This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3064799, IEEE Access

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Teaching Learning-based Optimization with Evolutionary Binarization Schemes for Tackling Feature Selection Problems

THAER THAHER[1], MAJDI MAFARJA[2], HAMZA TURABIEH [3], PEDRO A. CASTILLO[4], HOSSAM FARIS [5], IBRAHIM ALJARAH [6]

[1]Department of Engineering and Technology Sciences, *Arab American University, Ramallah, Palestine* (e-mail:thaer.thaher@gmail.com)
[2]Department of Computer Science, Birzeit University, Ramallah, Palestine (e-mail: mmafarja@birzeit.edu)
[3]Department of Information Technology, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia (e-mail: h.turabieh@tu.edu.sa)
[4]Department of Computer Architecture and Technology, University of Granada, 18071 Granada, Spain (e-mail:todos@geneura.ugr.es)
[5]Department of Business Information Technology, King Abdullah II School for Information Technology, The University of Jordan, Amman, Jordan

Corresponding author: Thaer Thaher (e-mail: t.thaher@student.aaup.edu).

ABSTRACT Machine learning techniques heavily rely on available training data in a data set. Certain features in the data can interfere with the learning process, so it is required to remove irrelevant and redundant features to build a robust training model. As such, several feature selection techniques are usually applied in a pre-processing phase to obtain the most appropriate set of features and improve the overall learning process. In this paper, a new feature selection approach is proposed based on a modified Teaching-Learning-based Optimization (TLBO) combined with four new binarization methods: the Elitist, the Elitist Roulette, the Elitist Tournament, and the Rank-based method. The influence of these binarization methods is studied and compared to other state-of-the-art techniques. The experimental results such as Shapiro-Wilk normality and Wilcoxon ranksum test show that both transfer functions and binarization approaches have a significant influence on the effectiveness of the binary TLBO. The experiments show that choosing a fitting transfer function along with a suitable binarization method has a substantial impact on the exploratory and exploitative potentials of the feature selection technique.

INDEX TERMS Teaching-Learning, Feature Selection, Metaheuristic, Transfer function, Binarization.

I. INTRODUCTION

The performance of Machine Learning (ML) techniques mainly depends on the nature of datasets, which often contain irrelevant or redundant features. such features could mislead or bias the learning process. Moreover, collecting data from different sources makes it possible to have redundant elements in the same dataset. To build a robust training model, therefore, the irrelevant and unnecessary features should be removed [1]. Feature Selection (FS), as a pre-processing step, has been widely used to search for the most informative features and increase the learning performance of a learning algorithm (e.g., classification). The importance of FS as a pre-processing step comes from the fact that there is a large number of features in a dataset; i.e., a large feature space, which requires a higher computational cost for the learning process.

FS methods can be broadly categorized into two classes: searching for the best feature combinations and evaluating those combinations. In the search stage, sequential forward, sequential backward, exhaustive, random, and heuristic selection are all examples of search strategies that can be used to search the feature space for finding the optimal or near optimal feature subsets [2]. Metaheuristic methods such as swarm intelligence algorithms (e.g., Particle Swarm Optimization (PSO) [3], Ant Colony Optimization (ACO) [4], Whale Optimization Algorithm [5], Harris hawks optimizer (HHO) [6], and Grey Wolf Optimizer (GWO) [7]), and Evolutionary Algorithms (e.g., Genetic Algorithm (GA) [8], Differential Evolution (DE) [9]) have been utilized by Chen et al. [10], Aljarah et al. [11], Xu et al. [12], Heidari et al. [13] as efficient search strategies in many optimization problems and especially for FS tasks.

From the evaluation perspective, FS methods are divided into three main categories; filters, wrappers, and embedded methods. Filter approaches (e.g., Chi-Square, Information Gain, Gain Ratio, and ReliefF) depend on finding the correlations between the features in evaluating the feature subset while no external evaluator participates in the evaluation process [14]. On the other hand, wrapper methods mainly depend on an external learning algorithm (e.g., classification algorithm, also known as induction algorithm) to evaluate the feature subsets [15]. However, the feature selection method is embedded in the learning process when considering the integrated approaches [16].

Wrapper approaches attracted the attention of many researchers in the literature, which is due to the involvement of the learning algorithm in the selection process, hence the selection of a feature is based on the resulting performance of the learning algorithm (e.g., classification accuracy for a specific classifier) [17]. Different classification algorithms (e.g., K-nearest Neighbor (KNN), Decision Tree (DT), and Artificial Neural Networks (ANN)) have been used in conjunction with different FS methods. Due to its simplicity, ease of implementation, and low time complexity, KNN is one of the most popular classification algorithms for the wrapper approaches.

TLBO is a popular social-inspired metaheuristic algorithm that was first introduced by Rao et al. [18]. Two phases of the optimizer are "Learner Phase" and "Teacher Phase", which bring superior performance for TLBO compared to other well-regarded algorithms when applied to different applications [19]. TLBO has been initially proposed to handle continuous optimization problems. To tackle FS, which is a binary optimization problem, TLBO requires adjustments and even new operators. The two-step binarization technique is popular in the literature utilized to transform continuous algorithms into binary form. In this technique, the fuzzy transfer functions are used firstly to map the continuous solutions into intermediate probability values within [0,1] while a binarization rule is applied as a second step to transform the intermediate solution into binary [20].

This work proposes an efficient wrapper-based feature selection approach that incorporates a modified binary TLBO as the search algorithm. This modification is accomplished in the algorithm at the level of the utilized binarization method in conjunction with two types of TFs. Four new binarization methods are introduced in our approach: the Elitist, the Elitist Roulette, the Elitist Tournament, and the Rank-based method. The influence of such methods is tested and compared to two other common binarization methods (i.e., the standard and the complement method).

The main contributions of this paper are summarized as follows:

- A new feature selection approach is proposed based on a modified binary TLBO.
- Four new binarization methods are introduced with TLBO: the Elitist, the Elitist Roulette, the Elitist Tournament, and the Rank-based method.

The rest of the paper is organized as follows: after introducing the main background in Section I, the recent FS approaches in the literature are analyzed, followed by a description of the used algorithms in this paper in Section II. A general overview of the TLBO algorithm is given in Section III. Section IV describes the details of the proposed approach. The results are discussed in Section V. Finally, the conclusion and the future directions are drawn in Section VI.

II. RELATED WORKS

There are a growing number of problems that need to be solved by analytical methods [21, 22, 23, 24, 25, 26, 27, 28]. Recently, various Swarm Intelligence (SI) algorithms have been utilized in various fields as alternative approaches [29, 30, 31, 32]. One of the areas is as search strategies in different wrapper FS methods [33, 34, 35]. As a primary SI algorithm, PSO has been widely used with FS methods. A combination of PSO and a micro GA approach was proposed by Mistry et al. [36] to perform FS. Another FS approach that is based on PSO-GA algorithms with the adaptive neuro-fuzzy inference systems (ANFIS) was proposed by Semero et al. [37]. Tran et al. [38] proposed first variable-length PSO to handle the feature selection problem. In addition, Wu et al. [39] solved the FS problem using a hybrid improved quantum-behavior PSO. Furthermore, a multi-objective PSO was used by Zhang et al. [40] to solve the feature selection problems. Mafarja et al. [41] and Mafarja and Sabar [42] proposed two recent approaches that employed two variants of PSO algorithm as searching strategies in wrapper FS methods. Also, a hybrid approach between PSO and Shuffled Frog Leaping Algorithm (SFLA) was proposed in [43] to improve the accuracy of fake reviews identification. Chen et al. [44] proposed an enhanced PSO approach with two crossover operators to tackle FS problems. De Souza et al. [45] proposed a new wrapper approach based in a v-shaped transfer function using one of recent meta algorithm called Crow Search Algorithm (CSA), the accuracy results of their approach were very good results. Ant Colony Optimization (ACO) algorithm was also applied in many FS methods. For instance, Shunmugapriya and Kanmani [46] proposed a hybrid FS approach that combines the characteristics of ACO with Artificial Bees Colony (ABC) (called AC-ABC) to enhance the search process. In AC-ABC, the ACO algorithm employs bees in the exploitation process, while ABC uses the ants as food sources in the search process. A combination of a modified binary coded ACO algorithm with GA was proposed by Wan et al. [47] as an FS method called MBACO. In MBACO, GA was used to generate either the visibility information or the initial pheromone information. Manbari et al. [48] proposed a filter FS approach that is based on a modified version of the binary ACO algorithm with a combination with a clustering technique.

The Salp Swarm Algorithm (SSA) is a recent metaheuristic algorithm that mimics the behavior of salps in nature. Although the SSA is still new, it has been used as a search strategy in many FS approaches. Aljarah et al. [49] and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3064799, IEEE Access

Faris et al. [50] proposed two SSA-based FS methods. The experimental results in both works proved the ability of the SSA to outperform other optimizers. Moreover, another SSA-based approach was proposed in [51]. In this approach, a set of chaotic maps is used to control the balance between exploration and exploitation in the SSA algorithm. Sayed et al. [52] proposed a chaotic based SSA for global optimization and FS.

In addition to the above-mentioned works, in which SI algorithms have been used as search strategies in FS methods, another algorithm widely used in this area is called Sine Cosine Algorithm (SCA) [53], which works based on sine and cosine functions in moving the positions of the solutions in the search space. Sindhu et al. [54] proposed a novel FS method that is based on an Improved SCA variant called (ICSA). In ICSA, an elitism strategy was used to select the global solution, and a new updating mechanism for the new solution was proposed. As other global optimization algorithms, SCA suffers from the stagnation in local optima. To overcome this drawback, Elaziz et al. [55] proposed a hybrid model between the SCA and the DE's operators that served as a local search method. This hybrid model helps the SCA algorithm to skip local optima.

Recently, a wide range of metaheuristics have been studied and integrated into different FS approaches [56]. One of the most interesting point about these approaches that they tend to significantly outperform the traditional approaches [57, 58]. For instance, Arora and Anand [59] proposed two FS approaches based on the binary Butterfly Optimization Algorithm (BOA), in wihch two transfer functions were used to convert the continuous version of the BOA to binary. In [60], another FS approach that is based on the binary Brain Storm Optimization (BSO) was proposed. In their work, the authors proposed eight variants of the BBSO by employing eight different transfer functions. The same algorithm (i.e., BSO) has been recently used in another FS approach by Pourpanah et al. [61]. A combination of BSO and the Fuzzy ARTMAP (FAM) model was proposed where the BSO was used as a selection strategy to search for the optimal feature subset from the prototype nodes that were incrementally produced by the FAM model. Ten datasets were used to evaluate the proposed BSO-FAM model, and the results were promising. A filter FS approach that is based on a binary version of the Differential Evolution (DE) as a searching strategy, and on the entropy as an evaluator, was proposed in [62].

In the past decades, metaheuristic algorithms were shown to be very successful for solving various optimization problems [63, 64, 65, 66, 67]. TLBO is a recent, nature-inspired metaheuristic, that has been widely used in tackling different optimization problems in many fields and different real-life applications [68]. Despite some drawbacks highlighted by Črepinšek et al. [69], Waghmare [70], Pickard et al. [71], Chinta et al. [72], many variants of TLBO have been proposed to tackle the FS problem in recent years. For instance, a multi-objective TLBO version, with different update mechanisms was proposed in [73] to find Pareto-optimal set of solutions for a multi-objective formulation of the FS problem. Another binary TLBO version was used with varying algorithms of classification in a wrapper FS approach in [74]. Moreover, Sevinç and Dökeroğlu [75] proposed a TLBO FS approach with the Extreme Learning Machines (ELM), called TLBO-ELM. For more details about the TLBO based methods, readers can refer to the surveys conducted by Rao [76] and Zou et al. [68] and the book written by Rao [77].

In the previous FS approaches, either the algorithm is binary by itself (e.g., GA), or a conversion method such as Transfer Function (TF) was used to convert the continuous feature vectors into binary in the internal process of the algorithms. In literature, there are two basic types of TFs: in the first one, the sigmoid function that was used by [78] to convert the PSO into a successful binary version. The second TF was called V-shaped TF, which was used with Gravitation Search Algorithm (GSA) by Rashedi et al. [79]. The main idea behind using the TFs is to utilize them as a conversion method based on a defined probability for updating each element in the continuous representation of the solution into 1 or 0 according to this probability. Following this step, a binarization rule is applied to map the value of TF into a binary one. The most commonly used techniques for this step are the standard and complement methods. In this work, we extend this research direction by proposing four new binarization methods and explore their effectiveness in combination with both V-shape and S-shape TFs.

III. TEACHING LEARNING-BASED OPTIMIZATION (TLBO)

TLBO is a successful human-inspired optimizer classified under the umbrella of metaheuristic methods [80]. Initially, Rao et al. [19] tried to mimic the communications and interactions between teachers and students in a classroom or any other location for developing a metaheuristic approach. In population-based TLBO, the population of students, which is also called learners, plays the role of search agents, while the teacher leads the search agents. The fitness value of each agent shows the level of that learner' results during the learning (optimization) process. The subjects that the teacher (a learner with the highest score) teaches are treated as the decision variables of the optimization problem. In TLBO, the exploratory and exploitative phases are done during two core processes: Teacher phase and Learner phase. In the teacher phase, the learning of the agents occurs based on the knowledge of teacher (leader) himself, while, the second phase is devoted to the interaction between the learners (following agents).

A. TEACHER PHASE

In this phase, the purpose is to increase the average grades of the learners in the classroom concerning the personal knowledge of the teacher. Hence, the best learner is selected as the teacher, which is the position of a learner agent with the lowest fitness value in a minimization scenario. Also, the average position of all agents is obtained. Then, the positions of all agents are updated using Eq. (1):

$$DM_{j,i} = r \times (X_{j,kbest,i} - T_f \times M_{j,i}) \tag{1}$$

$$X^{new}_{j,k,i} = X^{old}_{j,k,i} + DM_{j,i} \tag{2}$$

where i is iteration, j is the subject (dimension) ($j = 1, \ldots, m$), k is the learner (search agent) ($k = 1, \ldots, n$), r is a random number inside (0,1), $X_{j,kbest,i}$ is the score of the teacher in subject j, $M_{j,i}$ denotes the average score of all learners in subject j, $DM_{j,i}$ denotes the difference between the teacher score and the updated average score of the learner agents in each subject, $X_{j,k,i}$ denotes the score of learner k in subject j, $X^{new}_{j,k,i}$ is the updated position of the old position vector $X^{old}_{j,k,i}$, and T_f denotes the teaching factor, which is obtained as rule in Eq. (3):

$$T_f = round[1 + r'] \tag{3}$$

where r' is a random number inside (0, 1). Note that the value of T_f is 1 or 2 based on the obtained random value. Where T_f is set to 1 when $r' < 0.5$ and 2 when $r' \geq 0.5$. The T_f parameter controls the neighborhood size in the search space, which affects the exploitation and exploration abilities of the TLBO algorithm.

B. LEARNER PHASE

In the second phase, the way the learners interact with each other's is considered. The fact is that a learner can also acquire the information from other superior learners in the class. If we have two distinct learners, p and q, which is denoted by X_p and X_q, we can choose one of them randomly. Hence, the updated status of the learner X_p can be obtained using Eq. (4):

$$X^{new}_{j,p,i} = \begin{cases} X^{old}_{j,p,i} + r'' \left(X^{old}_{j,p,i} - X^{old}_{j,q,i}\right) & f(X_p) < f(X_q) \\ X^{old}_{j,p,i} - r'' \left(X^{old}_{j,q,i} - X^{old}_{j,p,i}\right) & f(X_q) < f(X_p) \end{cases} \tag{4}$$

where r'' is a random number inside (0,1), and $f(X_p)$ and $f(X_q)$ are the fitness values of X_p and X_q agents, respectively. Based on this rule, only the better quality agents are saved to be improved in the next iterations.

The pseudo-code of continuous TLBO is shown in Algorithm 1.

Algorithm 1 Pseudo-code of TLBO

```
Initialize number of agents N, dimensions D, and number of iterations (L)
Generate the candidate solutions (learners) X_i(i = 1, 2, ..., N)
Obtain the fitness value of all N agents
Set X_T as the best agent
Set l = 1
while (l ≤ L) do                                         ▷ Teacher phase
    Set the best learner as X_Teacher
    Obtain the mean value across the D design variables
    for (each learner (X^new_{j,k,i})) do
        Obtain T_f using Eq. (3)
        Update the positions using Eqs. (1) and (2)
    end for
    Evaluate the new learners
    Save the new agents if they are superior to the old one
    for (each learner (X^new_{j,k,i})) do                 ▷ Learner phase
        Randomly choose another learner
        Update the current agents using Eq. (4)
    end for
    Assess the new learners
    Save the new agents if they are superior to the old one
    Update X_T if there is a superior agent
    l = l + 1
end while
Return X_T
```

IV. THE PROPOSED APPROACH

The majority of metaheuristic algorithms have been proposed to optimize continuous optimization problems. To tackle binary optimization problems (e.g., FS), these algorithms require adjustments and even new operators. In the literature, three main groups of binarization techniques are used to convert continuous algorithms into the binary form. The first group is called the two-steps binarization techniques, in which the operators of the algorithms remain unchanged, and two steps take place to convert the continuous solution into the binary one after the original continuous iteration. In the second group called the continuous-binary operator transformation, however, the operators of the algorithm are reformulated, and the algebra of the search space is redefined [20]. Moreover, in the third category, a novel binarization method, that is based on a clustering technique (called K-means Transition Algorithm (KMTA)), was recently proposed by García et al. [81] as a general binarization method.

Transfer Functions (TF) and binarization are two-steps techniques that have been widely used to convert the continuous search space to binary pair in many algorithms (e.g., PSO [82], GSA [79]). In this technique, the TF is considered as the first step, which aims to produce an intermediate solution, with values in the interval [0, 1], that defines the probability of converting the corresponding dimension in the original solution into zero or one. The second step in these techniques is the binarization, where a binarization rule is applied to map the intermediate solution into a binary solution.

Kennedy and Eberhart [82] introduced the use of the sigmoid function (as in Eq. 5) to transform the continuous PSO into a binary version. In 2010, Rashedi et al. [79] introduced the use of the tanh function (as in Eq. (6)) to binarize the GSA. These two TFs belong to two different families that have distinguished based on their shape. These families were called the S-shaped (as in Fig. 1a) and the V-shaped (as in Fig. 1b).

$$T(x_j^i(t)) = \frac{1}{1+e^{-x_j^i(t)}} \tag{5}$$

$$T(x_j^i(t)) = |\tanh(x_j^i(t))| \tag{6}$$

In these works, two binarization methods were used; the standard and complement methods. In the standard techniques (see Eq. (7)), which was first used with the S-shaped TF as in Kennedy and Eberhart [82], a random number is generated, if its value is less than the probability value of the i^{th} element of the intermediate solution at the k^{th} iteration, then, i^{th} element of the binary solution is set to 1, otherwise, it is set to zero. In the complement method (see Eq. (8)), which was used with the V-shaped TF as in Rashedi et al. [79], the values (0 or 1) of the binary solution are set based on the benefits of the current solution, that is to say, based on the probability value ($T(v_i^k(t))$), the i^{th} element is either kept the same or flipped.

$$X_i^k(t+1) = \begin{cases} 1 & r < T(x_i^k(t)) \\ 0 & Otherwise \end{cases} \tag{7}$$

$$X_i^k(t+1) = \begin{cases} \backsim X_i^k(t) & r < T(x_i^k(t)) \\ X_i^k(t) & Otherwise \end{cases} \tag{8}$$

where r is a random number in [0, 1] interval.

In both TFs groups (i.e., S-shaped and V-shaped), the probability of updating the solution's element to 0 or 1 mainly depends on the step vector, which is considered as the only input to the TF. A higher probability value indicates that this solution is far from the best solution so far and requires an abrupt change (exploration). In contrast, a lower value indicates that the individual is very close to the best solution and requires smaller steps (exploitation) [83]. Therefore, the TF plays a significant role in balancing between exploration and exploitation for binary algorithms since different TFs have different behaviors when calculating the probability of updating the solution's element.

Mirjalili and Lewis [84] considered the same assumption of Kennedy and Eberhart [82] and Rashedi et al. [79], and used the standard Binarization Methods (BM) with four S-shaped functions, and the complement BM with four V-shaped functions. The standard method sets the solution's elements to 0 or 1 based on the calculated probability from the TF regardless of the current value in the solution. Which means that the solution may remain in its current position while we need to move it to achieve the exploration, and its position may be changed while we need to keep it to achieve the exploitation. However, the complement method considers the current value of the position to set the new value. For the large probability values, the solution is flipped to move it into a different region, while the small probability values keep the position value as is.

The main difference between the standard and the complement methods is the binarization mechanism, and revealed different results when used with different TFs. After a careful literature review, we found that most of the previous studies considered different TFs, while a few binarization methods were used. However, both TFs and binarization methods have a significant impact on the effectiveness of the optimization algorithm. Our experiments show that both using a suitable binarization mechanism with a TF has a substantial impact on the exploitative and exploratory potentials of the utilized binary algorithm. This motivated our attempts to propose different binarization methods.

As mentioned above, in both standard and complement methods, the updating mechanisms do not consider the best solution so far. Because the intermediate solution is a mutation probability of changing the solution and is based on the behavior of the evolutionary algorithms, the best solution so far (called elitist) may be used to re-position the current solution.

In this paper, four different binarization methods that consider other solution than the current one in the re-positioning process are proposed. In the proposed approaches, the guide solution is selected based on different selection criteria; best selection, where the solution with the best fitness value (called elitist) is selected, Roulette Wheel Selection (RWS) [85], Tournament Selection (TS) [86] and finally based on the solution's rank compared to other solutions in the population. Eq. (9) represents the general formula for using a selected solution to update the position of the current one. The mutation probability is calculated using the TF based on the selected solution. If a random number is less than that value, the dimension of the new solution will be the complement of the corresponding one of the selected solution. Otherwise, it will be set to the actual value of the selected solution.

$$X_{new}^k(t+1) = \begin{cases} \backsim X_{selected}^K(t) & r < T(x_i^k(t+1)) \\ X_{selected}^K(t) & Otherwise \end{cases} \tag{9}$$

where $\backsim$ represents the complement, $x_{selected}^K$ is the corresponding value of the selected solution.

The following remarks represent the brief description of the four BMs proposed in this paper:

1) BTLBO_E: Elitist method, where the best solution so far, according to the fitness value, is selected. In this mechanism, the position of the solution being processed is changed towards or away from the best solution. As the FS is a minimization problem, the solution with the minimum fitness value is selected. According to Eq. (9), if r is lower than $T(v_i^k(t+1)$, then, the solution is moved far from the best solution. Otherwise, the move will be towards that solution.
2) BTLBO_ERW: The name of this method is given based on the concept of Elitist Roulette. In this method, the selection process is based on the RWS mechanism. A chance to the other solutions in the population is given by employing the RWS to avoid moving all agents towards the best solution, especially in the last stages of the search process. Based on this fact, it gives a probability (p) for each solution to be selected according to its fitness value, where p is calculated

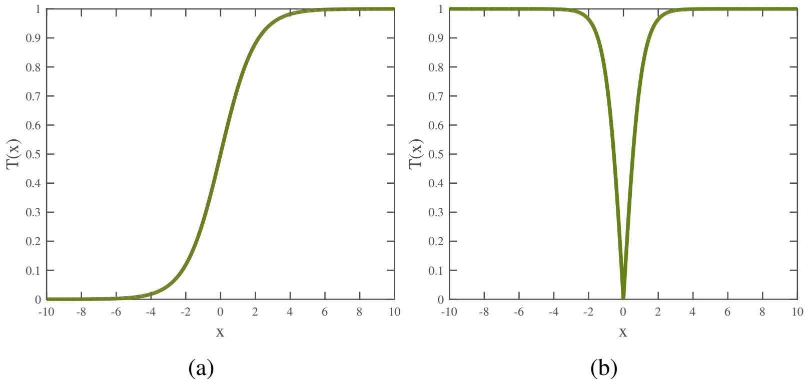


FIGURE 1: S-shaped and (b) V-shaped TFs

according to Eq. (10). Then, the selected solution is considered as a guiding solution in Eq. (9).

$$p_i = \frac{f_i}{\sum_{j=1}^{n} f_j} \tag{10}$$

where f_i is the fitness of the i^{th} solution, and n represents the population size.

3) BTLBO_ET: The name of this method is given based on the concept of Elitist Tournament. In this method, the TS mechanism is utilized to select a guiding solution instead of selecting the best one. In this mechanism, a set (with size τ) of solutions, which is called tournament, is randomly selected, then, the best solution in the tournament is picked up as the guiding solution. Then, the selected solution is considered as a guiding solution in Eq. (9). Figure 2 illustrates the process of selecting a solution following the TS mechanism.

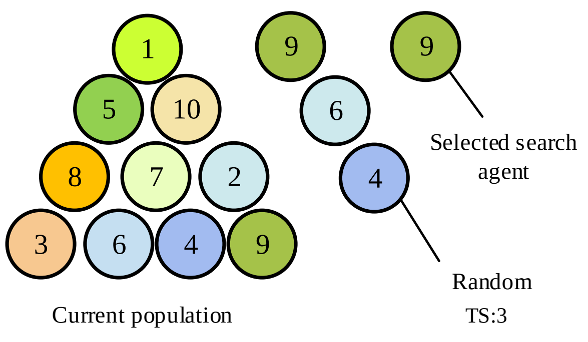


FIGURE 2: Tournament Selection mechanism

4) BTLBO_ER Rank-based method: Each solution in the population has a probability to be selected based on its rank in terms of the fitness value. In this method, each solution is given a rank from 1 to n based on the fitness value, where the best solution is given the rank n (recall that n is the population size), while the worst solution is given a rank of 1. Then, the probability of selecting each solution is calculated based on Eq. (11).

$$p_i = \frac{rank_i}{n \times (n-1)} \tag{11}$$

where $rank_i$ represents the rank of the i^{th} solution. The advantages of this method are that each solution is given a chance to be selected since the ranks of the individuals are scaled. If the fitness of the fittest solution is much higher than that of others, it would be chosen probably in most of the iterations. This mechanism can help the proposed variant to avoid the premature convergence event.

To make fair comparisons, the two basic binarization methods (standard and complement) will be investigated as follows:

1) BTLBO_S: Standard method as defined in Eq. (7).
2) BTLBO_C: Complement Method as defined in Eq. (8).

A. BTLBO FOR FS

One of the significant issues that should be considered when designing an optimization algorithm is the solution representation. As the FS is a binary optimization problem, a binary vector (with a length that is equal to the number of features in the original dataset) is used to represent a solution to a FS problem where a zero indicates that the corresponding feature is not selected and a one means that the relevant element is selected. In this work, two TFs are used to transform the TLBO algorithm into binary based on six different binarization methods.

Eq. (12) represents the fitness function adopted in the proposed feature selection approaches. As it can be seen the equation, the fitness function incorporates two important objectives which are the miss-classification rate of the underlying classifier (i.e., KNN classifier [87], and the reduction rate in the number of selected features by the optimizer.

$$\downarrow Fitness = \alpha\gamma_R(D) + \beta\frac{|R|}{|C|} \tag{12}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3064799, IEEE Access

where $\gamma_R(D)$ is the classification error rate resulted by the underlying induction algorithm, $|R|$ is the number of selected features by the optimizer, and $|C|$ is the total number of features in the original dataset, and α and β are weighting constants. The latter two are used to quantify the importance of the main objectives, which are the accuracy and the reduction rate. The value of α is set in [0,1], while $\beta = (1 - \alpha)$ [88].

V. EXPERIMENTAL RESULTS AND SIMULATIONS

A. EXPERIMENTAL SETUP

Eighteen well-regarded datasets obtained from UCI repository [89] are employed here to study the effectiveness of the proposed binary TLBO variants. These problems were chosen carefully with various details and properties (e.g., number of features, instances, and classes) to cover varied types of real-life tasks. Table 1 describes a brief explanation for each employed dataset.

TABLE 1: List of datasets

Dataset	No.of Features	No.of instances
Breastcancer	9	699
BreastEW	30	569
Exactly	13	1000
Exactly2	13	1000
HeartEW	13	270
Lymphography	18	148
M-of-n	13	1000
PenglungEW	325	73
SonarEW	60	208
SpectEW	22	267
CongressEW	16	435
IonosphereEW	34	351
KrvskpEW	36	3196
Tic-tac-toe	9	958
Vote	16	300
WaveformEW	40	5000
WineEW	13	178
Zoo	16	101

The same hardware and operating system configuration have been used to have a fair study. Details have been reported in Table 2.

TABLE 2: The system properties

Name	**Setting**
Hardware	
CPU	Intel Core(TM) i5-3210M
Frequency	2.5GHz
RAM	4GB
Hard drive	500 GB
Software	
Operating system	Windows 7
Language	MATLAB R2018a

All the optimizers are assessed using the same common configurations and settings (α = 0.99, β = 0.01, Number of runs = 30, and number of agents = 40, number of fitness function calls), as reported in Table 3. Please note that these settings were obtained from well-known FS approaches in the literature [90, 91] Since the TLBO algorithm calls the fitness function two times in each iteration, we executed it for the half number of iterations of the other algorithms. For the specific configurations mentioned in Table 3, we used the recommended values by other researchers in different papers, for instance, Rashedi et al. [79] recommended the value 10 for the parameter G_0 in BGSA, while the a parameter was recommend by Mirjalili et al. [7] to be from 2 to 0. The parameter values for the BBA algorithm were obtained from Mirjalili et al. [92]. The same case is with the parameters of the WOA algorithm which ordained form [5]. Because the experiments in this paper are devoted to meta-heuristic methods which incorporate randomness, we present the average results using 30 independent runs on each dataset. For for the value of K in KNN, previous works recommended that $K = 5$ so it was set to this value int this work for fair comparison as well [79, 88, 91, 93].

Please note that **bold** values in all reported tables show the best-obtained results. To identify if there is a significant difference between the solutions of different variants and competitors, we performed a Wilcoxon non-parametric statistical test [94] with significance level of 0.05. In order to judge the normality assumption of Wilcoxon test, we conducted Shapiro-Wilk (SW) test as a powerful and recommended procedure in the literature [95]. If the SW test is not applicable (i.e the sample standard deviation is zero), we performed Kolmogorov-Smirnov (KS) test.

TABLE 3: Experimental setup

Config. Name	Value
Fitness function	
α	0.99
β	0.01
Common Config.	
Number of runs	30
Number of agents	40
Number of iterations (for TLBO)	50
Number of iterations (for other optimizers)	100
Specific Config.	
G_0 (for BGSA)	10
a (for bGWO)	from 2 to 0
Q_{min} Frequency minimum (for BA)	0
Q_{max} Frequency maximum (for BA)	2
A Loudness (for BA)	0.5
r Pulse rate (for BA)	0.5
a (for WOA)	from 2 to 0
$a2$ (for WOA)	from -1 to -2
K for KNN	5
t for Tournament selection	10

B. RESULTS AND DISCUSSIONS

In this section, various extensive experiments are performed, and the results are presented in details to find the best variant of proposed BTLBO for solving FS datasets. First, we investigate the impact of each binarization method on the performance of the binary TLBO with S-shaped TFs according to different metrics. By these experiments, we can find the best binarization technique when using S-shaped TFs.

1) Different binarization methods with S-shaped TFs

Table 4 shows the accuracy results obtained using different binarization methods with S-shaped TFs. As per F-test results in Table 4, it is observed that the BTLBO_ET has attained the best results. It also provides 100% accuracy on 33.33% of datasets. It can be seen that there is a competition between the BTLBO_E, BTLBO_ERW, BTLBO_ET, and BTLBO_ER variants in terms of accuracy rates, while BTLBO_S and BTLBO_C variants show similar overall efficacy.

TABLE 4: Comparison between different binarization methods with S-shaped TFs in terms of average accuracy.

Benchmark	Measure	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ET	BTLBO_ER
Breastcancer	AVG	**1.0000**	0.9857	**1.0000**	0.9929	0.9786	0.9786
	STD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BreastEW	AVG	0.9851	0.9784	0.9936	0.9789	0.9877	**1.0000**
	STD	0.0041	0.0055	0.0039	0.0044	0.0055	0.0000
CongressEW	AVG	0.9885	0.9881	**1.0000**	0.9801	**1.0000**	0.9885
	STD	0.0000	0.0021	0.0000	0.0052	0.0000	0.0000
Exactly	AVG	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
	STD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Exactly2	AVG	0.7437	0.7873	0.7483	0.7820	**0.7995**	0.7463
	STD	0.0081	0.0134	0.0040	0.0173	0.0115	0.0043
HeartEW	AVG	0.9019	0.8741	0.8488	**0.9210**	0.8957	0.9086
	STD	0.0099	0.0141	0.0070	0.0083	0.0091	0.0161
IonosphereEW	AVG	0.9676	**0.9803**	0.9657	0.9244	0.9761	0.9775
	STD	0.0099	0.0079	0.0080	0.0069	0.0066	0.0070
KrvskpEW	AVG	0.9716	0.9781	0.9715	0.9763	0.9768	**0.9791**
	STD	0.0042	0.0044	0.0037	0.0053	0.0037	0.0045
Lymphography	AVG	0.9311	0.9398	**0.9589**	0.9539	0.9344	0.8877
	STD	0.0085	0.0134	0.0143	0.0163	0.0138	0.0163
M-of-n	AVG	0.9998	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
	STD	0.0009	0.0000	0.0000	0.0000	0.0000	0.0000
penglungEW	AVG	**1.0000**	**1.0000**	0.8565	0.8730	**1.0000**	0.9867
	STD	0.0000	0.0000	0.0228	0.0194	0.0000	0.0271
SonarEW	AVG	0.9802	0.9706	0.9857	**0.9992**	0.9976	0.9825
	STD	0.0110	0.0135	0.0119	0.0043	0.0073	0.0107
SpectEW	AVG	0.8914	0.9222	**0.9333**	0.8031	0.8599	0.9321
	STD	0.0106	0.0075	0.0104	0.0091	0.0093	0.0101
Tic-tac-toe	AVG	0.8385	0.8385	**0.8542**	0.8333	0.8281	0.8125
	STD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Vote	AVG	0.9522	0.9833	0.9844	**1.0000**	0.9878	0.9728
	STD	0.0058	0.0000	0.0042	0.0000	0.0075	0.0082
WaveformEW	AVG	0.7501	0.7513	0.7475	0.7513	**0.7609**	0.7532
	STD	0.0066	0.0081	0.0065	0.0049	0.0065	0.0060
WineEW	AVG	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
	STD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Zoo	AVG	**1.0000**	0.9524	**1.0000**	**1.0000**	**1.0000**	**1.0000**
	STD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Ranking	Best	5	5	9	7	8	6
Overall Ranking	F-Test	3.9167	3.6389	3.3333	3.5556	**3.0833**	3.4722

Table 5 compares the average number of features attained by different binarization methods with S-shaped TFs. According to the number of features, the BTLBO_E has shown the best efficacy, while BTLBO_ET has attained the next place.

Table 6 shows the average fitness values attained by different binarization methods with S-shaped TFs. Regarding the fitness results, the best variant is BTLBO_E technique. It has attained the minimum results on 44.44% of problems. We observe that the BTLBO_ET version is placed at the second stage.

Table 7 shows the average running time obtained by different binarization methods with S-shaped TFs. Based on running time, the fastest variant is BTLBO_S, while BTLBO_E and BTLBO_ERW are in the next stages.

The p-values of the normality test for accuracy results of variants with S-shaped TF are presented in Table 8. It is evident that most of the cases the p-value is less than 5% and the null hypothesis is rejected. This fact shows that there is evidence that the results of the different variants are not normally distributed.

Table 9 shows the p-values of the Wilcoxon test for the accuracy results of BTLBO-ET versus other techniques with S-shaped TF. The p-values evidently show that the recorded

TABLE 5: Comparison between different binarization methods with S-shaped TFs in terms of average number of features

Benchmark	Measure	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ET	BTLBO_ER
Breastcancer	AVG	**4.0000**	6.0000	7.0000	**4.0000**	**4.0000**	**4.0000**
	STD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BreastEW	AVG	12.4000	14.9000	12.5333	13.5000	11.9667	**11.6667**
	STD	3.1360	1.9888	2.3742	2.2399	2.2512	1.9357
CongressEW	AVG	5.4667	6.8667	5.5000	5.5000	6.3667	**5.4000**
	STD	1.2794	1.4794	0.9002	1.1963	0.7184	0.9322
Exactly	AVG	6.4667	6.3667	**6.2333**	6.4667	6.3333	6.4667
	STD	0.5074	0.4901	0.4302	0.5074	0.4795	0.5074
Exactly2	AVG	8.6333	8.3667	**4.7667**	8.5667	9.5000	7.9000
	STD	1.9561	2.5255	3.9713	2.1922	0.5724	1.4704
HeartEW	AVG	5.8667	5.8667	5.6333	6.7667	6.0333	**4.3667**
	STD	0.9371	1.0080	1.5862	1.0063	1.2726	1.2994
IonosphereEW	AVG	**10.9000**	13.5667	12.2333	12.8000	12.6667	12.4333
	STD	1.7685	2.1284	2.2997	2.7468	2.5641	2.2997
KrvskpEW	AVG	21.1000	20.8333	20.1667	21.4667	**18.8000**	22.2000
	STD	2.4544	2.7926	2.2450	2.5962	2.5784	3.0783
Lymphography	AVG	8.8667	7.7333	9.0000	8.6667	8.4667	**7.3000**
	STD	1.4559	1.3629	1.8383	1.2685	1.2521	1.6006
M-of-n	AVG	6.7667	6.7000	**6.2667**	6.4333	6.3000	6.4667
	STD	0.6261	0.5350	0.4498	0.5040	0.4661	0.5074
penglungEW	AVG	**125.1667**	132.3667	136.0667	135.2000	126.1667	142.0667
	STD	4.0606	6.4833	12.4123	8.7628	4.5719	17.0009
SonarEW	AVG	25.5667	27.3000	25.1333	**25.0000**	28.3000	27.1667
	STD	3.2129	3.2499	4.0830	2.4069	4.1285	2.4647
SpectEW	AVG	8.5333	10.8667	8.9667	**6.7000**	8.2333	11.0000
	STD	1.8333	2.0126	1.4735	2.0869	1.9241	2.2743
Tic-tac-toe	AVG	6.0000	6.0000	6.0000	6.0000	**5.0000**	6.0000
	STD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Vote	AVG	5.2000	**4.3000**	4.9667	6.3333	5.1667	5.0333
	STD	1.6274	0.9523	0.8899	0.8442	1.3153	1.2726
WaveformEW	AVG	**19.8333**	20.4000	19.8667	22.9667	20.9333	21.5000
	STD	2.5063	2.1107	3.3190	2.8343	2.9353	2.7885
WineEW	AVG	5.0000	4.5667	**2.1333**	5.7333	4.3333	3.7000
	STD	0.0000	0.5683	0.3457	0.6915	0.5467	0.5960
Zoo	AVG	6.0000	4.5000	**3.2000**	3.8667	3.5000	4.9667
	STD	0.5872	0.5085	0.4068	0.5074	0.5085	0.6149
Ranking	Best	4	1	5	3	3	5
Overall Ranking	F-Test	3.3889	3.0278	**4.2500**	2.8889	3.8611	3.5833

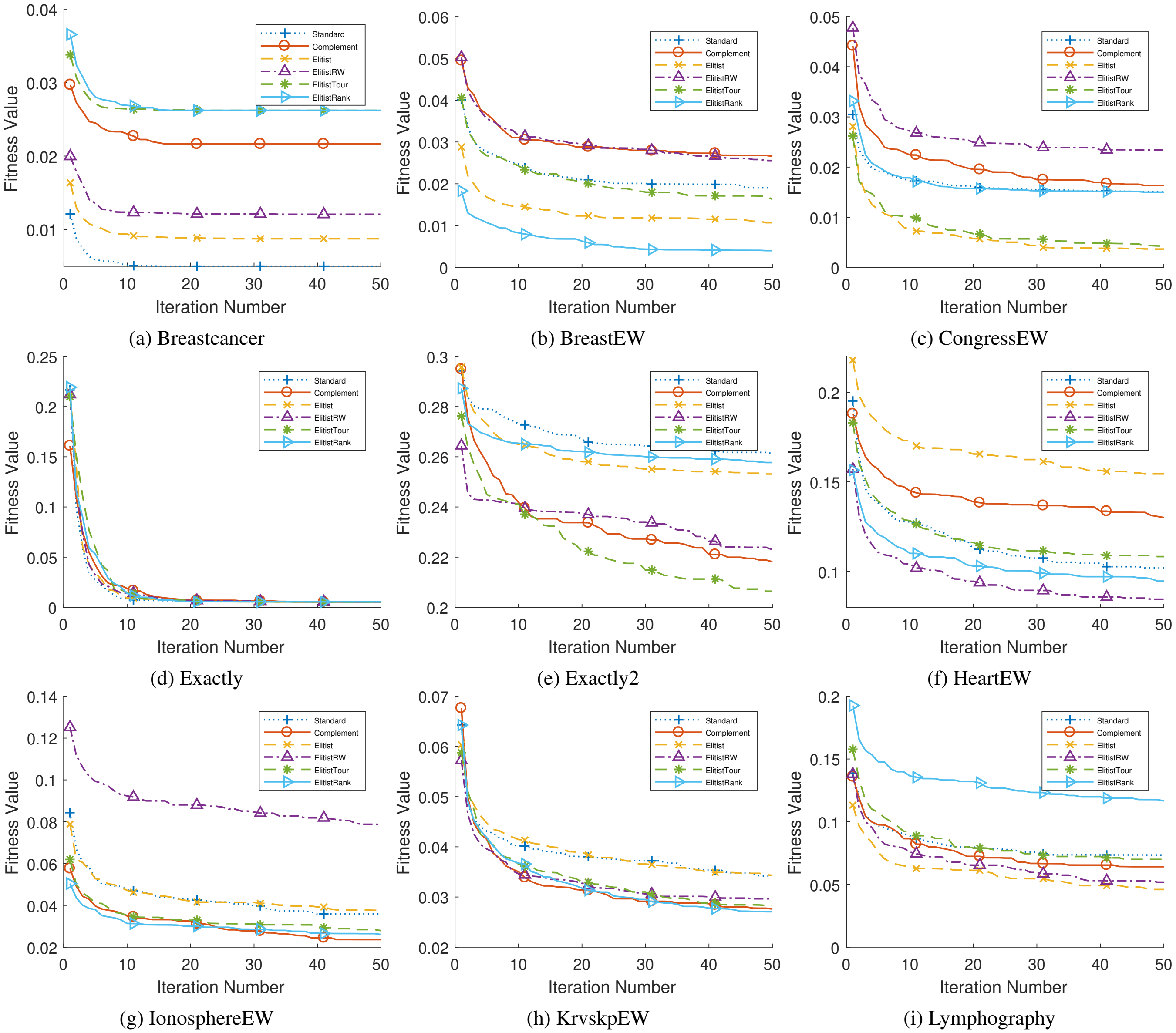
differences between the accuracy rates of the BTLBO-ET and other variants with S-shaped TFs are significantly meaningful in most of the cases.

Figures 3 and 4 demonstrate the convergence curves for BTLBO with different binarization approaches for S-shaped TFs in dealing with all datasets. According to convergence plots, firstly, it can be seen several patterns in convergence of different methods, while for some datasets like Exactly and M-of-n, the patterns are similar and there is a competition between different variants. Secondly, some variants show more stagnation drawbacks. If we consider all curves, it can be seen that the BTLBO_E technique has shown the fastest trends for majority of datasets. After BTLBO_E, the BTLBO_ERW variant also shows the second best convergence rate.

As per the average number of features and fitness values, it can be seen that the elitist method is the fittest binarization technique in the case of S-shaped TFs. The elitist approach also led to the best accuracy rates on nine datasets. This observation shows that when using S-shaped TFs, BTLBO with elitist method shows the best efficacy compared to other variants with other binarization techniques.

2) Different binarization methods with V-shaped TFs

In this subsection, we study the impact of each binarization method on the performance of the binary TLBO with V-shaped TFs using different performance measures. By these



(a) Breastcancer (b) BreastEW (c) CongressEW

(d) Exactly (e) Exactly2 (f) HeartEW

(g) IonosphereEW (h) KrvskpEW (i) Lymphography

FIGURE 3: Convergence curves for BTLBO with different binarization methods for S-shaped TFs on Breastcancer, BreastEW, CongressEW, Exactly, Exactly2, HeartEW, IonosphereEW, KrvskpEW, and Lymphography datasets.

experiments, it can be recognized as the most appropriate binarization approach when using V-shaped TFs.

Table 10 compares the accuracy results obtained by different binarization methods with V-shaped TFs. Based on accuracy rates in Table 10, the BTLBO_ER has scored first (see F-test results), whereas BTLBO_ERW also obtained the best results on 38.88 % of datasets. It is evident that BTLBO_ET has attained the best results on 50% of cases. Also, it can be seen that the BTLBO_C and BTLBO_E variants show no superiority on each other and has obtained the same overall place. If we consider the BTLBO_S variant, we observe that it is the last preference based on the accuracy results.

Table 11 exposes the average number of features found by different binarization methods with V-shaped TFs. As per number of features in Table 11, it can be seen that the method with lowest accuracy, BTLBO_S, is the best performing variant (superior results on 38.88%) in terms of average number of features.

Table 12 presents the average fitness results found by different binarization methods with V-shaped TFs. As per results in Table 12, we observe that BTLBO_ET has attained the minimum results on 38.88 % of cases, while BTLBO_ERW and BTLBO_ER are in the next places by finding the best results on 27.77% of problems. Based on F-test results, the BTLBO_ER is the ranked one approach, whereas BTLBO_ET, BTLBO_ERW, BTLBO_C, BTLBO_E, and BTLBO_S are in the next preferences, re-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3064799, IEEE Access

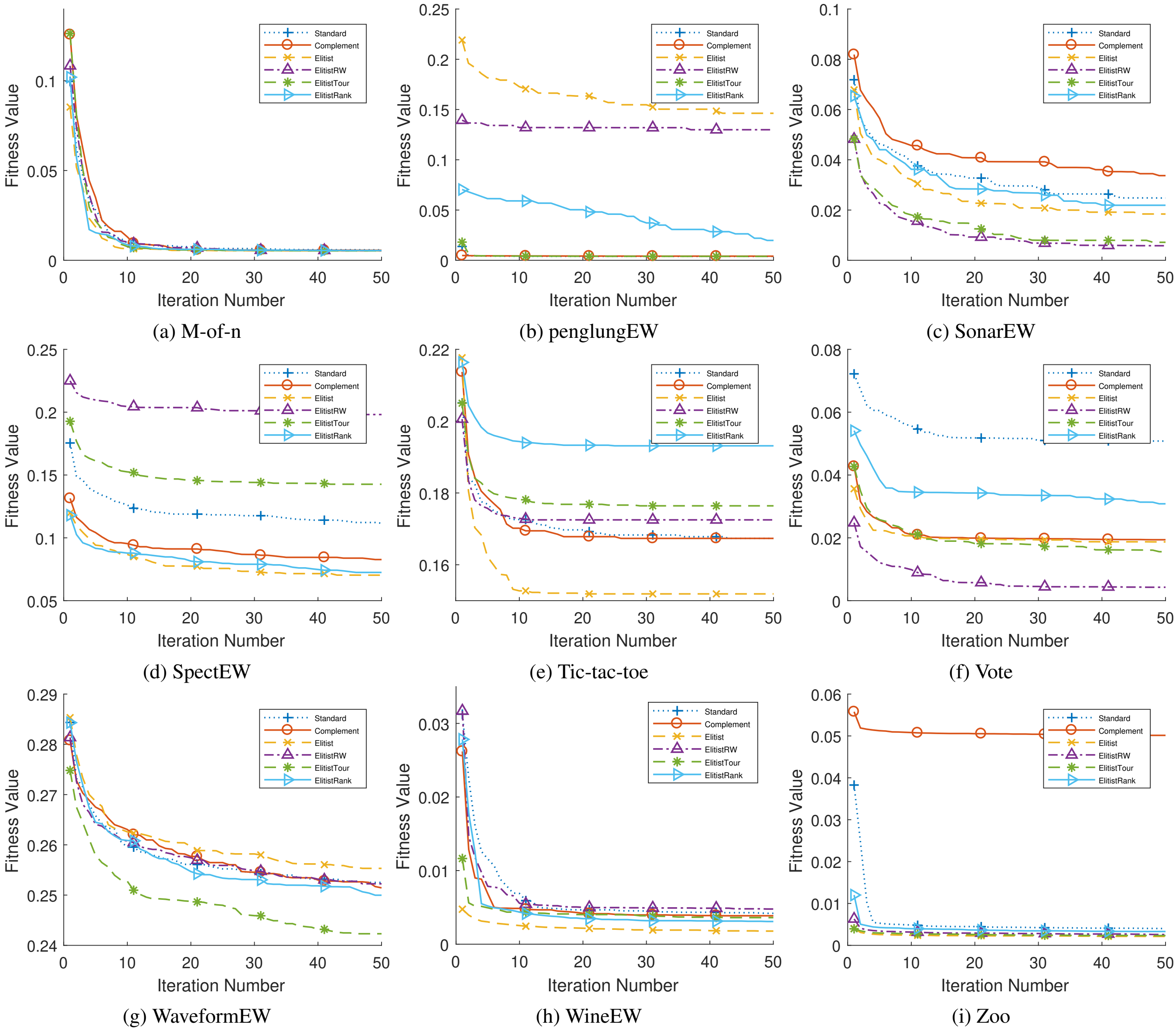


FIGURE 4: Convergence curves for BTLBO with different binarization methods for S-shaped TFs on M-of-n, penglungEW, SonarEW, SpectEW, Tic-tac-toe, Vote, WaveformEW, WineEW, and Zoo datasets.

spectively.

Table 13 shows the average running time spent by different binarization methods with V-shaped TFs. Based on CPU time analysis, the fastest version with V-shaped TFs on 83.33% of problems is still BTLBO_S, similarly to the observations in the variants with S-shaped TFs. For most of the cases, except the KrvskpEW, Tic-tac-toe, and WaveformEW, it is detected that the time gaps between various variants are not considerable.

The p-values of the normality test for accuracy results of variants with V-shaped TF are exposed in Table 14. We observe from Table 14 that the p-value is less than 5 % for most of the cases. Hence, the null hypothesis is not approved. This fact reveals that the obtained results follow a non-normal distribution.

Table 15 reveals the p-values of the Wilcoxon test for the accuracy results of BTLBO-ER compared to other peers when using V-shaped TF. The p-values clearly verify that the detected variations of the accuracy rates obtained by the BTLBO-ER and other variants with V-shaped TFs are statistically significant in most of the cases.

Figures 5 and 6 reveal the convergence behaviors for BTLBO with different binarization approaches for V-shaped TFs on all datasets. According to curves, it can be seen that BTLBO_ET shows the fastest rates in dealing with BreastEW, HeartEW, IonosphereEW, SpectEW, and penglungEW. As the next variants, the BTLBO_ERW and BTLBO_ER also show competitive rates on 27.77% of prob-

TABLE 6: Comparison between different binarization methods with S-shaped TFs in terms of average fitness

Benchmark	Measure	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ET	BTLBO_ER
Breastcancer	AVG	**0.0050**	0.0216	0.0088	0.0121	0.0262	0.0262
	STD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BreastEW	AVG	0.0190	0.0266	0.0107	0.0255	0.0163	**0.0040**
	STD	0.0037	0.0052	0.0035	0.0040	0.0050	0.0007
CongressEW	AVG	0.0150	0.0163	**0.0037**	0.0234	0.0042	0.0150
	STD	0.0009	0.0021	0.0006	0.0049	0.0005	0.0006
Exactly	AVG	0.0054	0.0053	**0.0052**	0.0054	0.0053	0.0054
	STD	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
Exactly2	AVG	0.2610	0.2175	0.2531	0.2230	**0.2064**	0.2577
	STD	0.0070	0.0120	0.0057	0.0159	0.0117	0.0041
HeartEW	AVG	0.1021	0.1296	0.1544	**0.0839**	0.1083	0.0941
	STD	0.0100	0.0135	0.0069	0.0084	0.0084	0.0166
IonosphereEW	AVG	0.0354	**0.0236**	0.0376	0.0787	0.0275	0.0261
	STD	0.0097	0.0077	0.0078	0.0067	0.0065	0.0067
KrvskpEW	AVG	0.0341	0.0276	0.0340	0.0296	0.0283	**0.0271**
	STD	0.0040	0.0045	0.0034	0.0050	0.0034	0.0041
Lymphography	AVG	0.0734	0.0641	**0.0460**	0.0507	0.0700	0.1155
	STD	0.0079	0.0135	0.0135	0.0158	0.0135	0.0157
M-of-n	AVG	0.0058	0.0056	**0.0052**	0.0054	0.0052	0.0054
	STD	0.0012	0.0004	0.0004	0.0004	0.0004	0.0004
penglungEW	AVG	**0.0039**	0.0041	0.1463	0.1299	0.0039	0.0176
	STD	0.0001	0.0002	0.0223	0.0191	0.0001	0.0265
SonarEW	AVG	0.0240	0.0337	0.0184	**0.0050**	0.0072	0.0219
	STD	0.0107	0.0130	0.0114	0.0042	0.0069	0.0105
SpectEW	AVG	0.1116	0.0822	**0.0703**	0.1981	0.1426	0.0725
	STD	0.0102	0.0071	0.0100	0.0082	0.0087	0.0094
Tic-tac-toe	AVG	0.1673	0.1673	**0.1519**	0.1725	0.1764	0.1931
	STD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Vote	AVG	0.0508	0.0194	0.0187	**0.0042**	0.0155	0.0303
	STD	0.0047	0.0006	0.0038	0.0006	0.0070	0.0074
WaveformEW	AVG	0.2525	0.2514	0.2551	0.2521	**0.2421**	0.2498
	STD	0.0065	0.0079	0.0065	0.0047	0.0065	0.0059
WineEW	AVG	0.0042	0.0038	**0.0018**	0.0048	0.0036	0.0031
	STD	0.0000	0.0005	0.0003	0.0006	0.0005	0.0005
Zoo	AVG	0.0040	0.0501	**0.0021**	0.0026	0.0023	0.0033
	STD	0.0004	0.0003	0.0003	0.0003	0.0003	0.0004
Ranking	Best	2	1	8	3	2	2
Overall Ranking	F-Test	2.6944	3.2778	**4.1944**	3.1944	4.1667	3.4722

lems. Among other variants, it can be seen that BTLBO_S shows the repetitive stagnation problems on the majority of cases.

Referring to the average accuracy rates and fitness values, we recognize that the rank-based elitist strategy is the best performing binarization technique in the case of V-shaped TFs. This observation reveals that when using V-shaped TFs, BTLBO with rank-based elitist method demonstrates the best performance compared to other peers with different binarization techniques.

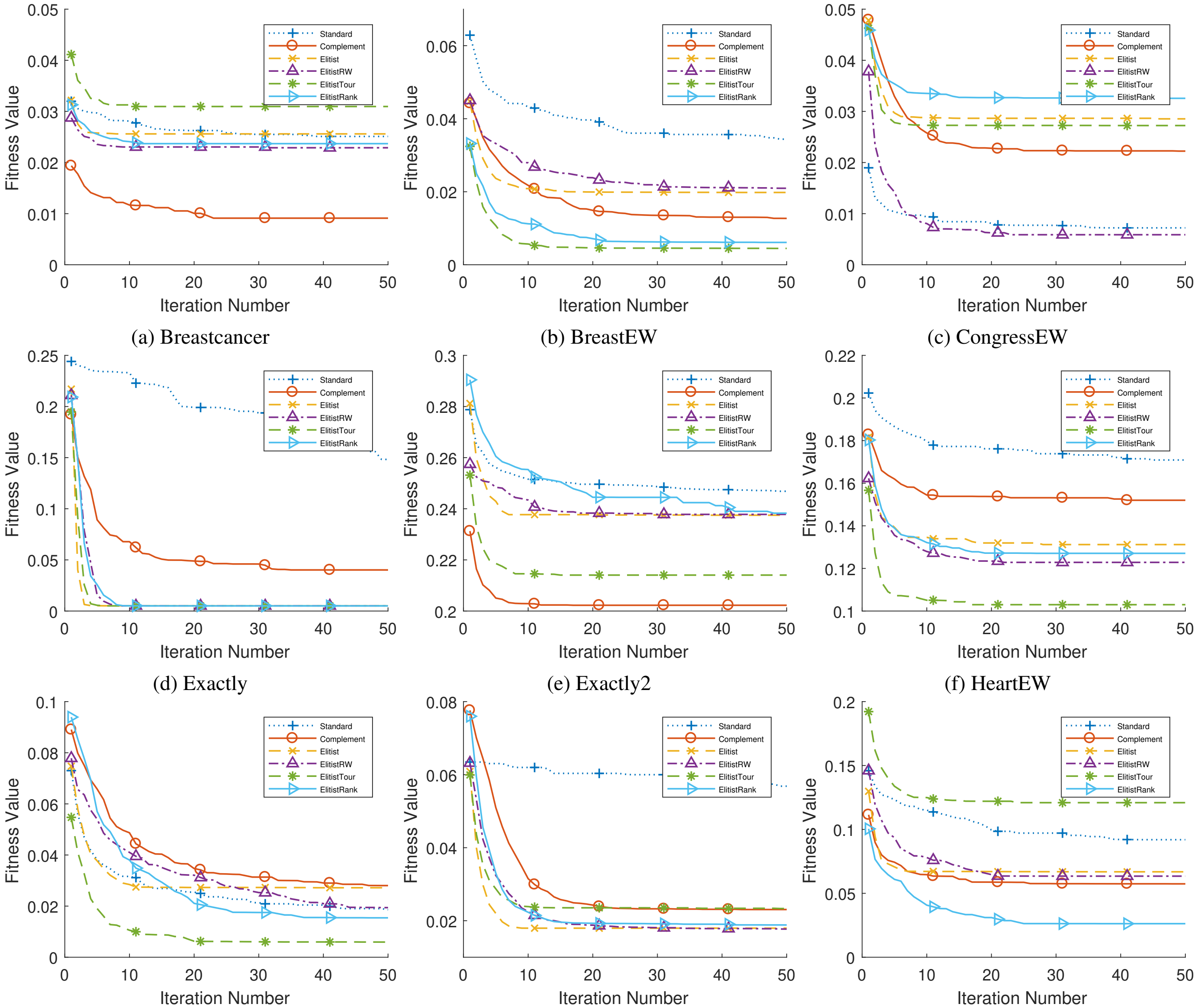
After all, the results and discussed showed that both the TF and binarization approach has a significant influence on the effectiveness of the binary TLBO. Hence, choosing a proper TF along with a fitting binarization scheme has a considerable impact on the exploratory and exploitative potentials of the final wrapper FS technique. One reason for improvements when using V-shaped TFs is that they follow an aggressive exploration tactic. V-shaped TFs allocate high mutation chances for both near and far optimal features, which this characteristic assist in outperforming on datasets with a lower number of features. In contrast, S-shaped TFs have a conservative exploration manner, and they provide high mutation chances only for far optimal features. This trait assists S-shaped TFs in delivering better results for datasets with a higher number of features.

TABLE 7: Comparison between different binarization methods with S-shaped TFs in terms of average running time

Benchmark	Measure	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ET	BTLBO_ER
Breastcancer	AVG	**22.0375**	22.2831	22.1016	22.2050	22.1826	25.4944
	STD	1.1964	1.1896	1.2504	1.2535	1.2375	5.6211
BreastEW	AVG	**22.8783**	23.0654	22.9440	22.9375	22.9969	26.1550
	STD	1.4343	1.3515	1.3704	1.3931	1.3898	4.5314
CongressEW	AVG	**20.1415**	20.2905	20.1971	20.3172	20.2081	22.5537
	STD	1.0633	1.1196	1.0476	1.0735	1.0665	3.3493
Exactly	AVG	**27.7219**	28.8969	28.0023	28.2458	28.0947	33.5767
	STD	1.5629	1.6901	1.6065	1.6412	1.6483	6.6495
Exactly2	AVG	30.1042	30.0318	**29.6147**	29.9449	29.6948	33.2720
	STD	1.7850	1.8091	1.7766	1.7816	1.7346	5.0766
HeartEW	AVG	**17.9813**	18.0761	17.9919	18.1071	18.0433	19.9286
	STD	0.8933	0.9318	0.8894	0.8604	0.8877	3.1225
IonosphereEW	AVG	18.8171	18.8568	18.9330	18.9223	**18.8120**	21.3733
	STD	1.0789	1.0519	1.0361	1.1205	1.0534	3.5469
KrvskpEW	AVG	258.7456	263.1133	**257.1450**	264.9291	259.0755	289.6172
	STD	39.1504	37.4012	31.8124	34.3285	37.3154	55.7511
Lymphography	AVG	17.1649	**16.8811**	17.1625	16.9105	16.9191	18.3764
	STD	0.7577	0.7462	0.7570	0.6859	0.5496	2.9433
M-of-n	AVG	**27.9288**	28.2807	28.0240	29.2921	29.0487	30.8442
	STD	1.4263	1.5863	1.6475	1.2430	1.1060	4.7666
penglungEW	AVG	**19.4497**	20.1905	20.1159	20.4472	21.6040	21.7908
	STD	0.9555	0.9715	1.1169	1.2970	1.6853	3.4438
SonarEW	AVG	17.6170	17.7034	**17.6044**	17.6382	19.7992	19.4067
	STD	0.8472	0.8607	0.8715	0.8749	3.6636	3.8151
SpectEW	AVG	17.9348	17.8566	17.9820	**17.8232**	19.7082	19.6119
	STD	0.8715	0.9238	0.9422	0.9396	2.9852	2.7552
Tic-tac-toe	AVG	**25.0070**	25.2593	25.2052	25.4149	28.4860	28.4177
	STD	1.3707	1.3828	1.4299	1.5358	6.2622	4.6939
Vote	AVG	18.4761	18.5653	18.4832	**18.3884**	20.2961	20.2867
	STD	0.8839	0.9171	0.8781	0.9364	3.0751	3.6949
WaveformEW	AVG	637.0569	669.8580	**636.1258**	661.9894	707.0919	694.0066
	STD	127.9187	131.5720	123.5059	104.9563	136.6539	174.2097
WineEW	AVG	17.1579	17.1697	17.1652	**17.0967**	19.1163	19.0430
	STD	0.8055	0.7264	0.8237	0.7362	3.5912	2.9233
Zoo	AVG	17.2283	17.6129	16.9378	**16.9288**	18.8791	19.3586
	STD	0.7030	0.7397	0.6707	0.7446	3.3728	3.7876
Ranking	Best	8	1	4	4	1	0
Overall Ranking	F-Test	**5.0000**	3.3333	4.6667	3.8333	2.8333	1.3333

TABLE 8: P-values of the Shapiro-Wilk and Kolmogorov-Smirnov normality tests for the classification accuracy results of methods with S-shaped TF (p $\leq$ 0.05 are shown in bold face)

dataset	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ET	BTLBO_ER
Breastcancer	**7.32E-20**	**1.06E-19**	**7.32E-20**	**8.80E-20**	**1.28E-19**	**1.28E-19**
BreastEW	**3.91E-08**	**4.73E-06**	**2.09E-08**	**1.43E-07**	**8.92E-06**	**7.32E-20**
CongressEW	**9.85E-20**	**7.77E-12**	**7.32E-20**	**2.09E-08**	**7.32E-20**	**9.85E-20**
Exactly	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**
Exactly2	**1.05E-05**	**1.55E-05**	**2.80E-09**	**4.63E-05**	**9.09E-05**	**2.74E-06**
HeartEW	**6.09E-08**	**5.45E-05**	**1.73E-09**	**2.09E-08**	**1.02E-07**	**2.26E-05**
IonosphereEW	**2.64E-04**	**3.11E-06**	**3.09E-06**	**1.02E-07**	**3.91E-08**	**1.43E-07**
KrvskpEW	8.31E-01	2.92E-01	4.50E-01	6.68E-02	8.05E-01	4.14E-01
Lymphography	**4.40E-11**	**7.18E-09**	**1.01E-08**	**2.65E-07**	**4.43E-08**	**1.33E-07**
M-of-n	**7.77E-12**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**
penglungEW	**7.32E-20**	**7.32E-20**	**3.19E-09**	**1.96E-10**	**7.32E-20**	**4.43E-09**
SonarEW	**1.25E-07**	**5.31E-06**	**1.43E-07**	**7.77E-12**	**1.78E-10**	**2.09E-08**
SpectEW	**6.38E-06**	**4.43E-09**	**3.11E-06**	**1.02E-07**	**1.82E-07**	**2.22E-06**
Tic-tac-toe	**5.83E-18**	**5.83E-18**	**3.75E-18**	**6.76E-18**	**7.84E-18**	**1.23E-17**
Vote	**5.98E-10**	**1.13E-19**	**4.40E-11**	**7.32E-20**	**2.09E-08**	**1.02E-07**
WaveformEW	**1.59E-02**	1.98E-01	5.11E-02	**4.91E-03**	6.07E-01	4.92E-01
WineEW	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**
Zoo	**7.32E-20**	**2.55E-19**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**



(a) Breastcancer (b) BreastEW (c) CongressEW

(d) Exactly (e) Exactly2 (f) HeartEW

(g) IonosphereEW (h) KrvskpEW (i) Lymphography

FIGURE 5: Convergence curves for BTLBO with different binarization methods for V-shaped TFs on Breastcancer, BreastEW, CongressEW, Exactly, Exactly2, HeartEW, IonosphereEW, KrvskpEW, and Lymphography datasets.

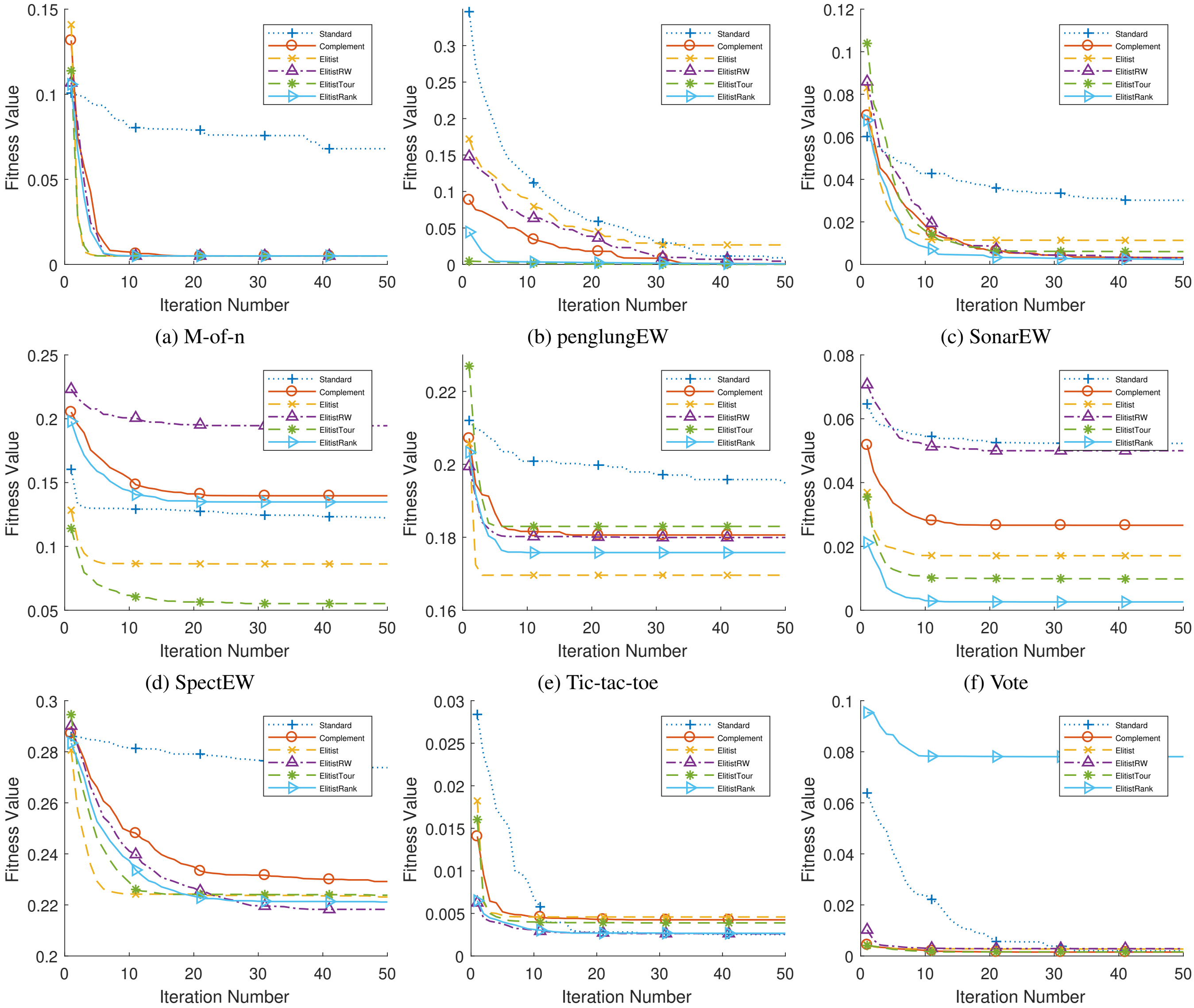
C. COMPARISON OF TOP VARIANTS OF BTLBO

The accuracy, number of features, fitness values, and running time of top variants, BTLBO-S-ET and BTLBO-V-ER are compared in Table 16.

Based on the results of top variants, it can be seen that the BTLBO-V-ER variant shows a better overall performance than BTLBO-S-ET in all metrics. In terms of accuracy rates, BTLBO-V-ER shows a superior efficacy on 55.55% of cases, and it obtains similar results on four problems: WineEW, M-of-n, penglungEW, and Exactly. Considering the number of features, the BTLBO-V-ER outperforms the BTLBO-S-ET on 83.33% of problems and only in three cases, BTLBO-S-ET finds better results. According to fitness and time results, BTLBO-V-ER outperforms the other peer on 77.77% of problems.

The main reason that the BTLBO_ER can carry out a smoother shift from the exploration to exploitation proclivity because of the V-shaped TF that assists the variant in aggressive exploring the feature space and allocating higher mutation chances for both near and far optimal features. It also utilizes a rank-based strategy to choose a solution and adopt the solutions in the next iteration. The advantage of rank-based selection scheme is that it helps the BTLBO variant to prevent rapid and premature convergence. Hence, the results are more enriched during more exploratory trends, and this led to more high-quality features.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3064799, IEEE Access



(a) M-of-n (b) penglungEW (c) SonarEW

(d) SpectEW (e) Tic-tac-toe (f) Vote

(g) WaveformEW (h) WineEW (i) Zoo

FIGURE 6: Convergence curves for BTLBO with different binarization methods for V-shaped TFs on M-of-n, penglungEW, SonarEW, SpectEW, Tic-tac-toe, Vote, WaveformEW, WineEW, and Zoo datasets.

D. COMPARISON OF BTLBO-V-ER WITH OTHER OPTIMIZERS

In this subsection, the performance of the BTLBO-V-ER variant is compared to other well-regarded optimizers from previous works. Numerical comparisons play a crucial role in detecting the overall ranks of developed methods [96, 97, 98, 99]. The performance of the proposed BTLBO-V-ER is compared to the well-established bGWO [91], BGSA [79], BBA [88], and WOA [90] optimizers in terms of average accuracy, the number of features, fitness values are presented in Tables 17-19, respectively. Its worth mentioning that these methods were implemented and executed in the same environment to make a fair comparisons with the proposed approaches.

As per accuracy results, it can be seen that the proposed BTLBO-V-ER has outperformed other peers on 60% of cases. F-test shows that the BTLBO-V-ER is ranked one, followed by bGWO, WOA, BGSA, and BBA techniques. It is seen that when the bGWO is ranked one (Breastcancer, CongressEW, M-of-n, SonarEW, WaveformEW, and Zoo), the results are very competitive and similar. We also observe that BBA cannot show a superior accuracy rate in dealing with any case.

Based on the average number of features in Table 18, the WOA has attained the best rates on 77.77% of cases. Based on F-test results, the BTLBO-V-ER is ranked three, followed by BBA and BGSA.

The p-values of the normality test for accuracy results of BTLBO-V-ER and other methods are reported in Table 20.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3064799, IEEE Access

TABLE 9: P-values of the Wilcoxon test for the classification accuracy results of BTLBO-ET versus other versions for S-shaped TF ($p \leq 0.05$ are shown in bold face, NaN: Not Applicable)

dataset	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ER
Breastcancer	**1.69E-14**	**1.69E-14**	**1.69E-14**	**1.69E-14**	NaN
BreastEW	5.19E-02	**3.47E-07**	**3.94E-05**	**1.47E-07**	**8.14E-12**
CongressEW	**1.69E-14**	**2.71E-14**	NaN	**2.43E-13**	**1.69E-14**
Exactly	NaN	NaN	NaN	NaN	NaN
Exactly2	**1.37E-11**	**1.16E-03**	**4.93E-12**	**3.45E-04**	**1.40E-11**
HeartEW	**8.25E-03**	**7.08E-08**	**1.83E-12**	**1.97E-10**	**1.28E-03**
IonosphereEW	**4.78E-04**	**3.38E-02**	**6.82E-06**	**3.83E-12**	4.26E-01
KrvskpEW	**1.52E-05**	2.64E-01	**4.02E-06**	3.48E-01	**4.79E-02**
Lymphography	4.91E-01	1.92E-01	**2.14E-07**	**4.25E-04**	**1.47E-11**
M-of-n	3.34E-01	NaN	NaN	NaN	NaN
penglungEW	NaN	NaN	**2.57E-13**	**6.13E-14**	**1.09E-02**
SonarEW	**7.21E-08**	**3.41E-10**	**5.89E-05**	3.13E-01	**8.43E-07**
SpectEW	**3.40E-11**	**2.63E-12**	**6.84E-12**	**5.45E-12**	**6.17E-12**
Tic-tac-toe	**1.69E-14**	**1.69E-14**	**1.69E-14**	**1.69E-14**	**1.69E-14**
Vote	**9.50E-13**	**2.70E-03**	**4.04E-02**	**5.36E-09**	**5.61E-08**
WaveformEW	**2.19E-07**	**1.06E-05**	**1.15E-08**	**9.79E-07**	**4.99E-05**
WineEW	NaN	NaN	NaN	NaN	NaN
Zoo	NaN	**1.69E-14**	NaN	NaN	NaN

TABLE 10: Comparison between different binarization methods with V-shaped TFs in terms of average accuracy

Benchmark	Measure	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ET	BTLBO_ER
Breastcancer	AVG	0.9802	**0.9957**	0.9781	0.9843	0.9729	0.9831
	STD	0.0031	0.0067	0.0018	0.0029	0.0066	0.0040
BreastEW	AVG	0.9675	0.9901	0.9828	0.9825	**0.9988**	0.9971
	STD	0.0066	0.0068	0.0054	0.0056	0.0038	0.0048
CongressEW	AVG	0.9950	0.9805	0.9751	**0.9973**	0.9751	0.9705
	STD	0.0058	0.0062	0.0074	0.0049	0.0044	0.0058
Exactly	AVG	0.8652	0.9645	**1.0000**	**1.0000**	**1.0000**	**1.0000**
	STD	0.1284	0.0911	0.0000	0.0000	0.0000	0.0000
Exactly2	AVG	0.7542	**0.7977**	0.7670	0.7637	0.7907	0.7627
	STD	0.0019	0.0064	0.0252	0.0133	0.0083	0.0177
HeartEW	AVG	0.8315	0.8519	0.8716	0.8815	**0.9000**	0.8759
	STD	0.0089	0.0129	0.0118	0.0104	0.0134	0.0099
IonosphereEW	AVG	0.9831	0.9742	0.9751	0.9831	**0.9967**	0.9869
	STD	0.0068	0.0105	0.0103	0.0068	0.0061	0.0082
KrvskpEW	AVG	0.9473	0.9818	**0.9873**	0.9867	0.9819	0.9855
	STD	0.0091	0.0044	0.0063	0.0045	0.0043	0.0027
Lymphography	AVG	0.9121	0.9464	0.9366	0.9398	0.8817	**0.9764**
	STD	0.0221	0.0166	0.0268	0.0184	0.0178	0.0251
M-of-n	AVG	0.9378	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
	STD	0.0573	0.0000	0.0000	0.0000	0.0000	0.0000
penglungEW	AVG	0.9911	**1.0000**	0.9733	0.9978	**1.0000**	**1.0000**
	STD	0.0231	0.0000	0.0414	0.0122	0.0000	0.0000
SonarEW	AVG	0.9714	0.9992	0.9921	**1.0000**	0.9968	**1.0000**
	STD	0.0115	0.0043	0.0144	0.0000	0.0103	0.0000
SpectEW	AVG	0.8784	0.8623	0.9173	0.8062	**0.9475**	0.8673
	STD	0.0105	0.0105	0.0152	0.0094	0.0195	0.0147
Tic-tac-toe	AVG	0.8108	0.8264	**0.8370**	0.8269	0.8227	0.8312
	STD	0.0188	0.0025	0.0140	0.0026	0.0108	0.0054
Vote	AVG	0.9500	0.9750	0.9850	0.9517	0.9939	**0.9994**
	STD	0.0076	0.0085	0.0051	0.0067	0.0082	0.0030
WaveformEW	AVG	0.7285	0.7735	0.7806	**0.7844**	0.7792	0.7820
	STD	0.0062	0.0081	0.0083	0.0050	0.0083	0.0062
WineEW	AVG	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
	STD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Zoo	AVG	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	0.9238
	STD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0237
Ranking	Best	2	6	6	7	9	7
Overall Ranking	F-Test	4.9444	3.5000	3.5000	3.0278	3.1111	**2.9167**

TABLE 11: Comparison between different binarization methods with V-shaped TFs in terms of average number of features

Benchmark	Measure	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ET	BTLBO_ER
Breastcancer	AVG	4.4333	3.9000	**3.1333**	5.8667	3.2667	5.5667
	STD	0.7279	0.6074	0.4342	1.0080	0.6915	0.8172
BreastEW	AVG	**6.5667**	8.2000	7.9000	10.4000	9.6333	9.3667
	STD	2.3589	2.5107	2.3540	2.0943	1.7905	2.8945
CongressEW	AVG	**3.4333**	4.3333	5.8333	4.8333	3.8667	5.0333
	STD	1.4308	1.0283	1.5992	1.8020	1.5253	1.8096
Exactly	AVG	7.3667	**6.0000**	**6.0000**	**6.0000**	**6.0000**	**6.0000**
	STD	2.2203	0.7878	0.0000	0.0000	0.0000	0.0000
Exactly2	AVG	4.1667	**2.3333**	8.1667	4.6000	8.1667	3.9667
	STD	0.9129	3.0324	1.3153	1.7927	1.2058	3.7277
HeartEW	AVG	4.8667	6.4000	4.9000	6.6000	**4.7667**	5.1000
	STD	0.8604	1.0372	1.4468	0.7701	1.9597	0.9229
IonosphereEW	AVG	**6.0333**	8.1333	8.2000	8.6333	8.8333	7.9333
	STD	1.3515	2.4738	2.2190	1.9025	2.1023	2.2273
KrvskpEW	AVG	16.5000	17.5667	19.0000	15.8000	19.0667	**15.5667**
	STD	5.5940	3.7202	1.9298	3.1666	2.1162	4.0911
Lymphography	AVG	6.3333	7.3667	6.9333	6.6333	6.4333	**4.9333**
	STD	1.8815	2.0759	1.1725	1.2172	1.1351	1.5071
M-of-n	AVG	7.7333	**6.0000**	**6.0000**	**6.0000**	**6.0000**	**6.0000**
	STD	1.5071	0.0000	0.0000	0.0000	0.0000	0.0000
penglungEW	AVG	9.7000	26.4667	10.5000	69.8667	**4.3333**	23.4333
	STD	2.8786	8.0590	3.5984	18.0683	0.9942	10.5950
SonarEW	AVG	**10.8000**	14.2667	20.4333	17.9000	16.8000	13.7667
	STD	3.0558	3.1724	3.9539	2.7082	3.8183	2.4731
SpectEW	AVG	**3.5333**	7.1000	9.1333	5.4667	6.8333	7.1333
	STD	0.7761	1.3734	2.4031	1.2521	1.0532	1.7760
Tic-tac-toe	AVG	**6.0000**	7.0000	6.5667	6.8667	**6.0000**	7.0000
	STD	0.7878	0.0000	0.5040	0.3457	0.0000	0.0000
Vote	AVG	4.1667	**2.8000**	3.3333	3.2000	5.6333	3.0667
	STD	1.5105	1.9191	0.7581	1.1861	1.7711	0.2537
WaveformEW	AVG	**13.4667**	19.1333	23.0333	18.8333	20.3333	20.7667
	STD	5.7819	2.7510	4.9374	2.5200	2.3973	3.1259
WineEW	AVG	**3.0333**	5.1000	5.5000	3.1333	4.6667	3.2000
	STD	0.1826	0.3051	0.8200	0.3457	0.9589	0.5509
Zoo	AVG	3.0000	**2.3000**	4.1000	4.3000	2.4000	3.9667
	STD	0.0000	0.4661	0.3051	0.4661	0.8137	1.0662
Ranking	Best	7	4	3	2	5	4
Overall Ranking	F-Test	**4.5278**	3.5833	2.7500	2.9444	3.5556	3.6389

We observe from Table 20 that the p-value is less than 5 % for most of the cases. Therefore, the null hypothesis is not accepted. This fact proves that the utilized results of 30 runs (sample) for the considered dataset are not normally distributed.

Table 21 indicates the p-values of the Wilcoxon test for the accuracy results of BTLBO-V-ER versus other peers. The p-values evidently confirm the meaningful variations of the accuracy results obtained by the BTLBO-V-ER and other competitors in most of the cases.

E. PERFORMANCE OF BTLBO-V-ER WITH DIFFERENT CLASSIFIERS

In this subsection, the performance of the BTLBO-V-ER variant with the KNN classifier is compared to Linear Discriminant Analysis (LDA), Decision Tree (DT), and Adaptive Boosting (AdaBoost) classifiers in terms of average accuracy, and time. Table 22 shows the performance results of BTLBO-V-ER with four different classifiers. Based on the results, it can be seen that the BTLBO-V-ER with KNN shows a good performance compared to BTLBO-V-ER with LDA, DT, and AdaBoost in terms of average accuracy, and time. In terms of accuracy rates, BTLBO-V-ER with KNN shows better performance on five datasets, and it obtains similar results on four datasets. According to time results, BTLBO-V-ER with

TABLE 12: Comparison between different binarization methods with V-shaped TFs in terms of average fitness

Benchmark	Measure	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ET	BTLBO_ER
Breastcancer	AVG	0.0251	**0.0091**	0.0256	0.0229	0.0310	0.0237
	STD	0.0028	0.0064	0.0018	0.0026	0.0067	0.0030
BreastEW	AVG	0.0344	0.0127	0.0198	0.0210	**0.0045**	0.0061
	STD	0.0061	0.0068	0.0052	0.0050	0.0039	0.0043
CongressEW	AVG	0.0072	0.0222	0.0285	**0.0059**	0.0272	0.0326
	STD	0.0049	0.0058	0.0067	0.0039	0.0035	0.0049
Exactly	AVG	0.1396	0.0401	**0.0050**	**0.0050**	**0.0050**	**0.0050**
	STD	0.1268	0.0897	0.0000	0.0000	0.0000	0.0000
Exactly2	AVG	0.2468	**0.2023**	0.2375	0.2378	0.2140	0.2383
	STD	0.0020	0.0040	0.0244	0.0121	0.0079	0.0199
HeartEW	AVG	0.1709	0.1520	0.1312	0.1228	**0.1030**	0.1271
	STD	0.0083	0.0126	0.0114	0.0099	0.0136	0.0098
IonosphereEW	AVG	0.0186	0.0280	0.0271	0.0193	**0.0059**	0.0154
	STD	0.0065	0.0100	0.0100	0.0067	0.0057	0.0081
KrvskpEW	AVG	0.0568	0.0231	0.0180	**0.0177**	0.0233	0.0188
	STD	0.0091	0.0037	0.0062	0.0040	0.0040	0.0019
Lymphography	AVG	0.0908	0.0574	0.0669	0.0635	0.1209	**0.0263**
	STD	0.0218	0.0160	0.0264	0.0180	0.0175	0.0252
M-of-n	AVG	0.0680	**0.0050**	**0.0050**	**0.0050**	**0.0050**	**0.0050**
	STD	0.0566	0.0000	0.0000	0.0000	0.0000	0.0000
penglungEW	AVG	0.0091	0.0008	0.0267	0.0044	**0.0001**	0.0007
	STD	0.0228	0.0002	0.0409	0.0122	0.0000	0.0003
SonarEW	AVG	0.0301	0.0032	0.0113	0.0030	0.0060	**0.0023**
	STD	0.0112	0.0043	0.0141	0.0005	0.0100	0.0004
SpectEW	AVG	0.1221	0.1397	0.0862	0.1945	**0.0552**	0.1348
	STD	0.0101	0.0102	0.0145	0.0089	0.0192	0.0138
Tic-tac-toe	AVG	0.1948	0.1806	**0.1696**	0.1799	0.1830	0.1758
	STD	0.0189	0.0025	0.0133	0.0023	0.0107	0.0053
Vote	AVG	0.0523	0.0266	0.0171	0.0500	0.0098	**0.0026**
	STD	0.0066	0.0072	0.0047	0.0062	0.0072	0.0031
WaveformEW	AVG	0.2722	0.2292	0.2231	**0.2182**	0.2238	0.2211
	STD	0.0068	0.0082	0.0089	0.0049	0.0084	0.0061
WineEW	AVG	**0.0025**	0.0043	0.0046	0.0026	0.0039	0.0027
	STD	0.0002	0.0003	0.0007	0.0003	0.0008	0.0005
Zoo	AVG	0.0020	**0.0015**	0.0027	0.0029	0.0016	0.0781
	STD	0.0000	0.0003	0.0002	0.0003	0.0005	0.0228
Ranking	Best	1	4	3	5	7	5
Overall Ranking	F-Test	2.2222	3.5000	3.2500	3.9167	3.9722	**4.1389**

KNN outperforms the other classifiers on 16 datasets.

F. COMPARISON WITH RESULTS OF LITERATURE

This subsection compares the results in term of classification rates with those obtained by previous well-established methods on a number of datasets. For this purpose, we compared the results of BTLBO-V-ER with BSSA_S3_CP proposed by Faris et al. [100], WOA-CM proposed by Mafarja and Mirjalili [90], BGOA_EPD_Tour proposed by afarja et al. [88], GA-based method proposed by Kashef and Nezamabadi-pour [101], PSO-based technique proposed by Kashef and Nezamabadi-pour [101], another GA-based method by Emary et al. [91], another method based on PSO Emary et al. [91], bGWO1 proposed by [91], bGWO2 developed by Emary et al. [91], HGSA designed by Taradeh et al. [102], BGOA-M method introduced by Mafarja et al. [103], BDA-TVv4 developed by Mafarja et al. [104], BGWOPSO technique developed by Al-Tashi et al. [105], and S-bBOA proposed by Arora and Anand [59]. Here, we focus on the final reported accuracy value of compared methods regardless of the same computing conditions and settings. We suppose that the reported rates in referred works represent the overall average accuracy of that method on the used datasets independent of settings and parameters.

From results of the BTLBO-V-ER in Table 23, it is observed that the developed method realizes the best re-

TABLE 13: Comparison between different binarization methods with V-shaped TFs in terms of average running time

Benchmark	Measure	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ET	BTLBO_ER
Breastcancer	AVG	**18.7310**	21.4529	20.6834	22.5252	20.8662	22.3319
	STD	2.6713	1.4272	1.3367	1.4168	1.2401	1.3481
BreastEW	AVG	**22.3971**	22.8115	22.4723	22.9240	23.1882	22.6132
	STD	3.1074	1.3732	1.5649	1.2749	1.5278	1.3679
CongressEW	AVG	**16.7382**	19.2396	19.5834	19.1872	18.4093	19.3533
	STD	2.4953	1.2630	1.3018	1.3505	1.4037	1.1233
Exactly	AVG	**23.4063**	26.2859	28.9297	26.9073	25.0171	29.8378
	STD	3.0762	1.7073	1.7585	1.5663	1.4067	1.7433
Exactly2	AVG	**22.1700**	22.8623	32.5834	26.0682	32.1811	28.9860
	STD	2.2559	2.8482	3.5536	2.7440	3.4014	5.1626
HeartEW	AVG	**16.3392**	17.8349	17.5552	18.0817	17.5022	17.6278
	STD	2.2122	0.8680	0.8570	0.8753	0.8552	0.7982
IonosphereEW	AVG	19.5924	19.1334	19.1999	19.0209	**18.9389**	19.1461
	STD	2.4869	1.1428	1.1031	1.0887	1.0485	1.0878
KrvskpEW	AVG	**160.4035**	248.3669	255.9400	239.2102	258.8687	233.9836
	STD	18.5969	38.1628	35.2922	33.1901	33.0553	27.3797
Lymphography	AVG	**16.1822**	16.5149	16.8944	16.5189	16.8772	16.4447
	STD	0.6741	0.7074	0.6957	0.7247	0.6655	0.7025
M-of-n	AVG	**22.4075**	27.1069	26.7875	26.8252	26.7828	27.1041
	STD	1.3899	1.4851	1.4718	1.4697	1.5501	1.5492
penglungEW	AVG	19.2187	20.0338	19.7522	20.3304	**19.0131**	19.5895
	STD	0.9490	0.9390	0.9550	1.0137	0.8975	0.8473
SonarEW	AVG	17.4315	17.2765	17.2845	17.4570	**17.2356**	17.4013
	STD	0.8252	0.8722	0.8025	0.9006	0.8422	0.8916
SpectEW	AVG	**13.0434**	17.9478	17.9829	17.9013	17.9296	18.1025
	STD	1.1524	0.9662	1.0834	0.9236	0.8947	0.8369
Tic-tac-toe	AVG	**19.2887**	28.6100	28.5727	28.6613	26.6711	29.4711
	STD	1.2303	1.9234	2.4464	1.8090	1.5450	1.8893
Vote	AVG	**13.3469**	17.0176	17.4495	17.4420	17.9854	17.5580
	STD	2.3758	1.2335	0.9272	0.8680	0.9895	0.8837
WaveformEW	AVG	**278.5031**	608.0048	675.6856	617.8048	619.8525	620.2033
	STD	38.9078	84.9576	159.5340	97.0335	108.0055	133.2125
WineEW	AVG	**14.2497**	16.9678	16.9209	16.6653	16.7992	16.5925
	STD	0.8466	0.6990	0.8001	0.7744	0.6856	0.7325
Zoo	AVG	**16.3994**	16.8177	17.0423	17.1344	16.6707	17.3569
	STD	0.7056	0.6195	0.7051	0.7856	0.7761	0.8070
Ranking	Best	15	0	0	0	3	0
Overall Ranking	F-Test	**5.4444**	3.3333	2.7222	2.9444	3.8333	2.7222

TABLE 14: P-values of the Shapiro-Wilk and Kolmogorov-Smirnov normality test for the classification accuracy results of V-shaped TF approaches ($p \leq 0.05$ are shown in bold face)

dataset	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ET	BTLBO_ER
Breastcancer	**1.01E-08**	**3.91E-08**	**4.40E-11**	**4.43E-09**	**1.71E-06**	**3.00E-07**
BreastEW	**2.46E-04**	**1.16E-04**	**1.23E-06**	**3.91E-05**	**1.93E-10**	**1.42E-07**
CongressEW	**1.82E-07**	**1.58E-06**	**3.60E-05**	**1.01E-08**	**1.73E-09**	**1.82E-07**
Exactly	**1.55E-04**	**7.46E-10**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**
Exactly2	**1.73E-09**	**2.15E-09**	**1.30E-04**	**3.83E-02**	**2.35E-03**	**3.96E-06**
HeartEW	**3.32E-07**	**8.52E-05**	**3.00E-06**	**5.74E-07**	**9.25E-08**	**1.58E-06**
IonosphereEW	**2.82E-07**	**6.42E-05**	**4.05E-04**	**2.82E-07**	**1.01E-08**	**9.25E-06**
KrvskpEW	9.43E-01	5.81E-02	**7.88E-05**	**2.35E-04**	**2.75E-05**	**4.68E-06**
Lymphography	**6.83E-05**	**3.05E-07**	**6.33E-05**	**5.21E-06**	**2.45E-05**	**2.72E-05**
M-of-n	**1.92E-04**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**
penglungEW	**5.98E-10**	**7.32E-20**	**3.84E-07**	**7.77E-12**	**7.32E-20**	**7.32E-20**
SonarEW	**2.82E-07**	**7.77E-12**	**6.39E-08**	**7.32E-20**	**1.93E-10**	**7.32E-20**
SpectEW	**3.09E-06**	**9.94E-07**	**8.16E-04**	**2.11E-07**	**3.29E-05**	**9.16E-04**
Tic-tac-toe	**4.91E-04**	**6.64E-08**	**8.46E-07**	**8.37E-09**	**6.42E-09**	**5.98E-10**
Vote	**1.45E-07**	**2.21E-07**	**1.78E-10**	**1.66E-08**	**1.02E-07**	**7.77E-12**
WaveformEW	3.29E-01	4.55E-01	7.78E-01	6.76E-01	8.54E-01	5.04E-01
WineEW	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**
Zoo	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**7.32E-20**	**1.43E-07**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3064799, IEEE Access

TABLE 15: P-values of the Wilcoxon test for the classification accuracy results of BTLBO-ER versus other versions for V-shaped Transfer Function ($p \leq 0.05$ are shown in bold face, NaN: Not Applicable)

dataset	BTLBO_S	BTLBO_C	BTLBO_E	BTLBO_ERW	BTLBO_ET
Breastcancer	**1.88E-03**	**6.81E-09**	**3.38E-07**	2.29E-01	**1.63E-08**
BreastEW	**6.11E-12**	**2.73E-05**	**9.35E-11**	**2.08E-10**	6.59E-02
CongressEW	**6.21E-12**	**6.33E-07**	**1.57E-02**	**3.20E-12**	**1.48E-03**
Exactly	**1.70E-08**	**1.10E-02**	NaN	NaN	NaN
Exactly2	**3.53E-03**	**2.50E-12**	2.24E-01	5.34E-01	**1.81E-10**
HeartEW	**3.33E-12**	**8.05E-09**	1.63E-01	**2.90E-02**	**1.17E-08**
IonosphereEW	6.52E-02	**7.32E-06**	**2.18E-05**	6.52E-02	**8.60E-06**
KrvskpEW	**1.74E-11**	**8.95E-04**	**4.60E-03**	**1.51E-02**	**1.53E-04**
Lymphography	**9.51E-11**	**7.63E-05**	**3.46E-06**	**6.65E-06**	**1.56E-11**
M-of-n	**1.30E-07**	NaN	NaN	NaN	NaN
penglungEW	**4.18E-02**	NaN	**6.39E-04**	3.34E-01	NaN
SonarEW	**9.94E-13**	3.34E-01	**2.75E-03**	NaN	8.15E-02
SpectEW	**2.51E-03**	1.75E-01	**2.43E-11**	**9.57E-12**	**1.31E-11**
Tic-tac-toe	**2.46E-09**	**1.97E-07**	1.05E-01	**1.13E-07**	**8.87E-08**
Vote	**2.39E-13**	**1.48E-12**	**2.68E-11**	**1.80E-13**	**1.42E-03**
WaveformEW	**2.92E-11**	**5.65E-05**	5.44E-01	1.40E-01	1.78E-01
WineEW	NaN	NaN	NaN	NaN	NaN
Zoo	**4.17E-13**	**4.17E-13**	**4.17E-13**	**4.17E-13**	**4.17E-13**

sults on nine datasets including Breastcancer, BreastEW, IonosphereEW, KrvskpEW, Lymphography, penglungEW, SonarEW, Tic-tac-toe, and Vote cases. There is a tie for three datasets. For WineEW case, which has 13 features and 178 instances, the proposed BTLBO-V-ER has the extreme accuracy rate of 100% similar to the obtained rate of BGWOPSO. For penglungEW that is a moderately larger scale dataset with 325 features, BTLBO-V-ER archives the ideal average accuracy of 100%. This observation indicates the boosted exploratory and exploitative capabilities of the proposed TLBO-based method and its more steady performance in harmonizing the exploration and exploitation drifts. It is seen that the accuracy of GA, PSO, bGWO1, and bGWO2 in [91] are not remarkable for this case, and the rates are located between the interval of [58, 60]. We observe that methods such as GA [101], PSO [101], GA [91], PSO [91], bGWO1 [91], bGWO2 [91], S-bBOA [59] have not achieved the relatively best rates in dealing with any of datasets. As per overall ranking rates (F-test), we observe that the BTLBO-V-ER attains the best place, followed by BGWOPSO, HGSA, BDA-TVv4, BGOA-M, BGOA_EPD_Tour, BSSA_S3_CP, S-bBOA, WOA-CM, bGWO2, PSO [101], bGWO1, PSO [91], GA [101], and GA [91].

These results also show that the designed modifications, V-shaped TF, and used rank-based selection structure have assisted this method in achieving high-quality solutions compared to the reported results in recent literature.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3064799, IEEE Access

TABLE 16: Comparison between the BTLBO-S-ET and BTLBO-V-ER based on accuracy, number of features, fitness, and running time

Benchmark	Mesure	Accuracy		Number of Features		Fitness		Time	
		BTLBO-S-ET	BTLBO-V-ER	BTLBO-S-ET	BTLBO-V-ER	BTLBO-S-ET	BTLBO-V-ER	BTLBO-S-ET	BTLBO-V-ER
Breastcancer	AVG	0.9786	**0.9831**	**4.0000**	5.5667	0.0262	**0.0237**	**22.1826**	22.3319
	STD	0.0000	0.0040	0.0000	0.8172	0.0000	0.0030	1.2375	1.3481
BreastEW	AVG	0.9877	**0.9971**	11.9667	**9.3667**	0.0163	**0.0061**	22.9969	**22.6132**
	STD	0.0055	0.0048	2.2512	2.8945	0.0050	0.0043	1.3898	1.3679
CongressEW	AVG	**1.0000**	0.9705	6.3667	**5.0333**	**0.0042**	0.0326	20.2081	**19.3533**
	STD	0.0000	0.0058	0.7184	1.8096	0.0005	0.0049	1.0665	1.1233
Exactly	AVG	**1.0000**	**1.0000**	6.3333	**6.0000**	0.0053	**0.0050**	**28.0947**	29.8378
	STD	0.0000	0.0000	0.4795	0.0000	0.0004	0.0000	1.6483	1.7433
Exactly2	AVG	**0.7995**	0.7627	9.5000	**3.9667**	**0.2064**	0.2383	29.6948	**28.9860**
	STD	0.0115	0.0177	0.5724	3.7277	0.0117	0.0199	1.7346	5.1626
HeartEW	AVG	**0.8957**	0.8759	6.0333	**5.1000**	**0.1083**	0.1271	18.0433	**17.6278**
	STD	0.0091	0.0099	1.2726	0.9229	0.0084	0.0098	0.8877	0.7982
IonosphereEW	AVG	0.9761	**0.9869**	12.6667	**7.9333**	0.0275	**0.0154**	**18.8120**	19.1461
	STD	0.0066	0.0082	2.5641	2.2273	0.0065	0.0081	1.0534	1.0878
KrvskpEW	AVG	0.9768	**0.9855**	18.8000	**15.5667**	0.0283	**0.0188**	259.0755	**233.9836**
	STD	0.0037	0.0027	2.5784	4.0911	0.0034	0.0019	37.3154	27.3797
Lymphography	AVG	0.9344	**0.9764**	8.4667	**4.9333**	0.0700	**0.0263**	16.9191	**16.4447**
	STD	0.0138	0.0251	1.2521	1.5071	0.0135	0.0252	0.5496	0.7025
M-of-n	AVG	**1.0000**	**1.0000**	6.3000	**6.0000**	0.0052	**0.0050**	29.0487	**27.1041**
	STD	0.0000	0.0000	0.4661	0.0000	0.0004	0.0000	1.1060	1.5492
penglungEW	AVG	**1.0000**	**1.0000**	126.1667	**23.4333**	0.0039	**0.0007**	21.6040	**19.5895**
	STD	0.0000	0.0000	4.5719	10.5950	0.0001	0.0003	1.6853	0.8473
SonarEW	AVG	0.9976	**1.0000**	28.3000	**13.7667**	0.0072	**0.0023**	19.7992	**17.4013**
	STD	0.0073	0.0000	4.1285	2.4731	0.0069	0.0004	3.6636	0.8916
SpectEW	AVG	0.8599	**0.8673**	8.2333	**7.1333**	0.1426	**0.1348**	19.7082	**18.1025**
	STD	0.0093	0.0147	1.9241	1.7760	0.0087	0.0138	2.9852	0.8369
Tic-tac-toe	AVG	0.8281	**0.8312**	**5.0000**	7.0000	0.1764	**0.1758**	**28.4860**	29.4711
	STD	0.0000	0.0054	0.0000	0.0000	0.0000	0.0053	6.2622	1.8893
Vote	AVG	0.9878	**0.9994**	5.1667	**3.0667**	0.0155	**0.0026**	20.2961	**17.5580**
	STD	0.0075	0.0030	1.3153	0.2537	0.0070	0.0031	3.0751	0.8837
WaveformEW	AVG	0.7609	**0.7820**	20.9333	**20.7667**	0.2421	**0.2211**	707.0919	**620.2033**
	STD	0.0065	0.0062	2.9353	3.1259	0.0065	0.0061	136.6539	133.2125
WineEW	AVG	**1.0000**	**1.0000**	4.3333	**3.2000**	0.0036	**0.0027**	19.1163	**16.5925**
	STD	0.0000	0.0000	0.5467	0.5509	0.0005	0.0005	3.5912	0.7325
Zoo	AVG	**1.0000**	0.9238	**3.5000**	3.9667	**0.0023**	0.0781	18.8791	**17.3569**
	STD	0.0000	0.0237	0.5085	1.0662	0.0003	0.0228	3.3728	0.8070
Ranking	W|T|L	4|4|10	**10|4|8**	3|0|15	**15|0|3**	4|0|14	**14|0|4**	4|0|14	**14|0|4**

TABLE 17: Comparison between BTLBO-V-ER and other methods in terms of average accuracy

Benchmark	Measure	BTLBO-V-ER	bGWO	BGSA	BBA	WOA
Breastcancer	AVG	0.9831	**0.9848**	0.9729	0.9650	0.9691
	STD	0.0040	0.0072	0.0078	0.0236	0.0034
BreastEW	AVG	**0.9971**	0.9781	0.9564	0.9108	0.9722
	STD	0.0048	0.0055	0.0085	0.0226	0.0089
CongressEW	AVG	0.9705	**0.9939**	0.9598	0.8487	0.9709
	STD	0.0058	0.0058	0.0058	0.1051	0.0058
Exactly	AVG	**1.0000**	0.9908	0.7930	0.6783	0.9298
	STD	0.0000	0.0502	0.1071	0.0989	0.1299
Exactly2	AVG	0.7627	0.7222	0.7157	0.6168	**0.7672**
	STD	0.0177	0.0130	0.0155	0.0649	0.0104
HeartEW	AVG	**0.8759**	0.8586	0.7932	0.7154	0.8679
	STD	0.0099	0.0150	0.0262	0.0704	0.0227
IonosphereEW	AVG	**0.9869**	0.9822	0.9174	0.8812	0.9737
	STD	0.0082	0.0082	0.0109	0.0350	0.0121
KrvskpEW	AVG	**0.9855**	0.9798	0.9402	0.8264	0.9546
	STD	0.0027	0.0068	0.0171	0.1153	0.0114
Lymphography	AVG	**0.9764**	0.9676	0.8838	0.8072	0.9388
	STD	0.0251	0.0140	0.0283	0.0915	0.0249
M-of-n	AVG	**1.0000**	**1.0000**	0.8947	0.7888	0.9650
	STD	0.0000	0.0000	0.0604	0.0953	0.0596
penglungEW	AVG	**1.0000**	0.9822	0.9311	0.8889	0.9689
	STD	0.0000	0.0300	0.0130	0.0404	0.0381
SonarEW	AVG	**1.0000**	**1.0000**	0.9436	0.8476	0.9222
	STD	0.0000	0.0000	0.0171	0.0479	0.0216
SpectEW	AVG	0.8673	0.8735	0.7932	0.7549	**0.8827**
	STD	0.0147	0.0169	0.0182	0.0603	0.0122
Tic-tac-toe	AVG	**0.8312**	0.8259	0.7816	0.7128	0.7944
	STD	0.0054	0.0093	0.0210	0.0870	0.0243
Vote	AVG	**0.9994**	0.9867	0.9589	0.9350	0.9983
	STD	0.0030	0.0134	0.0114	0.0411	0.0051
WaveformEW	AVG	0.7820	**0.7832**	0.7241	0.6801	0.7343
	STD	0.0062	0.0098	0.0116	0.0370	0.0114
WineEW	AVG	**1.0000**	0.9880	0.9843	0.8861	**1.0000**
	STD	0.0000	0.0140	0.0140	0.0807	0.0000
Zoo	AVG	0.9238	**1.0000**	**1.0000**	0.9037	**1.0000**
	STD	0.0237	0.0000	0.0000	0.1173	0.0000
Ranking	Best	**12**	6	1	0	4
Overall Ranking	F-Test	**1.6389**	2.0000	3.7778	5.0000	2.5833

TABLE 18: Comparison between BTLBO-V-ER and other meta-heuristics in terms of average number of features

Benchmark	Measure	BTLBO-V-ER	bGWO	BGSA	BBA	WOA
Breastcancer	AVG	5.5667	4.2667	5.6000	4.3000	**4.2000**
	STD	0.8172	0.7397	1.0372	1.2905	0.5509
BreastEW	AVG	9.3667	8.3333	14.1000	12.9000	**7.4000**
	STD	2.8945	2.2489	2.4403	2.3096	1.3287
CongressEW	AVG	5.0333	3.7000	6.3667	5.7000	**2.2333**
	STD	1.8096	0.7022	1.4259	1.5570	1.5241
Exactly	AVG	6.0000	5.9000	8.1000	6.9333	**5.5000**
	STD	0.0000	0.5477	1.7879	1.8742	1.4081
Exactly2	AVG	3.9667	8.2667	4.4667	5.8667	**3.4667**
	STD	3.7277	1.2015	2.7510	2.0800	0.5713
HeartEW	AVG	5.1000	5.5333	6.0000	5.7000	**4.7667**
	STD	0.9229	1.9070	1.7420	1.6640	0.8172
IonosphereEW	AVG	7.9333	7.2000	13.7333	12.4667	**4.2333**
	STD	2.2273	1.2429	2.7156	2.6618	0.8976
KrvskpEW	AVG	15.5667	14.2667	20.5667	15.9000	**10.0000**
	STD	4.0911	1.4606	2.9674	3.1552	3.4039
Lymphography	AVG	4.9333	5.7667	9.1333	8.9000	**4.7000**
	STD	1.5071	1.6121	2.1129	1.6887	1.3684
M-of-n	AVG	**6.0000**	**6.0000**	8.2000	6.7333	6.0667
	STD	0.0000	0.0000	1.3995	1.8925	0.5833
penglungEW	AVG	23.4333	10.1667	150.3333	127.0667	**7.2667**
	STD	10.5950	2.1669	9.0567	17.2705	1.4606
SonarEW	AVG	13.7667	10.6333	28.8667	25.1667	**10.5000**
	STD	2.4731	1.6291	4.5541	4.0691	3.3296
SpectEW	AVG	7.1333	7.0333	9.9667	8.9333	**4.3667**
	STD	1.7760	1.4499	2.3116	2.6773	1.5643
Tic-tac-toe	AVG	7.0000	6.4667	5.8333	**4.0667**	5.4000
	STD	0.0000	0.7303	0.5921	1.3374	0.4983
Vote	AVG	3.0667	4.8667	5.9667	6.8667	**2.9000**
	STD	0.2537	1.1059	1.7711	1.6344	0.7120
WaveformEW	AVG	20.7667	15.9333	22.0667	18.0333	**8.8000**
	STD	3.1259	2.1961	3.0050	3.1126	1.6692
WineEW	AVG	**3.2000**	5.6000	6.2333	5.1667	3.4333
	STD	0.5509	1.5888	1.3817	1.5332	0.5683
Zoo	AVG	3.9667	**2.7000**	7.1667	5.9333	5.4000
	STD	1.0662	0.5350	1.6626	1.7604	0.5632
Ranking	Best	2	2	0	1	**14**
Overall Ranking	F-Test	3.2500	3.5278	1.2778	2.2778	**4.6667**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3064799, IEEE Access

TABLE 19: Comparison between BTLBO-V-ER and other meta-heuristics in terms of average fitness

Benchmark	Measure	BTLBO-V-ER	bGWO	BGSA	BBA	WOA
Breastcancer	AVG	0.0237	0.0204	0.0339	**0.0166**	0.0359
	STD	0.0030	0.0063	0.0073	0.0042	0.0033
BreastEW	AVG	**0.0061**	0.0246	0.0480	0.0528	0.0301
	STD	0.0043	0.0054	0.0083	0.0104	0.0088
CongressEW	AVG	0.0326	**0.0085**	0.0441	0.0525	0.0303
	STD	0.0049	0.0055	0.0056	0.0084	0.0049
Exactly	AVG	**0.0050**	0.0140	0.2117	0.2225	0.0740
	STD	0.0000	0.0492	0.1056	0.1247	0.1278
Exactly2	AVG	0.2383	0.2819	0.2852	0.2993	**0.2334**
	STD	0.0199	0.0123	0.0167	0.0116	0.0103
HeartEW	AVG	**0.1271**	0.1446	0.2097	0.1963	0.1348
	STD	0.0098	0.0144	0.0255	0.0158	0.0221
IonosphereEW	AVG	**0.0154**	0.0198	0.0860	0.0774	0.0273
	STD	0.0081	0.0080	0.0109	0.0125	0.0120
KrvskpEW	AVG	**0.0188**	0.0241	0.0651	0.0636	0.0478
	STD	0.0019	0.0067	0.0165	0.0136	0.0107
Lymphography	AVG	**0.0263**	0.0355	0.1204	0.0906	0.0633
	STD	0.0252	0.0135	0.0276	0.0218	0.0246
M-of-n	AVG	**0.0050**	**0.0050**	0.1111	0.1038	0.0397
	STD	0.0000	0.0000	0.0593	0.0549	0.0590
penglungEW	AVG	**0.0007**	0.0179	0.0728	0.0696	0.0310
	STD	0.0003	0.0297	0.0129	0.0006	0.0377
SonarEW	AVG	0.0023	**0.0018**	0.0607	0.0779	0.0788
	STD	0.0004	0.0003	0.0167	0.0197	0.0210
SpectEW	AVG	0.1348	0.1286	0.2095	0.1893	**0.1182**
	STD	0.0138	0.0163	0.0180	0.0206	0.0117
Tic-tac-toe	AVG	**0.1758**	0.1805	0.2235	0.1846	0.2102
	STD	0.0053	0.0084	0.0206	0.0191	0.0236
Vote	AVG	**0.0026**	0.0164	0.0447	0.0309	0.0036
	STD	0.0031	0.0126	0.0111	0.0085	0.0046
WaveformEW	AVG	0.2211	**0.2187**	0.2788	0.2865	0.2653
	STD	0.0061	0.0095	0.0114	0.0146	0.0112
WineEW	AVG	**0.0027**	0.0166	0.0208	0.0307	0.0029
	STD	0.0005	0.0129	0.0133	0.0097	0.0005
Zoo	AVG	0.0781	**0.0018**	0.0048	0.0035	0.0036
	STD	0.0228	0.0004	0.0011	0.0010	0.0004
Ranking	Best	**11**	5	0	1	2
Overall Ranking	F-Test	**4.2500**	4.0278	1.5556	2.0000	3.1667

TABLE 20: P-values of the Shapiro-Wilk and Kolmogorov-Smirnov normality tests for the classification accuracy results obtained by BTLBO-V-ER and other meta-heuristics ($p \leq 0.05$ are bolded)

dataset	bGWO	BGSA	BBA	WOA	BTLBO-V-ER
Breastcancer	**1.82E-07**	**5.98E-04**	**7.56E-03**	**6.64E-08**	**3.00E-07**
BreastEW	**5.04E-05**	**4.24E-03**	5.27E-01	**1.91E-02**	**1.42E-07**
CongressEW	**2.11E-07**	**2.21E-07**	**3.04E-04**	**2.11E-07**	**1.82E-07**
Exactly	**7.77E-12**	**2.36E-04**	**1.81E-06**	**1.83E-08**	**7.32E-20**
Exactly2	**7.25E-04**	**8.58E-04**	**7.72E-03**	**2.13E-02**	**3.96E-06**
HeartEW	**3.04E-06**	**9.77E-03**	1.93E-01	**9.96E-03**	**1.58E-06**
IonosphereEW	**1.80E-06**	**7.33E-04**	1.29E-01	**4.20E-03**	**9.25E-06**
KrvskpEW	**1.47E-03**	9.46E-01	**2.20E-03**	2.15E-01	**4.68E-06**
Lymphography	**5.78E-08**	**6.02E-03**	1.20E-01	**5.91E-05**	**2.72E-05**
M-of-n	**7.32E-20**	**2.09E-03**	5.77E-01	**6.09E-08**	**7.32E-20**
penglungEW	**2.09E-08**	**1.06E-11**	**1.05E-05**	**1.55E-06**	**7.32E-20**
SonarEW	**7.32E-20**	**3.51E-04**	**2.35E-02**	**2.14E-03**	**7.32E-20**
SpectEW	**8.17E-05**	**8.53E-03**	1.89E-01	**9.04E-05**	**9.16E-04**
Tic-tac-toe	**5.26E-06**	**1.66E-02**	1.46E-01	**1.37E-06**	**5.98E-10**
Vote	**1.11E-04**	**1.50E-04**	**8.87E-03**	**1.78E-10**	**7.77E-12**
WaveformEW	9.70E-01	4.51E-01	**4.38E-03**	**4.09E-02**	5.04E-01
WineEW	**1.82E-07**	**1.82E-07**	**3.52E-03**	**7.32E-20**	**7.32E-20**
Zoo	**7.32E-20**	**7.32E-20**	**1.55E-05**	**7.32E-20**	**1.43E-07**

TABLE 21: P-values of the Wilcoxon test for the classification accuracy results obtained by BTLBO-V-ER versus other meta-heuristics ($p \leq 0.05$ are bolded), NaN: Not applicape

dataset	bGWO	BGSA	BBA	WOA
Breastcancer	6.01E-01	**1.14E-06**	**2.84E-04**	**1.03E-11**
BreastEW	**1.78E-11**	**7.85E-12**	**9.54E-12**	**1.89E-11**
CongressEW	**6.41E-12**	**1.53E-07**	**1.35E-11**	8.04E-01
Exactly	3.34E-01	**4.52E-12**	**1.64E-11**	**5.58E-03**
Exactly2	**1.17E-09**	**4.79E-10**	**1.59E-11**	5.20E-01
HeartEW	**7.16E-06**	**1.14E-11**	**1.26E-11**	2.48E-01
IonosphereEW	**3.91E-02**	**7.71E-12**	**1.16E-11**	**1.83E-05**
KrvskpEW	**1.25E-04**	**1.76E-11**	**1.76E-11**	**1.74E-11**
Lymphography	4.35E-01	**2.79E-11**	**3.38E-11**	**8.68E-06**
M-of-n	NaN	**1.20E-12**	**1.20E-12**	**2.79E-03**
penglungEW	**2.70E-03**	**1.77E-13**	**5.37E-13**	**5.80E-05**
SonarEW	NaN	**6.50E-13**	**1.07E-12**	**8.09E-13**
SpectEW	8.26E-02	**1.63E-11**	**3.87E-11**	**1.15E-04**
Tic-tac-toe	**1.03E-02**	**3.88E-12**	**1.76E-10**	**2.48E-07**
Vote	**3.34E-06**	**9.87E-13**	**6.96E-12**	3.13E-01
WaveformEW	5.59E-01	**2.95E-11**	**2.97E-11**	**2.96E-11**
WineEW	**5.59E-05**	**1.43E-06**	**1.10E-12**	NaN
Zoo	**4.17E-13**	**4.17E-13**	7.55E-01	**4.17E-13**

TABLE 22: Performance results of BTLBO-V-ER with KNN and with other classifiers (Linear Discriminant Analysis (LDA), Decision Tree (DT), and Adaptive Boosting (AdaBoost) in terms of average accuracy, and time

Benchmark	Measure	Accuracy				Time			
		KNN	LDA	DT	AdaBoost	KNN	LDA	DT	AdaBoost
Breastcancer	AVG	**0.9831**	0.9643	0.9750	0.9790	**22.33**	111.92	158.78	4255.13
	STD	0.0040	0.0000	0.0041	0.0042	1.35	18.71	83.04	237.24
BreastEW	AVG	**0.9971**	0.9567	0.9795	0.9664	**22.61**	96.50	114.55	4234.34
	STD	0.0048	0.0056	0.0042	0.0095	1.37	2.76	39.37	214.83
CongressEW	AVG	0.9705	0.9766	0.9805	**0.9866**	**19.35**	94.40	122.15	4015.53
	STD	0.0058	0.0021	0.0054	0.0053	1.12	2.79	55.67	208.26
Exactly	AVG	**1.0000**	0.6450	**1.0000**	**1.0000**	**29.84**	86.74	126.66	5264.17
	STD	0.0000	0.0000	0.0000	0.0000	1.74	6.26	44.63	694.31
Exactly2	AVG	0.7627	0.7750	**0.8080**	0.7678	**28.99**	86.59	122.47	4730.38
	STD	0.0177	0.0000	0.0057	0.0171	5.16	4.25	57.18	465.37
HeartEW	AVG	0.8759	0.8364	0.8759	**0.8815**	**17.63**	92.86	73.39	3995.69
	STD	0.0099	0.0173	0.0099	0.0248	0.80	3.16	2.17	165.39
IonosphereEW	AVG	0.9869	0.9498	0.9793	**0.9916**	**19.15**	95.90	79.05	4166.88
	STD	0.0082	0.0071	0.0080	0.0102	1.09	4.11	3.44	169.68
KrvskpEW	AVG	0.9855	0.9502	**0.9944**	0.9896	233.98	**115.85**	131.17	7785.30
	STD	0.0027	0.0009	0.0024	0.0028	27.38	4.27	8.08	359.71
Lymphography	AVG	**0.9764**	0.9333	0.8931	0.9344	**16.44**	128.79	72.64	4037.06
	STD	0.0251	0.0196	0.0138	0.0355	0.70	7.13	2.11	152.37
M-of-n	AVG	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**27.10**	99.20	78.17	4444.00
	STD	0.0000	0.0000	0.0000	0.0000	1.55	3.66	2.35	194.71
penglungEW	AVG	**1.0000**	**1.0000**	**1.0000**	0.8311	**19.59**	257.37	83.90	4122.97
	STD	0.0000	0.0000	0.0000	0.0694	0.85	10.01	2.71	181.70
SonarEW	AVG	**1.0000**	0.9540	0.9524	0.9373	**17.40**	95.66	80.47	4195.91
	STD	0.0000	0.0165	0.0234	0.0328	0.89	2.87	2.95	175.30
SpectEW	AVG	0.8673	**0.8914**	0.8907	0.8815	**18.10**	91.05	74.51	3963.96
	STD	0.0147	0.0106	0.0075	0.0210	0.84	2.54	1.82	166.66
Tic-tac-toe	AVG	0.8312	0.7134	0.8427	**0.9658**	**29.47**	95.69	91.04	5091.28
	STD	0.0054	0.0010	0.0029	0.0110	1.89	2.64	3.15	201.29
Vote	AVG	**0.9994**	0.9422	0.9978	0.9667	**17.56**	85.72	71.05	3529.50
	STD	0.0030	0.0085	0.0058	0.0000	0.88	5.51	3.99	125.57
WaveformEW	AVG	0.7820	0.8305	0.7791	**0.8454**	620.20	**202.69**	442.88	17052.41
	STD	0.0062	0.0022	0.0066	0.0053	133.21	11.12	44.62	535.72
WineEW	AVG	**1.0000**	**1.0000**	0.9722	**1.0000**	**16.59**	103.45	72.13	4784.70
	STD	0.0000	0.0000	0.0000	0.0000	0.73	3.66	2.33	163.29
Zoo	AVG	0.9238	0.9506	0.9508	**0.9673**	**17.36**	203.93	74.31	4938.91
	STD	0.0237	0.0010	0.0087	0.0346	0.81	8.54	2.04	143.30
Ranking	W|T|L	**5|4|9**	1|3|14	**2|3|13**	5|3|10	**16|0|2**	2|0|16	**0|0|18**	0|0|18

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3064799, IEEE Access

TABLE 23: Comparison of BTLBO-V-ER with other meta-heuristics from the literature in terms of average accuracy

Dataset	BTLBO-V-ER	BSSA_S3_CP[100]	WOA-CM [90]	BGOA_EPD_Tour [88]	GA [101]	PSO [101]	GA [91]	PSO [91]	bGWO1[91]	bGWO2[91]	HGSA [102]	BGOA-M[103]	BDA-TVv4[104]	BGWOPSO[105]	S-bBOA[59]
Breastcancer	**0.983**	0.977	0.968	0.980	0.957	0.949	0.968	0.967	0.976	0.975	0.974	0.974	0.977	0.980	0.969
BreastEW	**0.997**	0.948	0.971	0.947	0.923	0.933	0.939	0.933	0.924	0.935	0.971	0.970	0.974	0.970	0.971
CongressEW	0.970	0.963	0.792	0.964	0.898	0.937	0.932	0.928	0.935	0.938	0.966	0.976	**0.995**	0.980	0.959
Exactly	**1.000**	0.980	0.956	0.999	0.822	0.973	0.674	0.688	0.708	0.776	**1.000**	**1.000**	0.929	**1.000**	0.972
Exactly2	0.763	0.758	**1.000**	0.780	0.677	0.666	0.746	0.730	0.745	0.750	0.770	0.735	0.726	0.760	0.760
HeartEW	0.876	0.861	0.742	0.833	0.732	0.745	0.780	0.787	0.776	0.776	0.856	0.836	**0.886**	0.850	0.824
IonosphereEW	**0.987**	0.918	0.919	0.899	0.863	0.876	0.814	0.819	0.807	0.834	0.934	0.946	0.925	0.950	0.907
KrvskpEW	**0.985**	0.964	0.866	0.968	0.940	0.949	0.920	0.941	0.944	0.956	0.978	0.974	0.971	0.980	0.966
Lymphography	**0.976**	0.890	0.807	0.868	0.758	0.759	0.696	0.744	0.744	0.700	0.892	0.912	0.895	0.920	0.868
M-of-n	**1.000**	0.992	0.926	**1.000**	0.916	0.996	0.861	0.921	0.908	0.963	**1.000**	**1.000**	0.973	**1.000**	0.972
penglungEW	**1.000**	0.878	0.972	0.927	0.672	0.879	0.584	0.584	0.600	0.584	0.956	0.934	0.807	0.960	0.878
SonarEW	**1.000**	0.937	0.852	0.912	0.833	0.804	0.754	0.737	0.731	0.729	0.958	0.915	0.995	0.960	0.936
SpectEW	0.867	0.836	**0.991**	0.826	0.756	0.738	0.793	0.822	0.820	0.822	0.919	0.826	0.877	0.880	0.846
Tic-tac-toe	**0.831**	0.821	0.785	0.808	0.764	0.750	0.719	0.735	0.728	0.727	0.788	0.791	0.822	0.810	0.798
Vote	**0.999**	0.951	0.939	0.966	0.808	0.888	0.904	0.904	0.812	0.920	0.973	0.963	0.962	0.970	0.965
WaveformEW	0.782	0.734	0.753	0.737	0.712	0.732	0.733	0.762	0.786	0.789	**0.815**	0.751	0.749	0.800	0.743
WineEW	**1.000**	0.993	0.959	0.989	0.947	0.937	0.937	0.933	0.930	0.920	0.989	0.989	0.999	**1.000**	0.984
Zoo	0.924	**1.000**	0.980	0.993	0.946	0.963	0.855	0.861	0.879	0.879	0.932	0.958	0.983	**1.000**	0.978
Ranking (W|T|L)	**9|3|6**	0|1|17	2|0|16	0|1|17	0|0|18	0|0|18	0|0|18	0|0|18	0|0|18	0|0|18	1|2|15	0|2|16	2|0|16	0|4|14	0|0|18
Rank(F-Test)	**2.61**	6.28	8.14	6.03	12.50	11.00	12.78	12.06	11.86	10.97	4.39	5.89	5.25	3.00	7.25

VI. CONCLUSION AND FUTURE DIRECTIONS

In this work, an efficient wrapper-based feature selection approach based on a modified binary TLBO as a search algorithm was proposed for variant datasets. Four binarization methods were proposed: Elitist, the Elitist Roulette, the Elitist Tournament, and the Rank based approach. Their impact on the efficacy of different variants were compared to other common binarization methods. The experimental demonstrated that both TFs and binarization approaches have a significant influence on the effectiveness of the proposed binary TLBO, taking into account its exploratory and exploitative potentials, in comparison with well-regarded and recent feature selection methods. It was also noticed that the proposed binarization methods have a more significant impact on the performance of the TLBO algorithm than other methods used in the comparisons. Further investigation on the best combination between binarization methods and TFs revealed that Elitist Tournament is the best for S-shaped TF, while Elitist Rank-based is the best when combined with V-shaped TF. All in all, the BTLBO algorithm combined with Elitist Rank-based and V-shaped is recommended in terms of accuracy and feature reduction rates.

For future work, there are some research avenues. First, investigating other novel binarization methods that consider different strategies in repositioning the current solutions. Second, different TFs can be tested with the proposed binarization methods. This way, researchers can study the behavior of each TF with the different binarization methods. Moreover, other variants of TLBO and other SI algorithms can be tested with the new binarization methods.

ACKNOWLEDGMENT

The authors would like to acknowledge the support received from Taif University Researchers Supporting Project Number (TURSP-2020/125), Taif University, Taif, Saudi Arabia.

REFERENCES

[1] H. Liu and H. Motoda, Computational methods of feature selection. CRC Press, 2007.
[2] L. Huan and M. Hiroshi, Feature Selection for Knowledge Discovery and Data Mining. USA: Kluwer Academic Publishers, 1998.
[3] R. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory,” in Micro Machine and Human Science, 1995. MHS’95., Proceedings of the Sixth International Symposium on. IEEE, 1995, pp. 39–43.
[4] M. Dorigo, V. Maniezzo, and A. Colorni, “Ant system: optimization by a colony of cooperating agents,” IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 26, no. 1, pp. 29–41, 1996.
[5] S. Mirjalili and A. Lewis, “The whale optimization algorithm,” Advances in Engineering Software, vol. 95, pp. 51–67, 2016.
[6] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, “Harris hawks optimization: Algorithm and applications,” Future Generation Computer Systems, vol. 97, pp. 849 – 872, 2019.
[7] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey wolf optimizer,” Advances in Engineering Software, vol. 69, pp. 46–61, 2014.
[8] J. H. Holland, “Genetic algorithms,” Scientific american, vol. 267, no. 1, pp. 66–73, 1992.
[9] R. Storn and K. Price, “Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces,” Journal of global optimization, vol. 11, no. 4, pp. 341–359, 1997.
[10] H. Chen, S. Jiao, A. A. Heidari, M. Wang, X. Chen, and X. Zhao, “An opposition-based sine cosine approach with local search for parameter estimation of photovoltaic models,” Energy Conversion and Management, vol. 195, pp. 927–942, 2019.
[11] I. Aljarah, M. Mafarja, A. A. Heidari, H. Faris, and S. Mirjalili, “Clustering analysis using a novel locality-informed grey wolf-inspired clustering approach,” Knowledge and Information Systems, Apr 2019.
[12] Y. Xu, H. Chen, A. A. Heidari, J. Luo, Q. Zhang, X. Zhao, and C. Li, “An efficient chaotic mutative moth-flame-inspired optimizer for global optimization tasks,” Expert Systems with Applications, vol. 129, pp. 135 – 155, 2019.
[13] A. A. Heidari, I. Aljarah, H. Faris, H. Chen, J. Luo, and S. Mirjalili, “An enhanced associative learning-based exploratory whale optimizer for global optimization,” Neural Computing and Applications, Jan 2019.
[14] R. Jensen and Q. Shen, Computational intelligence and feature selection: rough and fuzzy approaches. John Wiley & Sons, 2008, vol. 8.
[15] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, “From data mining to knowledge discovery in databases,” AI magazine, vol. 17, no. 3, p. 37, 1996.
[16] V. Rodriguez-Galiano, J. Luque-Espinar, M. Chica-Olmo, and M. Mendes, “Feature selection approaches for predictive modelling of groundwater nitrate pollution: An evaluation of filters, embedded and wrapper methods,” Science of the Total Environment, vol. 624, pp. 661–672, 2018.
[17] R. Kohavi and G. H. John, “Wrappers for feature subset selection,” Artificial intelligence, vol. 97, no. 1-2, pp. 273–324, 1997.
[18] R. V. Rao, V. J. Savsani, and D. Vakharia, “Teaching–learning-based optimization: a novel method for constrained mechanical design optimization problems,” Computer-Aided Design, vol. 43, no. 3, pp. 303–315, 2011.
[19] R. Rao, V. Savsani, and D. Vakharia, “Teaching–learning-based optimization: An optimization method for continuous non-linear large scale problems,” Information Sciences, vol. 183, no. 1, pp. 1 – 15, 2012.
[20] B. Crawford, R. Soto, G. Astorga, J. García, C. Castro,

and F. Paredes, "Putting continuous metaheuristics to work in binary search spaces," Complexity, vol. 2017, 2017.
[21] J. Fan, D. Jiang, W. Liu, F. Wu, J. Chen, and J. Daemen, "Discontinuous fatigue of salt rock with low-stress intervals," International Journal of Rock Mechanics and Mining Sciences, vol. 115, pp. 77–86, 2019.
[22] W. Liu, Z. Zhang, J. Fan, D. Jiang, and J. Daemen, "Research on the stability and treatments of natural gas storage caverns with different shapes in bedded salt rocks," IEEE Access, vol. 8, p. 000507, 2020.
[23] W. Liu, Z. Zhang, J. Chen, J. Fan, D. Jiang, D. Jjk, and Y. Li, "Physical simulation of construction and control of two butted-well horizontal cavern energy storage using large molded rock salt specimens," Energy, vol. 185, pp. 682–694, 2019.
[24] W. Qiao, K. Huang, M. Azimi, and S. Han, "A novel hybrid prediction model for hourly gas consumption in supply side based on improved whale optimization algorithm and relevance vector machine," IEEE Access, vol. 7, pp. 88 218–88 230, 2019.
[25] W. Qiao, W. Tian, Y. Tian, Q. Yang, Y. Wang, and J. Zhang, "The forecasting of pm2.5 using a hybrid model based on wavelet transform and an improved deep learning algorithm," IEEE Access, vol. 7, pp. 142 814–142 825, 2019.
[26] L. Jinlong, X. Wenjie, Z. Jianjing, L. Wei, S. Xilin, and Y. Chunhe, "Modeling the mining of energy storage salt caverns using a structural dynamic mesh," Energy, vol. 193, p. 116730, 2020.
[27] W. Liu, Z. Zhang, J. Chen, J. Fan, D. Jiang, D. Jjk, and Y. Li, "Physical simulation of construction and control of two butted-well horizontal cavern energy storage using large molded rock salt specimens," Energy, vol. 185, pp. 682–694, 2019.
[28] J. F. Y. L. L. W. W. Liu, X. Zhang, "Evaluation of potential for salt cavern gas storage and integration of brine extraction: Cavern utilization, yangtze river delta region," Natural Resources Research, vol. 29, 2020.
[29] W. Qiao and Z. Yang, "Modified dolphin swarm algorithm based on chaotic maps for solving high-dimensional function optimization problems," IEEE Access, vol. 7, pp. 110 472–110 486, 2019.
[30] Q. Weibiao and Z. Yang, "Solving large-scale function optimization problem by using a new metaheuristic algorithm based on quantum dolphin swarm algorithm," IEEE Access, vol. 7, pp. 138 972–138 989, 2019.
[31] G. Zhou, H. Moayedi, M. Bahiraei, and Z. Lyu, "Employing artificial bee colony and particle swarm techniques for optimizing a neural network in prediction of heating and cooling loads of residential buildings," Journal of Cleaner Production, vol. 254, 2020.
[32] J. Chen, D. Lu, W. Liu, J. Fan, D. Jiang, L. Yi, and Y. Kang, "Stability study and optimization design of small-spacing two-well (sstw) salt caverns for natural gas storages," Journal of Energy Storage, vol. 27, p. 101131, 2020.
[33] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," Journal of machine learning research, vol. 3, no. Mar, pp. 1157–1182, 2003.
[34] E. Zorarpacı and S. A. Özel, "A hybrid approach of differential evolution and artificial bee colony for feature selection," Expert Systems with Applications, vol. 62, pp. 91–103, 2016.
[35] M. Mafarja, A. A. Heidari, M. Habib, H. Faris, T. Thaher, and I. Aljarah, "Augmented whale feature selection for iot attacks: Structure, analysis and applications," Future Generation Computer Systems, vol. 112 http://aliasgharheidari.com/, 05 2020.
[36] K. Mistry, L. Zhang, S. C. Neoh, C. P. Lim, and B. Fielding, "A micro-ga embedded pso feature selection approach to intelligent facial emotion recognition," IEEE transactions on cybernetics, vol. 47, no. 6, pp. 1496–1509, 2016.
[37] Y. K. Semero, J. Zhang, and D. Zheng, "Pv power forecasting using an integrated ga-pso-anfis approach and gaussian process regression based feature selection strategy," CSEE Journal of Power and Energy Systems, vol. 4, no. 2, pp. 210–218, 2018.
[38] B. Tran, B. Xue, and M. Zhang, "Variable-length particle swarm optimization for feature selection on high-dimensional classification," IEEE Transactions on Evolutionary Computation, vol. 23, no. 3, pp. 473–487, 2018.
[39] Q. Wu, Z. Ma, J. Fan, G. Xu, and Y. Shen, "A feature selection method based on hybrid improved binary quantum particle swarm optimization," IEEE Access, 2019.
[40] Y. Zhang, D.-w. Gong, and J. Cheng, "Multi-objective particle swarm optimization approach for cost-based feature selection in classification," IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB), vol. 14, no. 1, pp. 64–75, 2017.
[41] M. Mafarja, R. Jarrar, S. Ahmad, and A. A. Abusnaina, "Feature selection using binary particle swarm optimization with time varying inertia weight strategies," in Proceedings of the 2Nd International Conference on Future Networks and Distributed Systems, ser. ICFNDS '18. New York, NY, USA: ACM, 2018, pp. 18:1–18:9.
[42] M. Mafarja and N. R. Sabar, "Rank based binary particle swarm optimisation for feature selection in classification," in Proceedings of the 2Nd International Conference on Future Networks and Distributed Systems, ser. ICFNDS '18. New York, NY, USA: ACM, 2018, pp. 19:1–19:6.
[43] S. Rajamohana and K. Umamaheswari, "Hybrid approach of improved binary particle swarm optimization and shuffled frog leaping for feature selection," Computers & Electrical Engineering, vol. 67, pp. 497 – 508, 2018.

[44] Y. Chen, L. Li, J. Xiao, Y. Yang, J. Liang, and T. Li, "Particle swarm optimizer with crossover operation," Engineering Applications of Artificial Intelligence, vol. 70, pp. 159–169, 2018.
[45] R. C. T. De Souza, L. dos Santos Coelho, C. A. De Macedo, and J. Pierezan, "A v-shaped binary crow search algorithm for feature selection," in 2018 IEEE Congress on Evolutionary Computation (CEC). IEEE, 2018, pp. 1–8.
[46] P. Shunmugapriya and S. Kanmani, "A hybrid algorithm using ant and bee colony optimization for feature selection and classification (ac-abc hybrid)," Swarm and Evolutionary Computation, vol. 36, pp. 27–36, 2017.
[47] Y. Wan, M. Wang, Z. Ye, and X. Lai, "A feature selection method based on modified binary coded ant colony optimization algorithm," Applied Soft Computing, vol. 49, pp. 248–258, 2016.
[48] Z. Manbari, F. AkhlaghianTab, and C. Salavati, "Hybrid fast unsupervised feature selection for high-dimensional data," Expert Systems with Applications, vol. 124, pp. 97–118, 2019.
[49] I. Aljarah, M. Mafarja, A. A. Heidari, H. Faris, Y. Zhang, and S. Mirjalili, "Asynchronous accelerating multi-leader salp chains for feature selection," Applied Soft Computing, vol. 71, pp. 964–979, 2018.
[50] H. Faris, M. M. Mafarja, A. A. Heidari, I. Aljarah, A.-Z. Ala'M, S. Mirjalili, and H. Fujita, "An efficient binary salp swarm algorithm with crossover scheme for feature selection problems," Knowledge-Based Systems, vol. 154, pp. 43–67, 2018.
[51] S. Ahmed, M. Mafarja, H. Faris, and I. Aljarah, "Feature selection using salp swarm algorithm with chaos," in Proceedings of the 2nd International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence (2018). ACM, 2018, pp. 65–69.
[52] G. I. Sayed, G. Khoriba, and M. H. Haggag, "A novel chaotic salp swarm algorithm for global optimization and feature selection," Applied Intelligence, vol. 48, no. 10, pp. 3462–3481, Oct 2018.
[53] S. Mirjalili, "Sca: a sine cosine algorithm for solving optimization problems," Knowledge-Based Systems, vol. 96, pp. 120–133, 2016.
[54] R. Sindhu, R. Ngadiran, Y. M. Yacob, N. A. H. Zahri, and M. Hariharan, "Sine–cosine algorithm for feature selection with elitism strategy and new updating mechanism," Neural Computing and Applications, vol. 28, no. 10, pp. 2947–2958, 2017.
[55] M. E. A. Elaziz, A. A. Ewees, D. Oliva, P. Duan, and S. Xiong, "A hybrid method of sine cosine algorithm and differential evolution for feature selection," in International Conference on Neural Information Processing. Springer, 2017, pp. 145–155.
[56] J. Del Ser, E. Osaba, D. Molina, X.-S. Yang, S. Salcedo-Sanz, D. Camacho, S. Das, P. N. Suganthan, C. A. C. Coello, and F. Herrera, "Bio-inspired computation: where we stand and what's next," Swarm and Evolutionary Computation, vol. 48, pp. 220–250, 2019.
[57] Y. Hou, J. Li, H. Yu, and Z. Lia, "Biffoa: A novel binary improved fruit fly algorithm for feature selection," IEEE Access, 2019.
[58] F. Han, C. Yang, Y.-Q. Wu, J.-S. Zhu, Q.-H. Ling, Y.-Q. Song, and D.-S. Huang, "A gene selection method for microarray data based on binary pso encoding gene-to-class sensitivity information," IEEE/ACM transactions on computational biology and bioinformatics, vol. 14, no. 1, pp. 85–96, 2015.
[59] S. Arora and P. Anand, "Binary butterfly optimization approaches for feature selection," Expert Systems with Applications, vol. 116, pp. 147–160, 2019.
[60] J. P. Papa, G. H. Rosa, A. N. de Souza, and L. C. Afonso, "Feature selection through binary brain storm optimization," Computers & Electrical Engineering, vol. 72, pp. 468–481, 2018.
[61] F. Pourpanah, Y. Shi, C. P. Lim, Q. Hao, and C. J. Tan, "Feature selection based on brain storm optimization for data classification," Applied Soft Computing, vol. 80, pp. 761–775, 2019.
[62] T. Li, H. Dong, and J. Sun, "Binary differential evolution based on individual entropy for feature subset optimization," IEEE Access, 2019.
[63] M. Khishe and M. Mosavi, "Chimp optimization algorithm," Expert Systems with Applications, p. 113338, 2020.
[64] Y. Zhang and Z. Jin, "Group teaching optimization algorithm: A novel metaheuristic method for solving global optimization problems," Expert Systems with Applications, vol. 148, p. 113246, 2020.
[65] A. Faramarzi, M. Heidarinejad, B. Stephens, and S. Mirjalili, "Equilibrium optimizer: A novel optimization algorithm," Knowledge-Based Systems, vol. 191, p. 105190, 2020.
[66] E. H. de Vasconcelos Segundo, V. C. Mariani, and L. dos Santos Coelho, "Design of heat exchangers using falcon optimization algorithm," Applied Thermal Engineering, vol. 156, pp. 119 – 144, 2019.
[67] S. Shadravan, H. Naji, and V. Bardsiri, "The sailfish optimizer: A novel nature-inspired metaheuristic algorithm for solving constrained engineering optimization problems," Engineering Applications of Artificial Intelligence, vol. 80, pp. 20 – 34, 2019.
[68] F. Zou, D. Chen, and Q. Xu, "A survey of teaching–learning-based optimization," Neurocomputing, vol. 335, pp. 366–383, 2019.
[69] M. Črepinšek, S.-H. Liu, and L. Mernik, "A note on teaching–learning-based optimization algorithm," Information Sciences, vol. 212, pp. 79–93, 2012.
[70] G. Waghmare, "Comments on "a note on teaching–learning-based optimization algorithm"," Information Sciences, vol. 229, pp. 159–169, 2013.
[71] J. K. Pickard, J. A. Carretero, and V. C. Bhavsar,

"On the convergence and origin bias of the teaching-learning-based-optimization algorithm," Applied Soft Computing, vol. 46, pp. 115–127, 2016.

[72] S. Chinta, R. Kommadath, and P. Kotecha, "A note on multi-objective improved teaching–learning based optimization algorithm (mo-itlbo)," Information Sciences, vol. 373, pp. 337–350, 2016.

[73] H. E. Kiziloz, A. Deniz, T. Dokeroglu, and A. Cosar, "Novel multiobjective tlbo algorithms for the feature subset selection problem," Neurocomputing, vol. 306, pp. 94–107, 2018.

[74] M. Allam and M. Nandhini, "Optimal feature selection using binary teaching learning based optimization algorithm," Journal of King Saud University-Computer and Information Sciences, 2018.

[75] E. Sevinç and T. Dökeroğlu, "A novel hybrid teaching-learning-based optimization algorithm for the classification of data by using extreme learning machines," Turkish Journal of Electrical Engineering and Computer Sciences, vol. 27, pp. 1523–1533, 03 2019.

[76] R. Rao, "Review of applications of tlbo algorithm and a tutorial for beginners to solve the unconstrained and constrained optimization problems," Decision science letters, vol. 5, no. 1, pp. 1–30, 2016.

[77] R. V. Rao, "Applications of tlbo algorithm and its modifications to different engineering and science disciplines," in Teaching Learning Based Optimization Algorithm. Springer, 2016, pp. 223–267.

[78] J. Kennedy and R. Eberhart, "A new optimizer using particle swarm theory," in Micro Machine and Human Science, 1995. MHS '95., Proceedings of the Sixth International Symposium on, 1995, Conference Proceedings, pp. 39–43.

[79] E. Rashedi, H. Nezamabadi-Pour, and S. Saryazdi, "Bgsa: binary gravitational search algorithm," Natural Computing, vol. 9, no. 3, pp. 727–745, 2010.

[80] G. Zhou, H. Moayedi, and L. K. Foong, "Teaching-learning-based metaheuristic scheme for modifying neural computing in appraising energy performance of building," Engineering with Computers, vol. 36, 2020.

[81] J. García, B. Crawford, R. Soto, and G. Astorga, "A clustering algorithm applied to the binarization of swarm intelligence continuous metaheuristics," Swarm and evolutionary computation, vol. 44, pp. 646–664, 2019.

[82] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on, vol. 5. IEEE, 1997, pp. 4104–4108.

[83] S. Mirjalili, G.-G. Wang, and L. d. S. Coelho, "Binary optimization using hybrid particle swarm optimization and gravitational search algorithm," Neural Computing and Applications, vol. 25, no. 6, pp. 1423–1435, 2014.

[84] S. Mirjalili and A. Lewis, "S-shaped versus v-shaped transfer functions for binary particle swarm optimization," Swarm and Evolutionary Computation, vol. 9, pp. 1–14, 2013.

[85] D. E. Goldberg and J. H. Holland, "Genetic algorithms and machine learning," Machine learning, vol. 3, no. 2, pp. 95–99, 1988.

[86] D. E. Goldberg, "Messy genetic algorithms: Motivation analysis, and first results," Complex systems, vol. 4, pp. 415–444, 1989.

[87] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," The American Statistician, vol. 46, no. 3, pp. 175–185, 1992.

[88] M. Mafarja, I. Aljarah, A. A. Heidari, A. I. Hammouri, H. Faris, A. Al-Zoubi, and S. Mirjalili, "Evolutionary population dynamics and grasshopper optimization approaches for feature selection problems," Knowledge-Based Systems, vol. 145, pp. 25 – 45, 2018.

[89] D. Dheeru and E. Karra Taniskidou, "UCI machine learning repository," 2017.

[90] M. Mafarja and S. Mirjalili, "Whale optimization approaches for wrapper feature selection," Applied Soft Computing, vol. 62, pp. 441–453, 2018.

[91] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary grey wolf optimization approaches for feature selection," Neurocomputing, vol. 172, pp. 371–381, 2016.

[92] S. Mirjalili, S. M. Mirjalili, and X.-S. Yang, "Binary bat algorithm," Neural Computing and Applications, vol. 25, no. 3-4, pp. 663–681, 2014.

[93] M. M. Mafarja and S. Mirjalili, "Hybrid whale optimization algorithm with simulated annealing for feature selection," Neurocomputing, vol. 260, pp. 302–312, 2017.

[94] S. García, D. Molina, M. Lozano, and F. Herrera, "A study on the use of non-parametric tests for analyzing the evolutionary algorithms' behaviour: a case study on the cec'2005 special session on real parameter optimization," Journal of Heuristics, vol. 15, no. 6, p. 617, 2009.

[95] N. M. Razali, Y. B. Wah et al., "Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests," Journal of statistical modeling and analytics, vol. 2, no. 1, pp. 21–33, 2011.

[96] W. Gao, J. L. G. Guirao, M. Abdel-Aty, and W. Xi, "An independent set degree condition for fractional critical deleted graphs." Discrete & Continuous Dynamical Systems-Series S, vol. 12, no. 4&5, pp. 877–886, 2019.

[97] W. Gao, J. L. Guirao, B. Basavanagoud, and J. Wu, "Partial multi-dividing ontology learning algorithm," Information Sciences, vol. 467, pp. 35–58, 2018.

[98] W. Gao, W. Wang, D. Dimitrov, and Y. Wang, "Nano properties analysis via fourth multiplicative abc indicator calculating," Arabian journal of chemistry, vol. 11, no. 6, pp. 793–801, 2018.

[99] W. Gao, H. Wu, M. K. Siddiqui, and A. Q. Baig, "Study of biological networks using graph theory," Saudi journal of biological sciences, vol. 25, no. 6, pp. 1212–1219, 2018.

[100] H. Faris, M. Mafarja, A. A. Heidari, I. Aljarah, A. Al-Zoubi, S. Mirjalili, and H. Fujita, "An efficient binary salp swarm algorithm with crossover scheme for feature selection problems," Knowledge-Based Systems, vol. 154, pp. 43–67, 2018.

[101] S. Kashef and H. Nezamabadi-pour, "An advanced aco algorithm for feature subset selection," Neurocomputing, vol. 147, pp. 271–279, 2015.

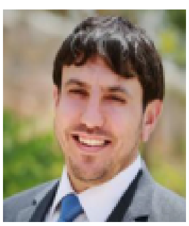
[102] M. Taradeh, M. Mafarja, A. A. Heidari, H. Faris, I. Aljarah, S. Mirjalili, and H. Fujita, "An evolutionary gravitational search-based feature selection," Information Sciences, vol. 497, pp. 219–239, 05 2019.

[103] M. Mafarja, I. Aljarah, H. Faris, A. Hammouri, A. Al-Zoubi, and S. Mirjalili, "Binary grasshopper optimisation algorithm approaches for feature selection problems," Expert Systems with Applications, vol. 117, pp. 267–286, 09 2018.

[104] M. Mafarja, I. Aljarah, A. A. Heidari, H. Faris, P. Fournier Viger, X. Li, and S. Mirjalili, "Binary dragonfly optimization for feature selection using time-varying transfer functions," Knowledge-Based Systems, vol. 161, pp. 185–204, 08 2018.

[105] Q. Al-Tashi, S. Jadid Abdulkadir, H. Rais, S. Mirjalili, and H. Alhussian, "Binary optimization using hybrid grey wolf optimization for feature selection," IEEE Access, vol. 7, pp. 39 496 – 39 508, 04 2019.

THAER THAHER is a PhD student in the Information Technology Engineering at Arab American University, Palestine Polytechnic University, and Al-Quds university, Palestine. He received his B.Sc in Computer Engineering and M.Sc in Advanced Computing from An-Najah National University, Palestine in 2007 and 2018 respectively. His research interests include Evolutionary Computation, Meta-heuristics, data mining, and machine learning.



MAJDI MAFARJA received his B.Sc in Software Engineering and M.Sc in Computer Information Systems from Philadelphia University and The Arab Academy for Banking and Financial Sciences, Jordan in 2005 and 2007 respectively. Dr. Mafarja did his PhD in Computer Science at the National University of Malaysia (UKM). He was a member in Datamining and Optimization Research Group (DMO). Now he is an Associate Professor at the Department of Computer Science at Birzeit University. His research interests include Evolutionary Computation, Meta-heuristics and Data mining.

HAMZA TURABIEH is an Associate professor at Computer Science Department- Faculty of Science and Information Technology- Taif University. Hamza Turabieh received his B.A., M.Sc. degrees in Computer Science from Balqa Applied University in 2004 and 2006 respectively, in Jordan. Turabieh obtained his Ph.D. from National University of Malaysia (UKM) in 2010, his research interests and activities lie at the interface of Computer Science and Operational Research. Intelligent decision support systems, search and optimization (combinatorial optimization, constraint optimization, multi-modal optimization, and multi-objective optimization) using heuristics, local search, meta-heuristics (in particular memetic algorithms, particle swarm optimization), hybrid approaches and their theoretical foundations.

PEDRO A. CASTILLO was born in Granada, in 1974. He received the B.Sc. degree in computer science and the Ph.D. degree from the University of Granada, Spain, in 1997 and 2000, respectively. He has worked as a Teaching Assistant with the Computer Science Department, University of Jaén, Spain, and a Visiting Researcher with Napier University, Edinburg, U.K., in July 1998, and the Santa Fe Institute, Santa Fe, NM, USA, in September 2000. He has led several research projects and directed five Ph.D. students. He is currently an Associate Professor with the Department of Computer Architecture and Technology, University of Granada. His main research interests include bio-inspired systems, hybrid systems, and the combination of evolutionary algorithms and neural networks.

HOSSAM FARIS is a Professor in the Information Technology Department at King Abdullah II School for Information Technology at The University of Jordan, Jordan. Hossam Faris received his B.A. and M.Sc. degrees in computer science from the Yarmouk University and Al-Balqa' Applied University in 2004 and 2008, respectively, in Jordan. He was awarded a full-time competition-based scholarship from the Italian Ministry of Education and Research to peruse his Ph.D. degrees in e-Business at the University of Salento, Italy, where he obtained his Ph.D. degree in 2011. In 2016, he worked as a postdoctoral researcher with the GeNeura team at the Information and Communication Technologies Research Center (CITIC), University of Granada, Spain. His research interests include applied computational intelligence, evolutionary computation, knowledge systems, data mining, semantic web, and ontologies.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3064799, IEEE Access

IBRAHIM ALJARAH is an associate professor of BIG Data Mining and Computational Intelligence at the University of Jordan-Department of Information Technology,Jordan. Currently, he is the Director of the Open Educational Resources and Blended Learning Center at The University of Jordan. He obtained his Ph.D. in computer science from the North Dakota State University, USA, in 2014. He also obtained the master degree in computer science and information systems from the Jordan University of Science and Technology — Jordan in 2006. He obtained the bachelor degree in Computer Science from Yarmouk University — Jordan, 2003. He participated in many conferences in the field of data mining, machine learning, and Big data such as CEC, GECCO, NTIT, CSIT, IEEE NABIC, CASON, and BIGDATA Congress. Furthermore, he contributed in many projects in USA such as Vehicle Class Detection System (VCDS), Pavement Analysis Via Vehicle Electronic Telemetry (PAVVET), and Farm Cloud Storage System(CSS) projects. He has published more than 60 papers in refereed international conferences and journals. His research focuses on Data Mining, Data Science, Machine Learning, Opinion Mining, Sentiment Analysis, Big Data, MapReduce, Hadoop, Swarm intelligence, Evolutionary Computation, and large-scale distributed algorithms.