test instances that were

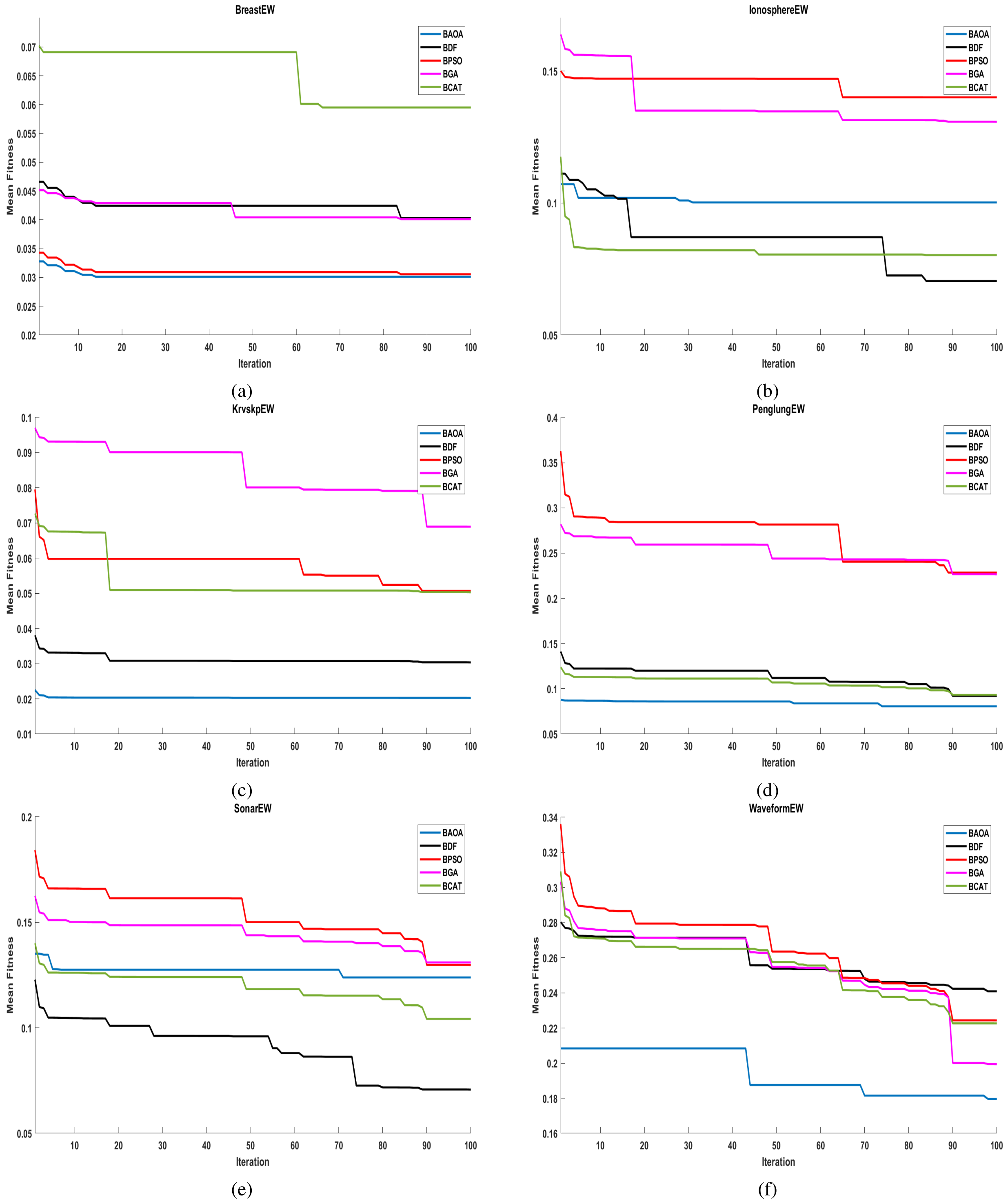


FIGURE 6. Convergence curves for BAOA and other state-of-art methods for large datasets.

tested, followed, BDF got five best results out of 18 test cases, BGA got four best results out of 18 test cases, BPSO got three best results out of 18 test cases, and BCAT got three best result out of 18 test cases. This is an excellent indication, and it's clear that the proposed BAOA was influential in obtaining the optimal number of selected features in these situations. In all of the examined datasets, it is clear that the given approach (BAOA) can identify new best outcomes, and the findings proved the suggested BAOA's ability to handle feature section challenges. In addition, the results revealed a substantial difference between the comparative techniques, particularly the suggested BAOA.Table 6 shows that the suggested BAOA

TABLE 6. The best fitness function of the proposed BAOA in comparison to different methods.

Datasets No.	BAOA	BDF	BPSO	BGA	BCAT
1	**0.01**	0.03	0.03	0.02	0.03
2	**0.02**	0.03	**0.02**	**0.02**	0.06
3	**0.02**	**0.02**	0.03	0.03	0.04
4	0.01	0.01	0.21	0.27	**0.00**
5	**0.20**	0.23	0.22	0.22	**0.20**
6	0.14	0.15	0.13	**0.12**	0.16
7	0.06	**0.04**	0.12	0.09	0.05
8	**0.01**	0.02	0.03	0.03	0.04
9	0.12	**0.07**	0.14	0.12	0.12
10	**0.00**	0.01	0.06	0.02	0.00
11	**0.02**	0.03	0.13	0.13	0.07
12	0.09	**0.03**	0.07	0.07	0.08
13	0.13	0.11	**0.10**	0.12	0.16
14	0.18	0.18	0.21	0.21	**0.17**
15	**0.01**	0.02	0.03	0.03	0.04
16	**0.16**	0.23	0.21	0.19	0.21
17	**0.00**	0.01	**0.00**	**0.00**	0.02
18	0.01	**0.00**	0.03	**0.00**	0.01
Total	**1.19**	1.20	1.77	1.69	1.46

TABLE 7. The worst fitness function of the proposed BAOA in comparison to different methods.

Datasets No.	BAOA	BDF	BPSO	BGA	BCAT
1	**0.03**	0.05	**0.03**	0.04	0.04
2	**0.04**	0.05	0.05	0.05	0.08
3	**0.04**	0.05	**0.04**	0.06	0.06
4	0.03	**0.02**	0.32	0.31	0.31
5	**0.23**	0.29	0.31	0.30	0.28
6	0.21	0.20	0.18	**0.14**	0.21
7	0.11	**0.09**	0.17	0.16	**0.09**
8	**0.04**	**0.04**	0.07	0.13	0.06
9	0.18	**0.17**	0.27	0.27	0.17
10	0.04	**0.01**	0.16	0.15	**0.01**
11	**0.11**	0.17	0.29	0.29	0.13
12	0.19	**0.11**	0.22	0.23	0.12
13	0.18	0.24	0.16	**0.15**	0.20
14	0.23	**0.07**	0.27	0.26	0.25
15	**0.05**	0.26	0.08	0.08	0.06
16	0.21	**0.04**	0.23	0.21	0.22
17	**0.03**	0.14	**0.03**	**0.03**	0.04
18	0.04	0.05	0.21	0.18	**0.03**
Total	**1.95**	1.97	3.09	3.04	2.36

had the best overall average best fitness function result (1.19), followed by BDF (1.20), BCAT (1.46), BGA (1.69), and finally, BPSO (1.77).

Table 7 shows the worst fitness function values obtained by the comparative methods (BDF, BPSO, BGA, and BCAT) using 18 datasets. It was shown that the proposed BAOA performed better in most of the tests based on the worst fitness values; it got the best worst-fitness function values in 8 out of 18 test cases, followed, BDF got eight best results out of 18 test cases, BGA got three best results out of 18 test cases, BPSO got three best results out of 18 test cases, and BCAT got three best result out of 18 test cases. Table 7 shows that the suggested BAOA had the best overall average worst-fitness function result (1.95), followed by BDF (1.97), BCAT (2.36), BGA (3.04), and finally, BPSO (3.09).

TABLE 8. Comparison between the proposed BAOA and state-of-the-art methods in term of computational times (Seconds).

Datasets	BAOA	BDF	BPSO	BGA	BCAT
Breast cancer	**5.08**	6.23	5.84	6.34	5.89
BreastEW	**5.11**	5.67	5.83	6.02	5.76
CongressEW	**3.80**	5.45	5.98	6.12	5.48
Exactly	**3.84**	5.76	6.11	6.23	5.79
Exactly2	**3.52**	5.65	5.91	5.99	5.74
HeartEW	**4.59**	5.32	5.88	5.89	5.43
IonosphereEW	**4.58**	5.56	5.78	5.91	5.53
KrvskpEW	**11.24**	20.32	23.45	27.87	21.54
Lymphography	**4.84**	5.66	6.12	6.43	5.58
M-of-n	**5.85**	5.76	5.99	6.65	5.66
PenglungEW	**5.52**	4.76	5.11	5.23	5.78
SonarEW	**5.24**	5.49	5.89	6.22	5.66
SpectEW	**3.45**	5.77	5.89	6.03	5.82
Tic-tac-toe	**5.80**	5.39	5.66	6.21	5.38
Vote	**4.53**	5.78	5.89	5.94	5.81
WaveformEW	**18.04**	37.23	41.21	48.22	38.45
WineEW	**5.08**	5.88	6.11	6.43	5.91
Zoo	**4.90**	5.43	5.87	5.98	5.55
Total	**105.01**	147.11	158.52	173.71	150.76

The simulation time of BAOA compared to the existing methods is shown in Table 8, inspecting the outcomes, the BAOA is the best approach that has the lowest simulation time among others. The total simulation time was 105.01, 147.11, 158.52, 173.71, and 150.76 seconds for BAOA, BDF, BPSO, BGA, and BCAT, respectively. The proposed BAOA method clearly saves processing time substantially. The statistical data clearly show a robust feature selection procedure when using BAOA. The suggested BAOA displayed an impressive ability to choose fewer attributes while maintaining accuracy. Furthermore, across all 18 datasets tested, the proposed BAOA requires much less time in seconds to calculate than other methods.

Figure 4 shows the results of the comparative methods using six large datasets in terms of average accuracy. The proposed method (BAOA) got better results in all test cases, as shown in this figure when testing the large datasets.

Figure 5 shows the comparative methods results using six large datasets in terms of the number of selected features. The proposed method (BAOA) got the smallest number of selected features in all test cases, as shown in this figure when testing the large datasets. The number of selected features is an attractive indicator to prove the proposed method's ability.

Figure 6 compares the convergence curve of the proposed BAOA to other current methods. As can be observed from the figure, the BAOA performs better than all algorithms while handling the majority of big datasets. On several datasets, including the BreastEW, PenglungEW, WaveformEW, and KrvskpEW datasets, the behaviors of the BCAT, BDF, BPSO, and BGA algorithms may be seen to exhibit premature convergence. Based on the findings made thus far, it is possible to infer that the BAOA has the capacity to strike a good balance between the exploration and exploitation stages. As a result, the algorithm's premature convergence and inactivity issues are significantly alleviated as compared

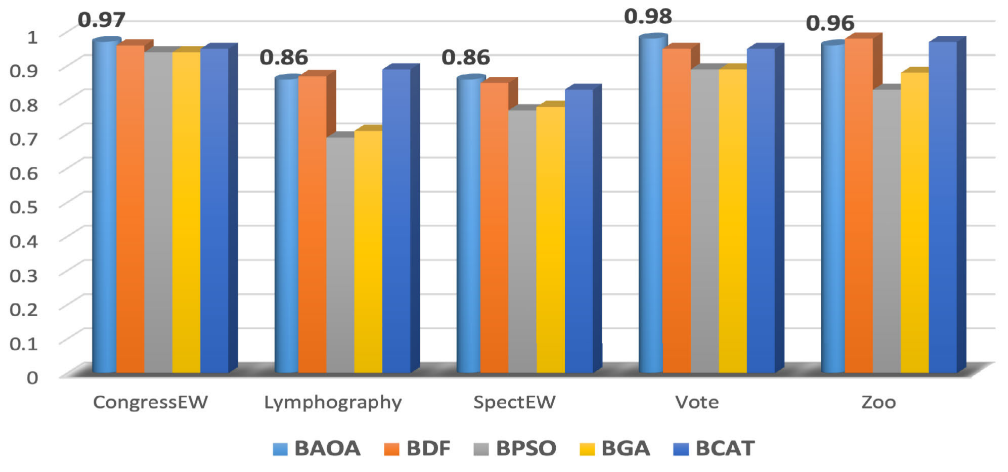


FIGURE 7. Comparison of average accuracy in medium data.

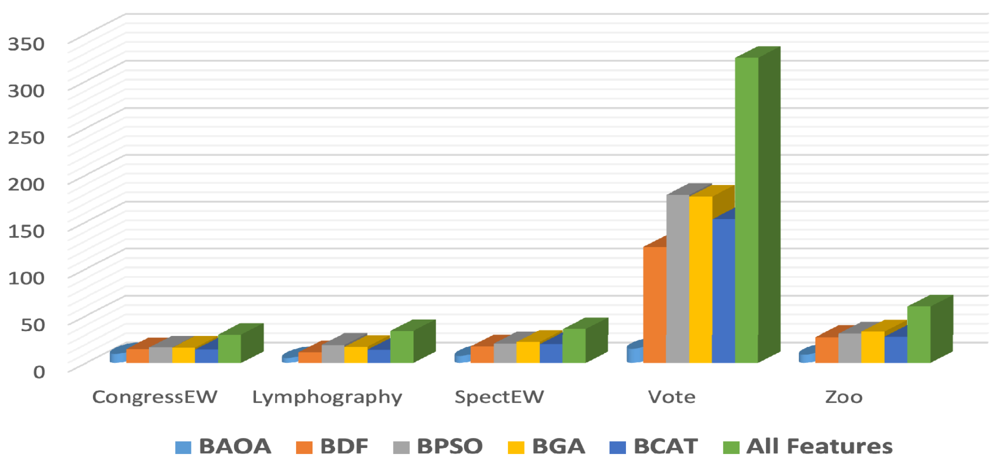


FIGURE 8. Comparison of selected features in medium data.

to the BCAT, BDF, BPSO, and BGA optimizers. Figure 7 depicts the average accuracy of the comparative methods applied to five medium-sized datasets. When testing medium datasets, the proposed method (BAOA) achieved better results in all test cases, as shown in this figure.

Using five medium datasets with varying numbers of selected features, the results of comparative methods are shown in Figure 8. This figure shows that the proposed method (BAOA) generated the fewest selected features in all cases when testing five medium datasets.

Figure 9 compares the convergence curve of the proposed BAOA to other current methods in medium datasets. As can be observed from the figure, the BAOA performs better than most of the existing algorithms in handling the majority of medium datasets. On several datasets, including the Zoo, Vote, and CongressEW datasets, the BCAT, BDF, BPSO, and BGA algorithms' behaviors in these three datasets show premature convergence. Therefore, the results show that BAOA has the capacity to strike a good balance between the exploration and exploitation stages compared to other existing methods when dealing with medium datasets.

Figure 10 illustrates the results of the comparison techniques utilizing seven small datasets in terms of average accuracy. As demonstrated in this image, the suggested approach (BAOA) produced better results in all test situations when testing tiny datasets.

Comparing seven small datasets in terms of their number of selected characteristics is depicted in Figure 11. This figure shows that the suggested method (BAOA) obtained the least number of chosen features in all test instances when testing seven small datasets.

Figure 12 compares the convergence curve of the proposed BAOA to other current methods in small datasets. The proposed BAOA outperforms the existing methods in handling most small datasets. On five small datasets out of seven, including the Breastcancer, Exactly, Exactly2, Tic Tac Toe, and WineEW datasets, the behaviors of the BCAT, BDF, BPSO, and BGA algorithms in these five datasets are prone to

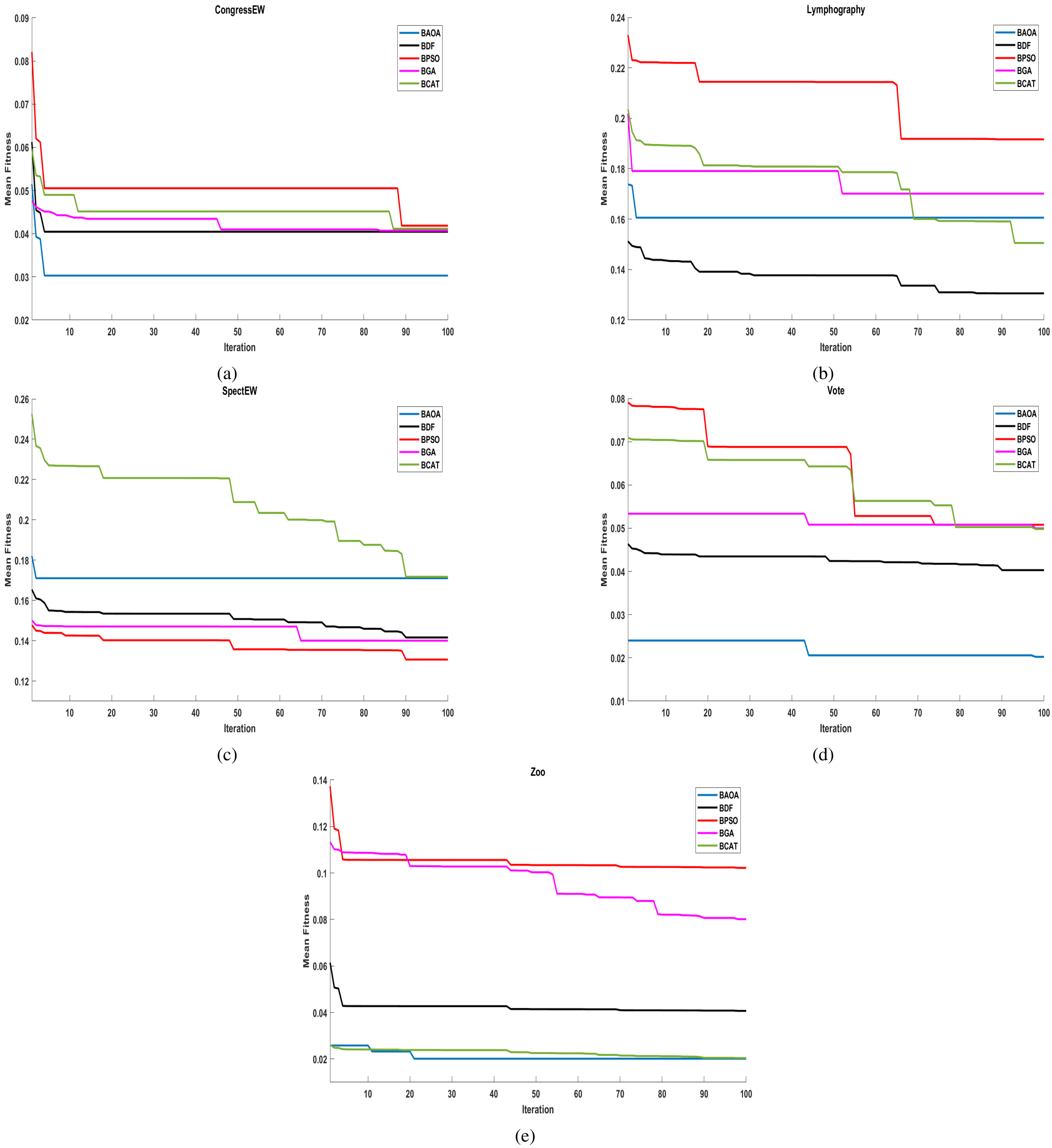


FIGURE 9. **Convergence curves for BAOA and other state-of-art methods for medium datasets.**

premature convergence. Therefore, the results clearly indicate the superiority of the BAOA in maintaining a good balance between the exploration and exploitation phases compared to other existing methods when dealing with small datasets.

Figure 13 shows the results' conclusion for all the tested methods using 18 datasets in different evaluation measures. An additional observation is made to demonstrate the method's effectiveness in resolving feature selection issues.

A. STATISTICAL ANALYSIS ON SAMPLE DATASETS

This section performs an in-depth analysis to check the effectiveness and superiority of the proposed method BAOA.

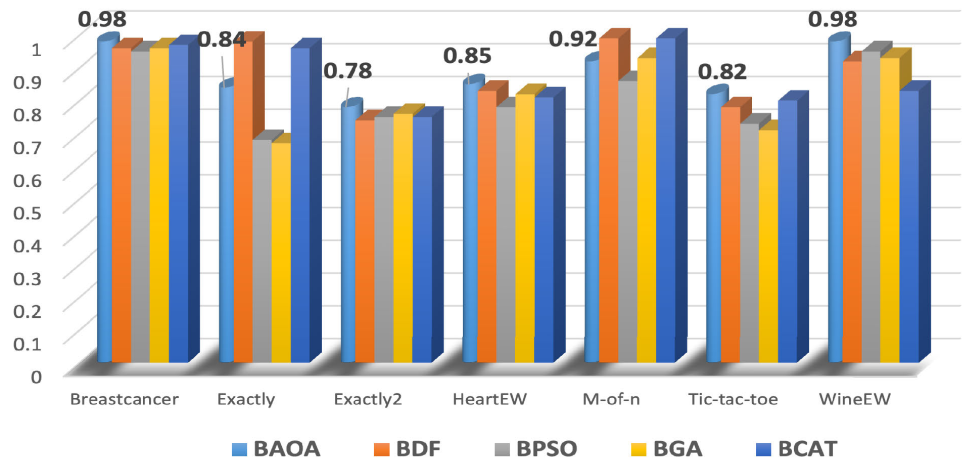


FIGURE 10. Comparison of average accuracy in small data.

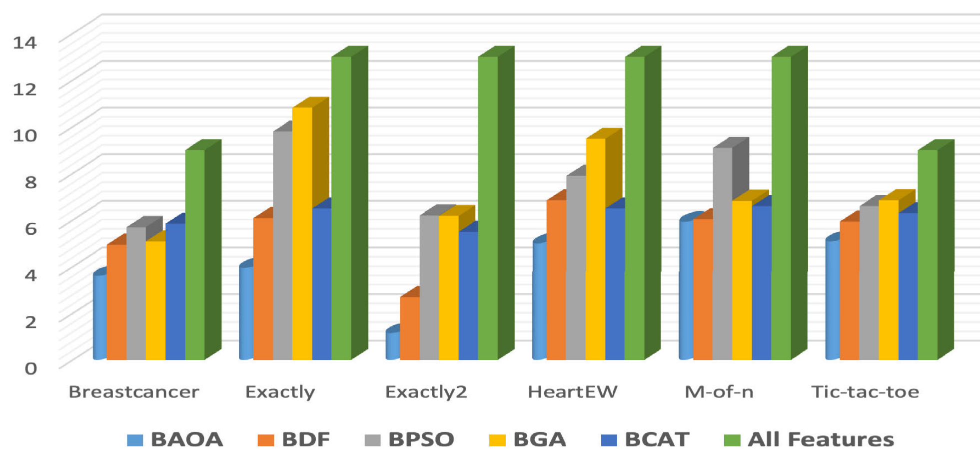


FIGURE 11. Comparison of selected features in small data.

TABLE 9. Statistical analysis based on classification accuracy using the PengLung dataset.

	BAOA	BDF	BPSO	BGA	BCAT
Number of values	10	10	10	10	10
Minimum	0.96	0.87	0.7	0.7	0.91
25% Percentile	0.96	0.89	0.72	0.7	0.91
Median	0.96	0.89	0.72	0.7	0.91
75% Percentile	0.96	0.89	0.72	0.7125	0.9125
Maximum	0.961	0.9	0.73	0.73	0.93
Range	0.001	0.03	0.03	0.03	0.02
10% Percentile	0.96	0.872	0.702	0.7	0.91
90% Percentile	0.9609	0.899	0.729	0.729	0.929
95% CI of median					
Actual confidence level	97.85%	97.85%	97.85%	97.85%	97.85%
Lower confidence limit	0.96	0.89	0.72	0.7	0.91
Upper confidence limit	0.96	0.89	0.72	0.72	0.92
Mean	0.9601	0.889	0.719	0.706	0.913
Std. Deviation	0.0003162	0.007379	0.007379	0.01075	0.006749
Std. Error of Mean	0.0001	0.002333	0.002333	0.003399	0.002134

This analysis is based on the analysis of variance (ANOVA) test and Wilcoxon signed rank test. As seen from Table 9, in the PengLung dataset, in terms of classification accuracy, the proposed BAOA outperforms the state-of-the-art methods on the following statistical measures: minimum, median, maximum, range, mean, and std. the error of mean with the values of 0.96, 0.96, 0.961, 0.001, 0.9601, and 0.0001,

TABLE 10. Analysis of variance test based on classification accuracy using the PengLung dataset.

ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
Treatment (between columns)	0.5472	4	0.1368	F (4, 45) = 2532	P<0.0001
Residual (within columns)	0.002431	45	0.00005402		
Total	0.5496	49			

respectively. The ANOVA test was utilized to evaluate the statistical significance of the variations in the classification accuracy obtained from the proposed BAOA algorithm and other optimizers. This analysis aims to determine the independence of the results obtained from multiple algorithms. The null hypothesis assumes no significant difference exists between the classification accuracy obtained from the proposed BAOA optimizer and other optimizers. A significance level greater than 5 % confirms the null hypothesis, while a significance level less than 5% rejects it. In terms of significant difference using classification

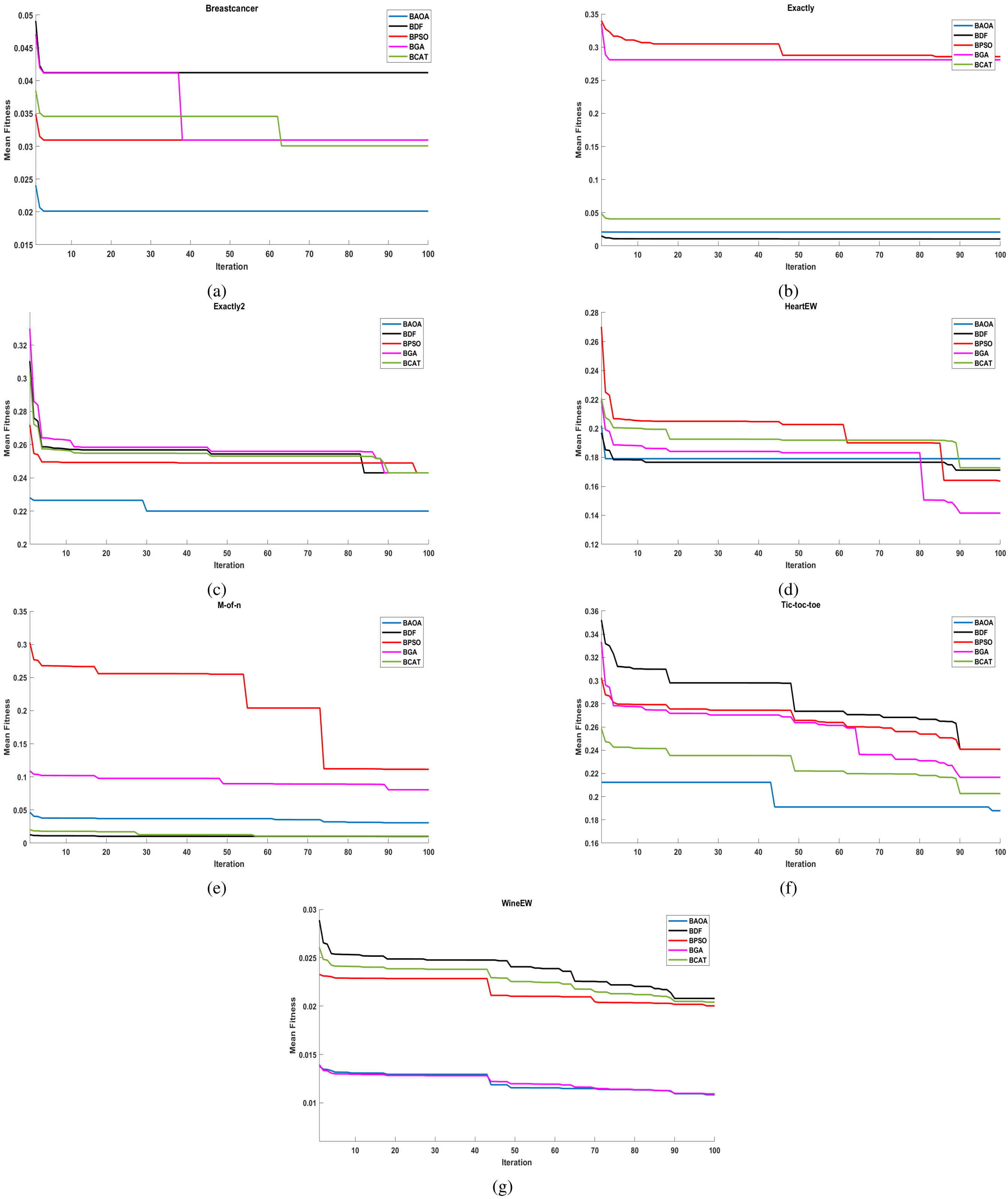


FIGURE 12. Convergence curves for BAOA and other state-of-art methods for small datasets.

accuracy as a measure, Table 10 shows that based on ANOVA the proposed BAOA significantly differ from other algorithms with a p-value (0.0001). Similarly, based on Wilcoxon signed rank test in Table 11, there is a significant

FIGURE 13. Total best, mean, and worst function values obtained from BAOA and other methods overall 18 datasets.

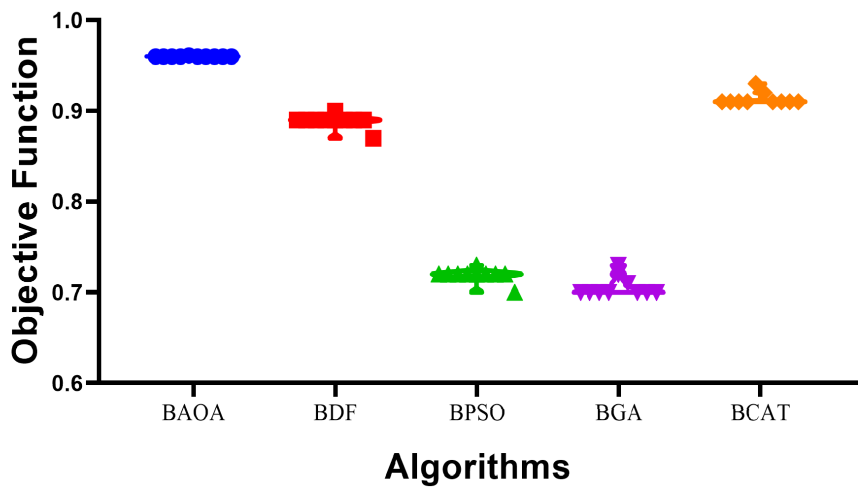


FIGURE 14. Average error of the results of the PengLung dataset.

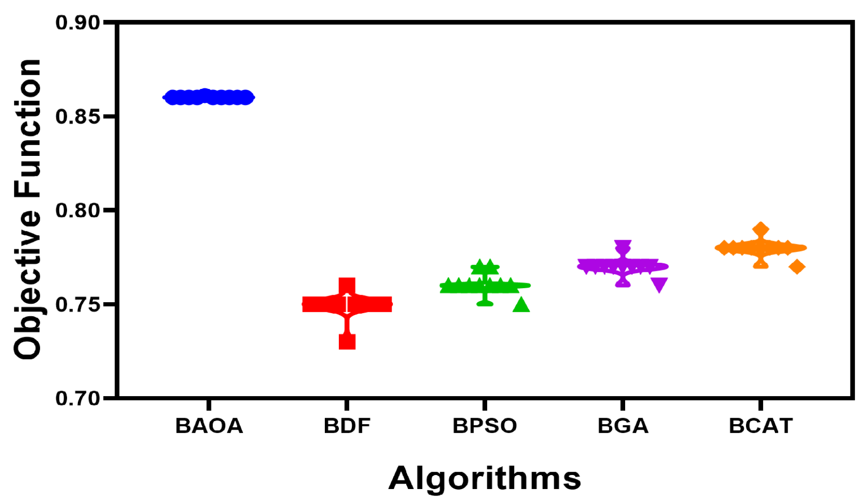


FIGURE 15. Average error of the results of the waveform dataset.

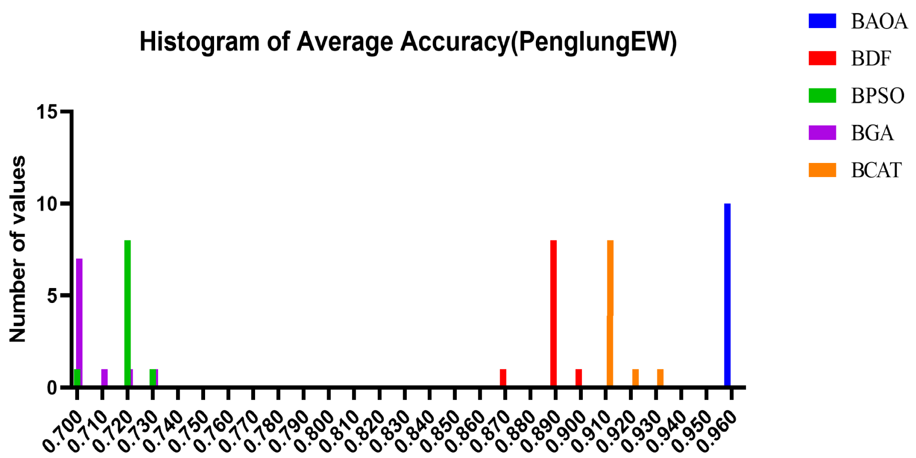


FIGURE 16. Histogram of the error in the results of the PengLung dataset.

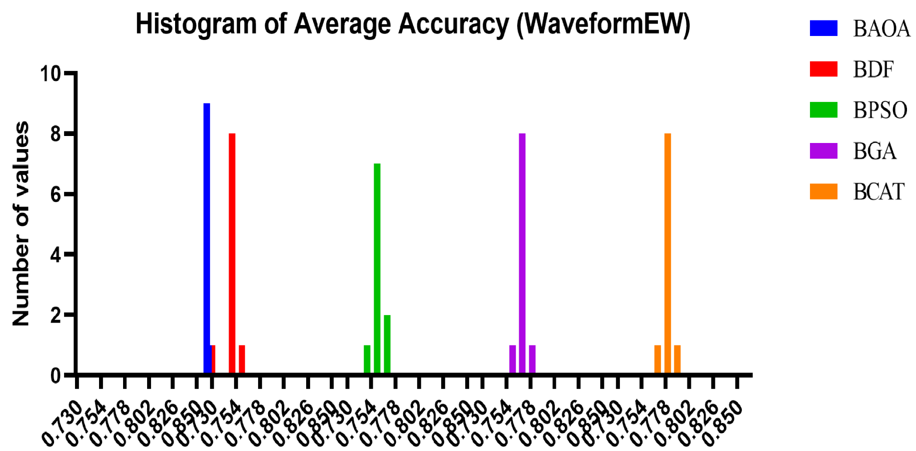


FIGURE 17. Histogram of the error in the results of the waveform dataset.

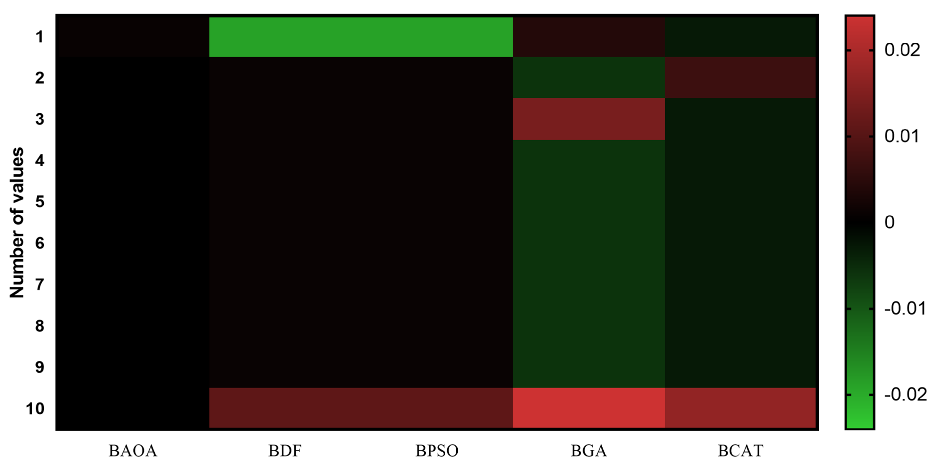


FIGURE 18. Heatmap of the analysis applied to the PengLung dataset.

difference and the proposed BAOA is better in terms of actual median and discrepancy. In addition, the visual representation of the results of the ANOVA is also presented in terms of a set of plots. Figure 14 shows the average error of the

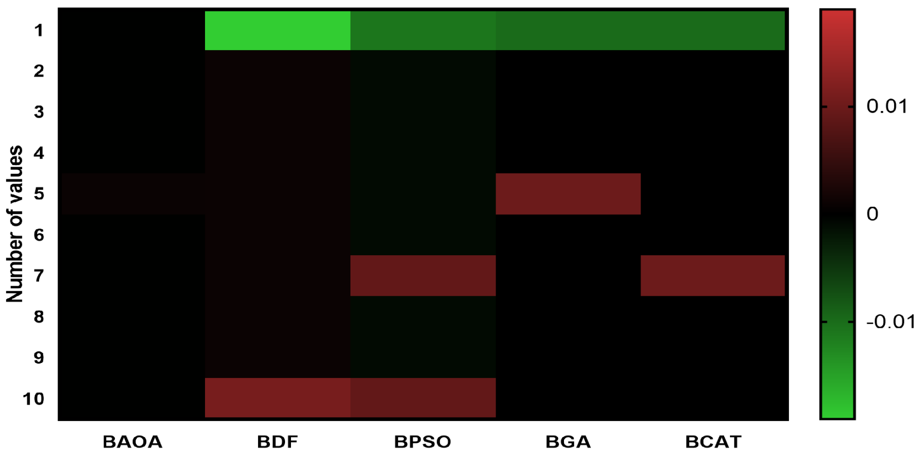


FIGURE 19. Heatmap of the analysis applied to the waveform dataset.

TABLE 11. Wilcoxon signed rank test based on classification accuracy using the PengLung dataset.

	BAOA	BDF	BPSO	BGA	BCAT
Theoretical median	0	0	0	0	0
Actual median	0.96	0.89	0.72	0.7	0.91
Number of values	10	10	10	10	10
Wilcoxon Signed Rank Test					
Sum of signed ranks (W)	55	55	55	55	55
Sum of positive ranks	55	55	55	55	55
Sum of negative ranks	0	0	0	0	0
P value (two-tailed)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Exact or an estimate?	Exact	Exact	Exact	Exact	Exact
P value summary	**	**	**	**	**
Significant (alpha=0.05)?	Yes	Yes	Yes	Yes	Yes
How big is the discrepancy?					
Discrepancy	0.96	0.89	0.72	0.7	0.91

TABLE 12. Statistical analysis based on classification accuracy using the waveform dataset.

	BAOA	BDF	BPSO	BGA	BCAT
Number of values	10	10	10	10	10
Minimum	0.86	0.73	0.75	0.76	0.77
25% Percentile	0.86	0.75	0.76	0.77	0.78
Median	0.86	0.75	0.76	0.77	0.78
75% Percentile	0.86	0.75	0.7625	0.77	0.78
Maximum	0.861	0.76	0.77	0.78	0.79
Range	0.001	0.03	0.02	0.02	0.02
10% Percentile	0.86	0.732	0.751	0.761	0.771
90% Percentile	0.8609	0.759	0.77	0.779	0.789
95% CI of median					
Actual confidence level	97.85%	97.85%	97.85%	97.85%	97.85%
Lower confidence limit	0.86	0.75	0.76	0.77	0.78
Upper confidence limit	0.86	0.75	0.77	0.77	0.78
Mean	0.8601	0.749	0.761	0.77	0.78
Std. Deviation	0.0003162	0.007379	0.005676	0.004714	0.004714
Std. Error of Mean	0.0001	0.002333	0.001795	0.001491	0.001491

TABLE 13. Analysis of variance test based on classification accuracy using the waveform dataset.

ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
Treatment	0.07757	4	0.01939	F (4, 45) = 739.0	P<0.0001
Residual	0.001181	45	0.00002624		
Total	0.07875	49			

results of the proposed BAOA (blue color) compared to those existing in the literature, which indicates better performance. Similarly, in terms of average accuracy, Figure 15 shows better performance (0.96) of the proposed BAOA compared to the existing algorithms. The heatmap in Figure 18, illustrates clearly that the proposed BAOA surpassed the existing algorithms on the PengLung dataset.

Similarly, in the Waveform dataset, as can be seen from Table 12, the proposed BAOA shows better performance than the other methods in terms of classification accuracy for the following statistical measures: minimum, median, maximum, range, mean, and std. the error of mean with the values of 0.86, 0.86, 0.861, 0.001, 0.8601, and 0.0001, respectively. In terms of significant difference using classification accuracy measure, Table 13 shows that based on ANOVA the proposed BAOA significantly differ from the-state-of-the arts in terms of classification accuracy with a p-value (0.0001). Also, based on Wilcoxon signed rank test in Table 14, there is a significant difference and the proposed BAOA is better in terms of actual median and discrepancy.

The graphic illustration of the results of the ANOVA is also presented in terms of a set of different plots for the Waveform dataset. Figure 17 shows the average error of the results of the proposed BAOA (blue color) compared to those existing in the literature, which indicates better performance. Similarly, in terms of average accuracy, Figure 18 shows better performance (0.86) of the proposed BAOA compared to the existing algorithms. The heatmap in Figure 19, illustrates clearly that the proposed BAOA surpassed the existing algorithms for the Waveform dataset. Finally, we statistically find the p-value based on classification accuracy for all datasets as shown in Table 15. These statistical analyses emphasize the effectiveness of the proposed BAOA, which clearly indicates that BAOA can handle feature selection

TABLE 14. Wilcoxon signed rank test based on classification accuracy using the waveform dataset.

	BAOA	BDF	BPSO	BGA	BCAT
Theoretical median	0	0	0	0	0
Actual median	0.86	0.75	0.76	0.77	0.78
Number of values	10	10	10	10	10
Wilcoxon Signed Rank Test					
Sum of signed ranks (W)	55	55	55	55	55
Sum of positive ranks	55	55	55	55	55
Sum of negative ranks	0	0	0	0	0
P value (two-tailed)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Exact or an estimate?	Exact	Exact	Exact	Exact	Exact
P value summary	**	**	**	**	**
Significant (alpha=0.05)?	Yes	Yes	Yes	Yes	Yes
How big is the discrepancy?					
Discrepancy	0.86	0.75	0.76	0.77	0.78

TABLE 15. P-values of the statistical analysis based on classification accuracy for all datasets.

Dataset	BAOA	BDF	BPSO	BGA	BCAT
1	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
2	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
3	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
4	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
5	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
6	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
7	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
8	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
9	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
10	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
11	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
12	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
13	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
14	<0.0001	<0.0001	<0.0001	<0.0001	5.44E-02
15	<0.0001	<0.0001	6.12E-02	<0.0001	<0.0001
16	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
17	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
18	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

problems very effectively and achieve the main aim of feature selection by reducing the number of features and maximizing classification accuracy.

VI. CONCLUSION

In this study, the problem of feature selection was addressed by a binary variant of the Arithmetic Optimization Algorithm (BAOA). In order to verify the usefulness and efficiency of the proposed method, 18 common benchmark datasets from UCI were used. The proposed method was evaluated using a set of evaluation measures. Binary Dragonfly Algorithm (BDF), Binary Particle Swarm Optimization (BPSO), Binary Genetic Algorithm (BGA), and Binary Cat Swarm Optimization (BCAT), four feature selection algorithms, were compared to the suggested method. The proposed method outperformed a wide range of methods in terms of accuracy and the number of features selected on the majority of datasets. This method outperformed existing state-of-the-art methods by a considerable margin when compared statistically using the metrics Mean, Best, and Worst Fitness, amongst others. The suggested BAOA method's superior results show it can manage the trade-off between exploratory and exploitative tendencies during optimization iterations.

In future studies, it would be worthwhile to apply real-world problems, such as optimization in engineering issues and different transfer functions, and see how BAOA behaves accordingly. Additionally, the proposed technique can be compared to other popular classifiers, SVM and Artificial Neural Networks, which are significant competitors of KNN, to determine whether performance remains constant or differs.

ACKNOWLEDGMENT

Authors wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

REFERENCES

[1] S. Tabakhi, P. Moradi, and F. Akhlaghian, "An unsupervised feature selection algorithm based on ant colony optimization," *Eng. Appl. Artif. Intell.*, vol. 32, pp. 112–123, Jun. 2014.

[2] S.-H. Liao, "Knowledge management technologies and applications—Literature review from 1995 to 2002," *Exp. Syst. Appl.*, vol. 25, no. 2, pp. 155–164, Aug. 2003.

[3] M.-S. Chen, J. Han, and P. S. Yu, "Data mining: An overview from a database perspective," *IEEE Trans. Knowl. Data Eng.*, vol. 8, no. 6, pp. 866–883, Dec. 1996.

[4] N. Khodadadi, E. Harati, F. De Caso, and A. Nanni, "Optimizing truss structures using composite materials under natural frequency constraints with a new hybrid algorithm based on cuckoo search and stochastic paint optimizer (CSSPO)," *Buildings*, vol. 13, no. 6, p. 1551, Jun. 2023.

[5] E.-S.-M. El-kenawy, F. Albalawi, S. A. Ward, S. S. M. Ghoneim, M. M. Eid, A. A. Abdelhamid, N. Bailek, and A. Ibrahim, "Feature selection and classification of transformer faults based on novel meta-heuristic algorithm," *Mathematics*, vol. 10, no. 17, p. 3144, Sep. 2022.

[6] N. Khodadadi, A. Ö. Çiftçioğlu, S. Mirjalili, and A. Nanni, "A comparison performance analysis of eight meta-heuristic algorithms for optimal design of truss structures with static constraints," *Decis. Analytics J.*, vol. 8, Sep. 2023, Art. no. 100266.

[7] A. Alqushaibi, S. J. Abdulkadir, H. M. Rais, Q. Al-Tashi, M. G. Ragab, and H. Alhussian, "Enhanced weight-optimized recurrent neural networks based on sine cosine algorithm for wave height prediction," *J. Mar. Sci. Eng.*, vol. 9, no. 5, p. 524, May 2021.

[8] N. Khodadadi, L. Abualigah, and S. Mirjalili, "Multi-objective stochastic paint optimizer (MOSPO)," *Neural Comput. Appl.*, vol. 34, pp. 18035–18058, Jun. 2022.

[9] L. Abualigah, M. A. Elaziz, N. Khodadadi, A. Forestiero, H. Jia, and A. H. Gandomi, "Aquila optimizer based PSO swarm intelligence for IoT task scheduling application in cloud computing," in *Integrating Meta-Heuristics and Machine Learning for Real-World Optimization Problems*. Cham, Switzerland: Springer, 2022, pp. 481–497.

[10] S. Sharma, N. Khodadadi, A. K. Saha, F. S. Gharehchopogh, and S. Mirjalili, "Non-dominated sorting advanced butterfly optimization algorithm for multi-objective problems," *J. Bionic Eng.*, vol. 20, no. 2, pp. 819–843, Mar. 2023.

[11] A. A. Abdelhamid, E.-S.-M. El-Kenawy, N. Khodadadi, S. Mirjalili, D. S. Khafaga, A. H. Alharbi, A. Ibrahim, M. M. Eid, and M. Saber, "Classification of monkeypox images based on transfer learning and the Al-biruni Earth radius optimization algorithm," *Mathematics*, vol. 10, no. 19, p. 3614, Oct. 2022.

[12] A. Kaveh, S. Talatahari, and N. Khodadadi, "Stochastic paint optimizer: Theory and application in civil engineering," *Eng. Comput.*, vol. 38, pp. 1921–1952, Oct. 2020.
[13] T. R. Farshi, "Battle royale optimization algorithm," *Neural Comput. Appl.*, vol. 33, no. 4, pp. 1139–1157, Feb. 2021.
[14] B. Abdollahzadeh, F. S. Gharehchopogh, N. Khodadadi, and S. Mirjalili, "Mountain gazelle optimizer: A new nature-inspired metaheuristic algorithm for global optimization problems," *Adv. Eng. Softw.*, vol. 174, Dec. 2022, Art. no. 103282.
[15] M. M. Fouad, A. I. El-Desouky, R. Al-Hajj, and E. M. El-Kenawy, "Dynamic group-based cooperative optimization algorithm," *IEEE Access*, vol. 8, pp. 148378–148403, 2020.
[16] N. Khodadadi, S. Talatahari, and A. H. Gandomi, "ANNA: Advanced neural network algorithm for optimisation of structures," *Proc. Inst. Civil Engineers Struct. Buildings*, Apr. 2023, pp. 1–23.
[17] H. Karami, M. V. Anaraki, S. Farzin, and S. Mirjalili, "Flow direction algorithm (FDA): A novel optimization approach for solving optimization problems," *Comput. Ind. Eng.*, vol. 156, Jun. 2021, Art. no. 107224.
[18] A. A. Alhussan, D. S. Khafaga, E. M. El-Kenawy, A. Ibrahim, M. M. Eid, and A. A. Abdelhamid, "Pothole and plain road classification using adaptive mutation dipper throated optimization and transfer learning for self driving cars," *IEEE Access*, vol. 10, pp. 84188–84211, 2022.
[19] A. A. Abdelhamid, S. K. Towfek, N. Khodadadi, A. A. Alhussan, D. S. Khafaga, M. M. Eid, and A. Ibrahim, "Waterwheel plant algorithm: A novel metaheuristic optimization method," *Processes*, vol. 11, no. 5, p. 1502, May 2023.
[20] H.-N. Dai, H. Wang, G. Xu, J. Wan, and M. Imran, "Big data analytics for manufacturing Internet of Things: Opportunities, challenges and enabling technologies," *Enterprise Inf. Syst.*, vol. 14, nos. 9–10, pp. 1279–1303, Nov. 2020.
[21] Y.-P. Chen, Y. Li, G. Wang, Y.-F. Zheng, Q. Xu, J.-H. Fan, and X.-T. Cui, "A novel bacterial foraging optimization algorithm for feature selection," *Exp. Syst. Appl.*, vol. 83, pp. 1–17, Oct. 2017.
[22] S. Arora and P. Anand, "Binary butterfly optimization approaches for feature selection," *Exp. Syst. Appl.*, vol. 116, pp. 147–160, Feb. 2019.
[23] Q. Al-Tashi, S. J. Abdul 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.
[24] Z. Sadeghian, E. Akbari, and H. Nematzadeh, "A hybrid feature selection method based on information theory and binary butterfly optimization algorithm," *Eng. Appl. Artif. Intell.*, vol. 97, Jan. 2021, Art. no. 104079.
[25] Q. Al-Tashi, S. J. Abdulkadir, H. M. Rais, S. Mirjalili, H. Alhussian, M. G. Ragab, and A. Alqushaibi, "Binary multi-objective grey wolf optimizer for feature selection in classification," *IEEE Access*, vol. 8, pp. 106247–106263, 2020.
[26] K. Chen, F.-Y. Zhou, and X.-F. Yuan, "Hybrid particle swarm optimization with spiral-shaped mechanism for feature selection," *Exp. Syst. Appl.*, vol. 128, pp. 140–156, Aug. 2019.
[27] Q. Pang and L. Zhang, "A recursive feature retention method for semi-supervised feature selection," *Int. J. Mach. Learn. Cybern.*, vol. 12, pp. 2639–2657, May 2021.
[28] M. M. Kabir, M. Shahjahan, and K. Murase, "A new hybrid ant colony optimization algorithm for feature selection," *Exp. Syst. Appl.*, vol. 39, no. 3, pp. 3747–3763, Feb. 2012.
[29] Q. Al-Tashi, S. J. Abdulkadir, H. M. Rais, S. Mirjalili, and H. Alhussian, "Approaches to multi-objective feature selection: A systematic literature review," *IEEE Access*, vol. 8, pp. 125076–125096, 2020.
[30] Z. Chang, L. Lei, Z. Zhou, S. Mao, and T. Ristaniemi, "Learn to cache: Machine learning for network edge caching in the big data era," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 28–35, Jun. 2018.
[31] Q. Chen, M. Zhang, and B. Xue, "Feature selection to improve generalization of genetic programming for high-dimensional symbolic regression," *IEEE Trans. Evol. Comput.*, vol. 21, no. 5, pp. 792–806, Oct. 2017.
[32] M. Hamraz, A. Ali, W. K. Mashwani, S. Aldahmani, and Z. Khan, "Feature selection for high dimensional microarray gene expression data via weighted signal to noise ratio," *PLoS ONE*, vol. 18, no. 4, Apr. 2023, Art. no. e0284619.
[33] M. Dash and H. Liu, "Consistency-based search in feature selection," *Artif. Intell.*, vol. 151, nos. 1–2, pp. 155–176, Dec. 2003.
[34] H. Liu, M. Dash, and H. Motoda, "Consistency based feature selection," in *Knowledge Discovery and Data Mining. Current Issues and New Applications: Current Issues and New Applications*. Kyoto, Japan: Springer, 2007, p. 98.
[35] M. Hamraz, Z. Khan, D. M. Khan, N. Gul, A. Ali, and S. Aldahmani, "Gene selection in binary classification problems within functional genomics experiments via robust Fisher score," *IEEE Access*, vol. 10, pp. 51682–51692, 2022.
[36] A. Jovic, K. Brkic, and N. Bogunovic, "A review of feature selection methods with applications," in *Proc. 38th Int. Conv. Inf. Commun. Technol., Electron. Microelectron. (MIPRO)*, May 2015, pp. 1200–1205.
[37] Q. Al-Tashi, H. M. Rais, S. J. Abdulkadir, and S. Mirjalili, "Feature selection based on grey wolf optimizer for oil & gas reservoir classification," in *Proc. Int. Conf. Comput. Intell. (ICCI)*, Oct. 2020, pp. 211–216.
[38] B. M. E. El-Den and M. M. Eid, "Watermarking models and artificial intelligence," *J. Artif. Intell. Metaheuristics*, vol. 1, no. 2, pp. 24–30, 2022.
[39] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Comput. Elect. Eng.*, vol. 40, no. 1, pp. 16–28, Jan. 2014.
[40] A. Zakeri and A. Hokmabadi, "Efficient feature selection method using real-valued grasshopper optimization algorithm," *Exp. Syst. Appl.*, vol. 119, pp. 61–72, Apr. 2019.
[41] M. A. Elaziz, A. Dahou, L. Abualigah, L. Yu, M. Alshinwan, A. M. Khasawneh, and S. Lu, "Advanced metaheuristic optimization techniques in applications of deep neural networks: A review," *Neural Comput. Appl.*, vol. 33, pp. 14079–14099, Apr. 2021.
[42] Q. Al-Tashi, M. B. Saad, A. Muneer, R. Qureshi, S. Mirjalili, A. Sheshadri, X. Le, N. I. Vokes, J. Zhang, and J. Wu, "Machine learning models for the identification of prognostic and predictive cancer biomarkers: A systematic review," *Int. J. Mol. Sci.*, vol. 24, no. 9, p. 7781, Apr. 2023.
[43] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, Apr. 1997.
[44] L. Abualigah, A. Diabat, S. Mirjalili, M. A. Elaziz, and A. H. Gandomi, "The arithmetic optimization algorithm," *Comput. Methods Appl. Mech. Eng.*, vol. 376, Apr. 2021, Art. no. 113609.
[45] Y. Jiang, Q. Luo, Y. Wei, L. Abualigah, and Y. Zhou, "An efficient binary gradient-based optimizer for feature selection," *Math. Biosciences Eng.*, vol. 18, no. 4, pp. 3813–3854, 2021.
[46] Q. Al-Tashi, H. M. Rais, S. J. Abdulkadir, S. Mirjalili, and H. Alhussian, "A review of grey wolf optimizer-based feature selection methods for classification," in *Evolutionary Machine Learning Techniques* (Algorithms for Intelligent Systems), S. Mirjalili, H. Faris, and I. Aljarah, Eds. Singapore: Springer, 2020, doi: 10.1007/978-981-32-9990-0_13.
[47] 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.
[48] Q. Al-Tashi, T M Shami, S. J. Abdulkadir, E. A. P. Akhir, A. Alwadain, H. Alhussain, A. Alqushaibi, H. M. Rais, A. Muneer, M. B. Saad, J. Wu, and S. Mirjalili, "Enhanced multi-objective grey wolf optimizer with lévy flight and mutation operators for feature selection," *Comput. Syst. Sci. Eng.*, vol. 47, no. 2, pp. 1937–1966, 2023.
[49] A. A. Ewees, M. A. El Aziz, and A. E. Hassanien, "Chaotic multi-verse optimizer-based feature selection," *Neural Comput. Appl.*, vol. 31, no. 4, pp. 991–1006, Apr. 2019.
[50] S. Gu, R. Cheng, and Y. Jin, "Feature selection for high-dimensional classification using a competitive swarm optimizer," *Soft Comput.*, vol. 22, no. 3, pp. 811–822, Feb. 2018.
[51] C. B. Şahin and L. Abualigah, "A novel deep learning-based feature selection model for improving the static analysis of vulnerability detection," *Neural Comput. Appl.*, vol. 33, pp. 14049–14067, May 2021.
[52] Q. Al-Tashi, M. B. Saad, A. Sheshadri, C. C. Wu, J. Y. Chang, B. Al-Lazikani, C. Gibbons, N. I. Vokes, J. Zhang, J. J. Lee, J. V. Heymach, D. Jaffray, S. Mirjalili, and J. Wu, "SwarmDeepSurv: Swarm intelligence advances deep survival network for prognostic radiomics signatures in four solid cancers," *Patterns*, vol. 4, no. 8, Aug. 2023, Art. no. 100777.
[53] E.-S.-M. El-kenawy, A. Ibrahim, N. Bailek, K. Bouchouicha, M. A. Hassan, M. Jamei, and N. Al-Ansari, "Sunshine duration measurements and predictions in Saharan Algeria region: An improved ensemble learning approach," *Theor. Appl. Climatol.*, vol. 147, nos. 3–4, pp. 1015–1031, Nov. 2021.
[54] E.-S.-M. El-Kenawy, S. Mirjalili, A. A. Abdelhamid, A. Ibrahim, N. Khodadadi, and M. M. Eid, "Meta-heuristic optimization and keystroke dynamics for authentication of smartphone users," *Mathematics*, vol. 10, no. 16, p. 2912, Aug. 2022.

[55] M. Marwa and R. M. Zaki, "Classification of student performance based on ensemble optimized using dipper throated optimization," *J. Artif. Intell. Metaheuristics*, vol. 2, no. 1, pp. 36–45, 2022.

[56] K.-C. Lin, K.-Y. Zhang, Y.-H. Huang, J. C. Hung, and N. Yen, "Feature selection based on an improved cat swarm optimization algorithm for big data classification," *J. Supercomput.*, vol. 72, no. 8, pp. 3210–3221, Aug. 2016.

[57] A. A. Ewees, L. Abualigah, D. Yousri, Z. Y. Algamal, M. A. Al-Qaness, R. A. Ibrahim, and M. A. Elaziz, "Improved slime mould algorithm based on firefly algorithm for feature selection: A case study on QSAR model," *Eng. Comput.*, vol. 38, pp. 2407–2421, Mar. 2021.

[58] M. M. Mafarja, D. Eleyan, I. Jaber, A. Hammouri, and S. Mirjalili, "Binary dragonfly algorithm for feature selection," in *Proc. Int. Conf. New Trends Comput. Sci. (ICTCS)*, Oct. 2017, pp. 12–17.

[59] L. Abualigah, B. Alsalibi, M. Shehab, M. Alshinwan, A. M. Khasawneh, and H. Alabool, "A parallel hybrid Krill Herd algorithm for feature selection," *Int. J. Mach. Learn. Cybern.*, vol. 12, no. 3, pp. 783–806, Mar. 2021.

[60] L. Abualigah and A. J. Dulaimi, "A novel feature selection method for data mining tasks using hybrid sine cosine algorithm and genetic algorithm," *Cluster Comput.*, vol. 24, pp. 2161–2176, Feb. 2021.

[61] Q. Al-Tashi, S. Mirjalili, J. Wu, S. J. Abdulkadir, T. M. Shami, N. Khodadadi, and A. Alqushaibi, "Moth-flame optimization algorithm for feature selection: A review and future trends," in *Handbook of Moth-Flame Optimization Algorithm*. Taylor & Francis, 2022, pp. 11–34.

[62] R. Alwajih, S. J. Abdulkadir, H. A. Hussian, N. Aziz, Q. Al-Tashi, S. Mirjalili, and A. Alqushaibi, "Hybrid binary whale with Harris Hawks for feature selection," *Neural Comput. Appl.*, vol. 34, pp. 19377–19395, Jul. 2022.

[63] O. A. Alomari, A. T. Khader, M. A. Al-Betar, and L. M. Abualigah, "Gene selection for cancer classification by combining minimum redundancy maximum relevancy and bat-inspired algorithm," *Int. J. Data Mining Bioinf.*, vol. 19, no. 1, pp. 32–51, 2017.

[64] L. Sun, S. Si, J. Zhao, J. Xu, Y. Lin, and Z. Lv, "Feature selection using binary monarch butterfly optimization," *Appl. Intell.*, vol. 53, pp. 706–727, Apr. 2022.

[65] B. Turkoglu, S. A. Uymaz, and E. Kaya, "Binary artificial algae algorithm for feature selection," *Appl. Soft Comput.*, vol. 120, May 2022, Art. no. 108630.

[66] J. Too and S. Mirjalili, "A hyper learning binary dragonfly algorithm for feature selection: A COVID-19 case study," *Knowl.-Based Syst.*, vol. 212, Jan. 2021, Art. no. 106553.

[67] A. Chaudhuri and T. P. Sahu, "Feature selection using binary crow search algorithm with time varying flight length," *Exp. Syst. Appl.*, vol. 168, Apr. 2021, Art. no. 114288.

[68] R. A. Khurma, I. Aljarah, A. Sharieh, M. A. Elaziz, R. Damaševičius, and T. Krilavičius, "A review of the modification strategies of the nature inspired algorithms for feature selection problem," *Mathematics*, vol. 10, no. 3, p. 464, Jan. 2022.

[69] N. Neggaz, E. H. Houssein, and K. Hussain, "An efficient Henry gas solubility optimization for feature selection," *Exp. Syst. Appl.*, vol. 152, Aug. 2020, Art. no. 113364.

[70] B. Nouri-Moghaddam, M. Ghazanfari, and M. Fathian, "A novel multi-objective forest optimization algorithm for wrapper feature selection," *Exp. Syst. Appl.*, vol. 175, Aug. 2021, Art. no. 114737.

[71] N. M. Hijazi, H. Faris, and I. Aljarah, "A parallel metaheuristic approach for ensemble feature selection based on multi-core architectures," *Exp. Syst. Appl.*, vol. 182, Nov. 2021, Art. no. 115290.

[72] Y. Lu, M. Liang, Z. Ye, and L. Cao, "Improved particle swarm optimization algorithm and its application in text feature selection," *Appl. Soft Comput.*, vol. 35, pp. 629–636, Oct. 2015.

[73] L.-Y. Chuang, S.-W. Tsai, and C.-H. Yang, "Improved binary particle swarm optimization using catfish effect for feature selection," *Exp. Syst. Appl.*, vol. 38, no. 10, pp. 12699–12707, Sep. 2011.

[74] Y. Zheng, Y. Li, G. Wang, Y. Chen, Q. Xu, J. Fan, and X. Cui, "A novel hybrid algorithm for feature selection based on whale optimization algorithm," *IEEE Access*, vol. 7, pp. 14908–14923, 2019.

[75] M. Abdel-Basset, W. Ding, and D. El-Shahat, "A hybrid Harris Hawks optimization algorithm with simulated annealing for feature selection," *Artif. Intell. Rev.*, vol. 54, no. 1, pp. 593–637, Jan. 2021.

[76] C. L. Blake and C. J. Merz, "Uci repository of machine learning databases, 1998," Univ. California, Tech. Rep., 1998.

[77] E. R. Ziegel, "The elements of statistical learning," *Technometrics*, vol. 45, no. 3, pp. 267–268, 2003, doi: 10.1198/tech.2003.s770.

NIMA KHODADADI (Member, IEEE) received the bachelor's degree in civil engineering and the master's degree in structural civil engineering from the University of Tabriz (one of the ten top universities in Iran), and the Ph.D. degree from the Iran University of Science and Technology (IUST). He is currently pursuing the Ph.D. degree with the University of Miami (UM), Miami, FL, USA. He is also a Research Assistant. He has published more than 45 publications with more than 500 citations. His research interests include wind engineering and steel structures, particularly in the experimental and numerical investigation of steel-braced frames. In addition, he has been actively involved in engineering optimization, especially in evolutionary algorithms. Using comprehensive finite element analyses, he has also investigated the use of differently shaped sections in real steel frames. Recently, he has been engaged in research in single and multi-objective engineering optimization, especially in solving large-scale and practical structural design problems.

EHSAN KHODADADI received the B.Sc. and M.Sc. degrees in plant breeding with a specialization in molecular genetics and genetic engineering from the Islamic Azad University of Ardabil, in 2007 and 2009, respectively, and the Ph.D. degree in molecular genetics and genetic engineering from the University of Ilam, in 2017. He is currently pursuing the Ph.D. degree in material science and engineering. He joined the University of Arkansas, in 2023. He has made significant contributions to the scientific community throughout his career. He has published more than 30 papers covering diverse fields, such as bioinformatics, proteomics, post-translational modifications, systems biology, and optimization algorithms. These publications highlight his expertise and research interests across various domains. In addition to his research accomplishments, he was an Assistant Professor with the Islamic Azad University of Tabriz. This role allowed him to share his knowledge and passion for genetics and biotechnology by teaching science courses.

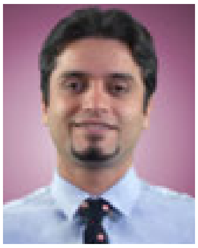
QASEM AL-TASHI received the B.Sc. degree in software engineering from Universiti Teknologi Malaysia (UTM), in 2012, the M.Sc. degree in software engineering from Universiti Kebangsaan Malaysia (UKM), in 2017, and the Ph.D. degree in information technology from Universiti Teknologi PETRONAS, in 2021. He was a Research Scientist with Universiti Teknologi PETRONAS and a Postdoctoral Fellow with The University of Texas MD Anderson Cancer Center, Houston, TX, USA, where he is currently a Research Investigator with the Department of Physics. His research interests include feature selection, swarm intelligence, biomarker discovery, survival analysis, multi-objective optimization, and artificial intelligence in healthcare. He is an Editor of the *Journal of Applied Artificial Intelligence* and the *Journal of Information Technology and Computing*. He is a Reviewer of several high-impact factor journals, such as *Artificial Intelligence Review*, IEEE ACCESS, *Knowledge-Based Systems*, *Soft Computing*, *Journal of Ambient Intelligence and Humanized Computing*, *Applied Soft Computing*, *Neurocomputing*, *Applied Artificial Intelligence*, and *PLOS One*.

EL-SAYED M. EL-KENAWY (Senior Member, IEEE) is currently an Assistant Professor with the Delta Higher Institute for Engineering and Technology (DHIET), Mansoura, Egypt. He has published more than 88 papers with more than 2400 citations and an H-index of 32. He has launched and pioneered independent research programs. He motivates and inspires his students in different ways by providing a thorough understanding of various computer concepts. He explains complex concepts in an easy-to-understand manner. His research interests include artificial intelligence, machine learning, optimization, deep learning, digital marketing, and data science. He is a Reviewer of *Computers, Materials and Continua* journal, IEEE ACCESS, and other journals.

LAITH ABUALIGAH (Senior Member, IEEE) received the degree in computer information system and the master's degree in computer science from Al al-Bayt University, Jordan, in 2011 and 2014, respectively, and the Ph.D. degree from the School of Computer Science, Universiti Sains Malaysia (USM), Malaysia, in 2018. He is currently an Associate Professor with Al-Ahliyya Amman University, Jordan. He is also a Distinguished Researcher with the School of Computer Science, USM. According to the report published by Stanford University, in 2020, he is one of the 2% influential scholars, which depicts the 100,000 top scientists in the world. He has published more than 220 journal articles and books, which collectively have been cited more than 7000 times (H-index = 39). His research interests include arithmetic optimization algorithm (AOA), bio-inspired computing, nature-inspired computing, swarm intelligence, artificial intelligence, meta-heuristic modeling, optimization algorithms, evolutionary computations, information retrieval, text clustering, feature selection, combinatorial problems, optimization, advanced machine learning, big data, and natural language processing. He serves as an Associate Editor for *Cluster Computing* (Springer), *Soft Computing* (Springer), and *Journal of King Saud University-Computer and Information Sciences* (Elsevier).

SAID JADID ABDULKADIR (Senior Member, IEEE) received the B.Sc. degree in computer science from Moi University, the M.Sc. degree in computer science from Universiti Teknologi Malaysia, and the Ph.D. degree in information technology from Universiti Teknologi PETRONAS. He is currently an Associate Professor with the Department of Computer and Information Sciences, Universiti Teknologi PETRONAS. His research interests include machine learning, deep learning architectures, and predictive and streaming analytics applications. He is serving in the executive committee for the IEEE Computational Intelligence Society Malaysia Chapter and the Editor-in-Chief for platform journal.

ALAWI ALQUSHAIBI received the B.Sc. degree in computer networks and security from Universiti Teknologi Malaysia, in 2012, and the master's degree in research from Universiti Teknologi PETRONAS, in 2021. During his academic journey, he has acquired knowledge and skills in conducting independent research, producing academic writing, and teaching computer science courses. He is currently an Academic Researcher, who has focused his research interests in the areas of machine learning, data science, optimization, feature selection, classification, data analytics, and image processing, specifically in generative adversarial networks (GANs).



SEYEDALI MIRJALILI (Senior Member, IEEE) is currently a Professor and the Founding Director of the Centre for Artificial Intelligence Research and Optimization, Torrens University Australia. He is internationally recognized for his advances in optimization and swarm intelligence, including the first set of algorithms from a synthetic intelligence standpoint-a radical departure from how natural systems are typically understood-and a systematic design framework to reliably benchmark, evaluate, and propose computationally cheap robust optimization algorithms. He has published more than 300 publications with more than 40,000 citations and an H-index of 70. As the most cited researcher in robust optimization, he has been on the list of 1% highly-cited researchers and has been named as one of the most influential researchers in the world by the Web of Science, for three consecutive years, since 2019. In 2020, he was ranked 21st across all disciplines and 4th in artificial intelligence and image processing in Stanford University's list of the World's Top Scientists. In 2021, The Australian newspaper named him the top researcher in Australia, in three fields artificial intelligence, evolutionary computation, and fuzzy systems. His research interests include optimization, swarm intelligence, evolutionary algorithms, and machine learning. He is an Associate Editor of several AI journals, including *Neurocomputing*, *Applied Soft Computing*, *Advances in Engineering Software*, *Computers in Biology and Medicine*, *Healthcare Analytics*, *Applied Intelligence*, and IEEE ACCESS.