Molecular Diversity
https://doi.org/10.1007/s11030-022-10410-y

ORIGINAL ARTICLE

A novel nonlinear time-varying sigmoid transfer function in binary whale optimization algorithm for descriptors selection in drug classification

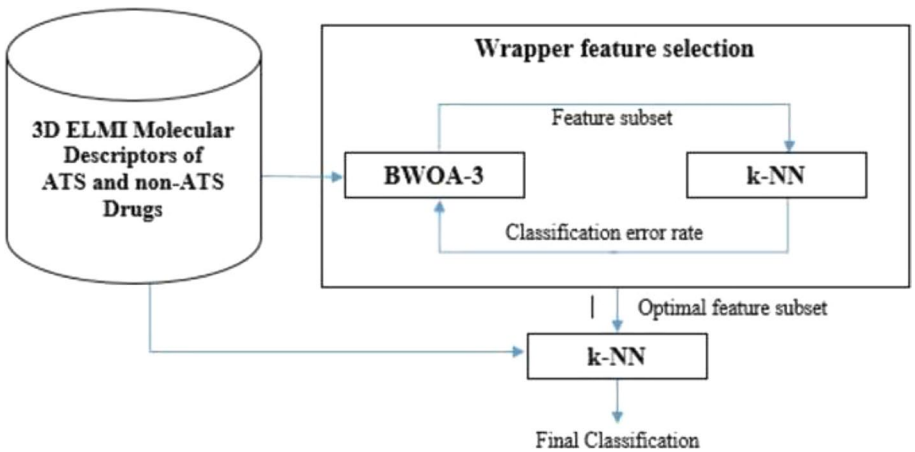
Norfadzlia Mohd Yusof[1] · Azah Kamilah Muda[2] · Satrya Fajri Pratama[2] · Ajith Abraham[3]

Received: 28 December 2021 / Accepted: 15 February 2022

Abstract

In computational chemistry, the high-dimensional molecular descriptors contribute to the curse of dimensionality issue. Binary whale optimization algorithm (BWOA) is a recently proposed metaheuristic optimization algorithm that has been efficiently applied in feature selection. The main contribution of this paper is a new version of the nonlinear time-varying Sigmoid transfer function to improve the exploitation and exploration activities in the standard whale optimization algorithm (WOA). A new BWOA algorithm, namely BWOA-3, is introduced to solve the descriptors selection problem, which becomes the second contribution. To validate BWOA-3 performance, a high-dimensional drug dataset is employed. The proficiency of the proposed BWOA-3 and the comparative optimization algorithms are measured based on convergence speed, the length of the selected feature subset, and classification performance (accuracy, specificity, sensitivity, and f-measure). In addition, statistical significance tests are also conducted using the Friedman test and Wilcoxon signed-rank test. The comparative optimization algorithms include two BWOA variants, binary bat algorithm (BBA), binary gray wolf algorithm (BGWOA), and binary manta-ray foraging algorithm (BMRFO). As the final contribution, from all experiments, this study has successfully revealed the superiority of BWOA-3 in solving the descriptors selection problem and improving the Amphetamine-type Stimulants (ATS) drug classification performance.

Graphical abstract



Keywords Metaheuristic · Feature selection · Descriptors selection · Time-varying transfer function · Binary whale optimization algorithm

✉ Norfadzlia Mohd Yusof
norfadzlia@utem.edu.my

Extended author information available on the last page of the article

Published online: 07 March 2022

Introduction

It was reported in World Drug Report 2021 [1] that there was an increase of more than tenfold the amount of seized Amphetamine-type Stimulants (ATS) drugs for 2015–2019. Methamphetamine was the dominant kind of ATS seized in 2019, followed by amphetamine and "ecstasy." The previous decade has shown the availability of numerous substances in the drug market with the growth of the dynamic market for the non-medical use of prescription and synthetic drugs. Therefore, developing a fast and efficient analytical method for seized ATS drugs and biological samples would be a great tool to prevent ATS drug smuggling and trafficking. It would also contribute to establishing a drug-free society [2].

Cheminformatics can provide a computational method in ATS drug analysis and testing that is much cheaper and faster. The important element in computational chemistry is molecular descriptors. Molecular descriptors keep information about the molecule and have become a support to many contemporary computational models. However, too much information stored in the molecular descriptors may cause the computational model to suffer from the curse of dimensionality. Feature or descriptor selection is a preprocessing technique that is generally used in the cheminformatics domain to determine a relevant descriptors subset from a problem domain but still uphold suitably high accuracy in indicating the original descriptors [3]. Descriptor selection aims to get rid of those noisy, irrelevant, or misleading descriptors for the model-building tasks envisage.

Metaheuristic algorithm, specifically the swarm-intelligence (SI) algorithm, has garnered considerable attention in the cheminformatics domain and achieved competitive results when solving descriptors selection problems. Generally, the SI algorithm is integrated within the wrapper feature selection technique to optimize the feature searching process. Table 1 outlines some of the successful implementation of SI algorithms as molecular descriptors selection techniques.

A binary whale optimization algorithm (BWOA) is proposed in this paper to resolve a molecule descriptors selection problem for ATS drug classification. To generate the BWOA, the newly introduced version of the nonlinear time-varying Sigmoid transfer function is employed. In addition, the recently developed 3D molecular descriptors, namely Three Dimensional Exact Legendre Moment Invariants (3D ELMI) molecular descriptors by Pratama et al. [12], are employed to represent the ATS and non-ATS drug compounds. The calculated molecular descriptors are then used as the dataset to validate the performance of the proposed BWOA.

The structure of this paper is as follows: Sect. 2 gives a detailed description of the whale optimization algorithm (WOA). Section 3 briefly describes the proposed transfer function and its application in BWOA for feature selection. Section 4 provides details regarding the method and materials used to run experiments. The experimental results are discussed in Sect. 5. Lastly, the concluding remarks and opportunities for future work are stated in Sect. 6.

Whale optimization algorithm (WOA)

In 2016, Mirjalili and Lewis proposed an algorithm that is inspired by the hunting mechanism of humpback whales called bubble-net foraging, known as a whale optimization algorithm (WOA) [13]. Initially, the WOA algorithm will assume the target prey as the best search agent (whale) that is near to the optimum. Then, other search agents will update their positions based on the best search agent. WOA swarming behavior is simulated in mathematical formulations below:

Table 1 Previous works on SI-based Descriptors Selection

Year	Application	Swarm Intelligence Algorithms	Paper
2017	QSAR modeling	Salps algorithm	[4]
2018	Drug classification	Chaotic dragonfly algorithm (CDA)	[5]
2020	QSAR modeling	Binary grasshopper optimization algorithm (BGOA)	[6]
2020	QSAR/QSPR classification modeling	Binary pigeon optimization algorithm (BPO)	[7]
2020	Drug classification	Binary particle swarm optimization algorithm (BPSO), BWOA, and binary manta-ray optimization (BMRFO)	[8]
2020	QSAR modeling	Harris hawks optimization (HHO) algorithm	[9]
2020	QSAR modeling	Hybrid Harris hawks optimization with cuckoo search and chaotic map (CHHO–CS)	[10]
2021	QSAR modeling	Seagull optimization algorithm (SOA)	[11]

$$D = \left|C \cdot \text{Whale}^*(t) - \text{Whale}(t)\right|, \tag{1}$$

$$\text{Whale}(t+1) = \text{Whale}^*(t) - A \cdot D, \tag{2}$$

where t is the iteration number. $Whale(t)$ denotes the candidate search agent at iteration number t and $Whale^*(t)$ indicates the best search agent (prey) so far. On the other hand, A and C are coefficient numbers mathematically formulated via Eqs. 3 and 4. Moreover, D indicates the distance vector between the search agent and prey. In each iteration, $Whale^*(t)$ is updated when there is a better solution.

$$A = 2 \cdot a \cdot r + a, \tag{3}$$

$$C = 2 \cdot r, \tag{4}$$

where r is a random vector in [0, 1], while the value of a linearly decreases from 2 to 0 over iterations. The bubble-net behavior of humpback whales in the exploitation phase is designed based on two mechanisms: (1) Shrinking encircling of prey: The humpback move in a shrinking encircling along a spiral-shaped path toward the prey by decreasing a variable value in Eq. 5. A is a random value in the interval $(-a, a)$,

$$a = 2 - t\frac{2}{T}, \tag{5}$$

where t indicates the iteration number and T is the maximum number of iterations. (2) Spiral updating position: A logarithmic spiral function is used to imitate the helix-shaped movement of humpback whales between the candidate search agent $Whale(t)$ and the prey $Whale^*(t)$ so far. This procedure is mathematically expressed in Eq. 7.

$$D^* = \left|\text{Whale}^*(t) - \text{Whale}(t)\right|, \tag{6}$$

$$\text{Whale}(t+1) = D^* \cdot e^{bl} \cdot \cos(2\pi l) + \text{Whale}^*(t), \tag{7}$$

where b is a constant and l is a random number in the range between -1 and 1.

During the optimization phase, an assumption of 50% probability is used to choose between these two mechanisms to update the whales' position. The mathematical formulation to model this behavior is established as follows:

$$\text{Whale}(t+1) = \begin{cases} \text{Whale}^*(t) - A \cdot D, & \text{if } p < 0.5, \\ D^* \cdot e^{bl} \cdot \cos(2\pi l) + \text{Whale}^*(t), & \text{if } p \geq 0.5, \end{cases} \tag{8}$$

where p is a random number in (0, 1).

In the exploration phase, the hunt for prey is conducted at random. Contradicting with the exploitation phase, a search agent position is updated following a random search agent. A contains a random value that is either greater than 1 or less than -1. These values will urge the search agent to move far away from the best search agent. With this mechanism and $|A| > 1$, it allows WOA to perform a global search in overcoming the problem of the local optima. Finally, Eq. 10 describes the mathematical formulation:

$$D = \left|C \cdot \text{Whale}_{\text{rand}} - \text{Whale}\right|, \tag{9}$$

$$\text{Whale}(t+1) = \text{Whale}_{\text{rand}} - A \cdot D, \tag{10}$$

where $Whale_{rand}$ indicates a search agent that is randomly chosen from the current population.

The proposed binary whale optimization algorithm (BWOA)

This study proposed a binary whale optimization algorithm (BWOA) that aims to solve two issues: (1) High-dimensional molecular descriptors and low classification accuracy, (2) Slow convergence rate and difficulty in balancing exploration and exploitation in the existing BWOA. In addition, a new version of the time-varying Sigmoid transfer function is proposed as the improvement strategy in the proposed BWOA and is described in detail in the next section.

New version of nonlinear time-varying sigmoid transfer function

In binary optimization problems, transfer functions were used to control the two basic activities in the swarm intelligence (SI) algorithms: exploration and exploitation [14]. The Sigmoid transfer function was widely used in binary optimization algorithms, but it cannot provide enough balance between these two activities. To overcome the problem, Islam et al. [15] proposed a time-varying transfer function as given below:

$$\text{Sigmoid}\left(\overrightarrow{\text{Whale}}(t+1)\right) = \frac{1}{1 + e^{\left(-\overrightarrow{\text{Whale}}(t+1)/Tv\right)}}, \tag{11}$$

where Tv denotes a control parameter of the time-varying that decreases over iterations. Each element in the whale's position vector is transformed based on Eq. 12 proposed by Kennedy and Eberhart [16] according to the probability value $\text{Sigmoid}\left(\overrightarrow{\text{Whale}}(t+1)\right)$ calculated using Eq. 11:

$$\overrightarrow{\text{Whale}}(t+1) = \begin{cases} 1, \textit{if} \, \text{r and} < \text{Sigmoid}\left(\overrightarrow{\text{Whale}}(t+1)\right), \\ 0, \text{otherwise}, \end{cases} \tag{12}$$

rand is a random number in [0, 1].

The time-varying (Tv) with a linear update is utilized as follows:

$$Tv(t) = Tv_{\max} + \left(Tv_{\min} - Tv_{\max}\right)\left(\frac{t}{T}\right), \tag{13}$$

where Tv_{min}, Tv_{max} are the control parameter's minimum and maximum values, t is the current iteration, while T indicates the maximum number of iterations.

Recently, Kahya et al. proposed two time-varying update strategies [17]: The nonlinear and the decreasing shape and implemented in BWOA for feature selection problems. This points out that the time-varying update techniques over transfer function are very critical to avoid stagnation to local optima. Moreover, this method is employed to explore numerous regions in the search space discovering the global optima. Their findings revealed that BWOA with a nonlinear time-varying update strategy had provided good convergence, a small selected feature subset, and high classification accuracy. It also had overcome the linear and the proposed decreasing shape update strategies.

Therefore, this study proposes a new version of the time-varying Sigmoid transfer function in Eq. 14 that is not presented in [17] to compute the probability of shifting the continuous whales' positions to binary. The standard version of the transfer function in Eq. 13 is employed by Heba to produce a BWOA in [18].

$$\text{Sigmoid}\left(\overrightarrow{\text{Whale}}(t+1)\right) = \frac{1}{1+e^{-10\left(\overrightarrow{\text{Whale}}(t+1)-0.5\right)}}. \tag{13}$$

$$\text{Sigmoid}\left(\vec{W}(t+1)\right) = \frac{1}{1+e^{-10\left(\overrightarrow{\text{Whale}}(t+1)/Tv-0.5\right)}}. \tag{14}$$

A nonlinear time-varying (Tv) update formulation is utilized, and the formulation denotes in Eq. 15 [17]:

$$Tv(t) = Tv_{\max} + \left(Tv_{\min} - Tv_{\max}\right)\left(\frac{t}{T}\right)^{\alpha}. \tag{15}$$

In this study, 4 and 0.01 were set to Tv_{max} and Tv_{min} [19, 20], respectively. Alpha (α) is set to 0.5 similar to the one in [17].

Fitness function

The two objectives of feature selection are to attain high classification accuracy and fewer features [21]. Since wrapper-based feature selection technique is used, a classification algorithm participates in the feature evaluation process. Equation 16 represents the fitness function used to evaluate solutions in the feature selection process where it is designed to have a balance between the two objectives. The small fitness value (small classification error rate and a less number of selected features) implies the best feature subset.

$$\downarrow \text{Fitness} = \alpha \times \text{CER} + \beta \times \frac{\left|F_{\text{select}}\right|}{\left|F_{\text{actual}}\right|}, \tag{16}$$

where CER represents the classification error rate calculated by the classification algorithm. $\left|F_{\text{select}}\right|$ is the length of the selected feature subset, and $\left|F_{\text{actual}}\right|$ is the original size of features in the dataset, α, and β are the two parameters corresponding to the importance of classification quality and feature subset length. $\alpha \in [1, 0]$ and $\beta = (1 - \alpha)$ are adopted from [21, 22]. We considered the classification performance to be the most important metric. Thus, α and β are set to 0.99 and 0.01 [19, 21] (Fig. 1).

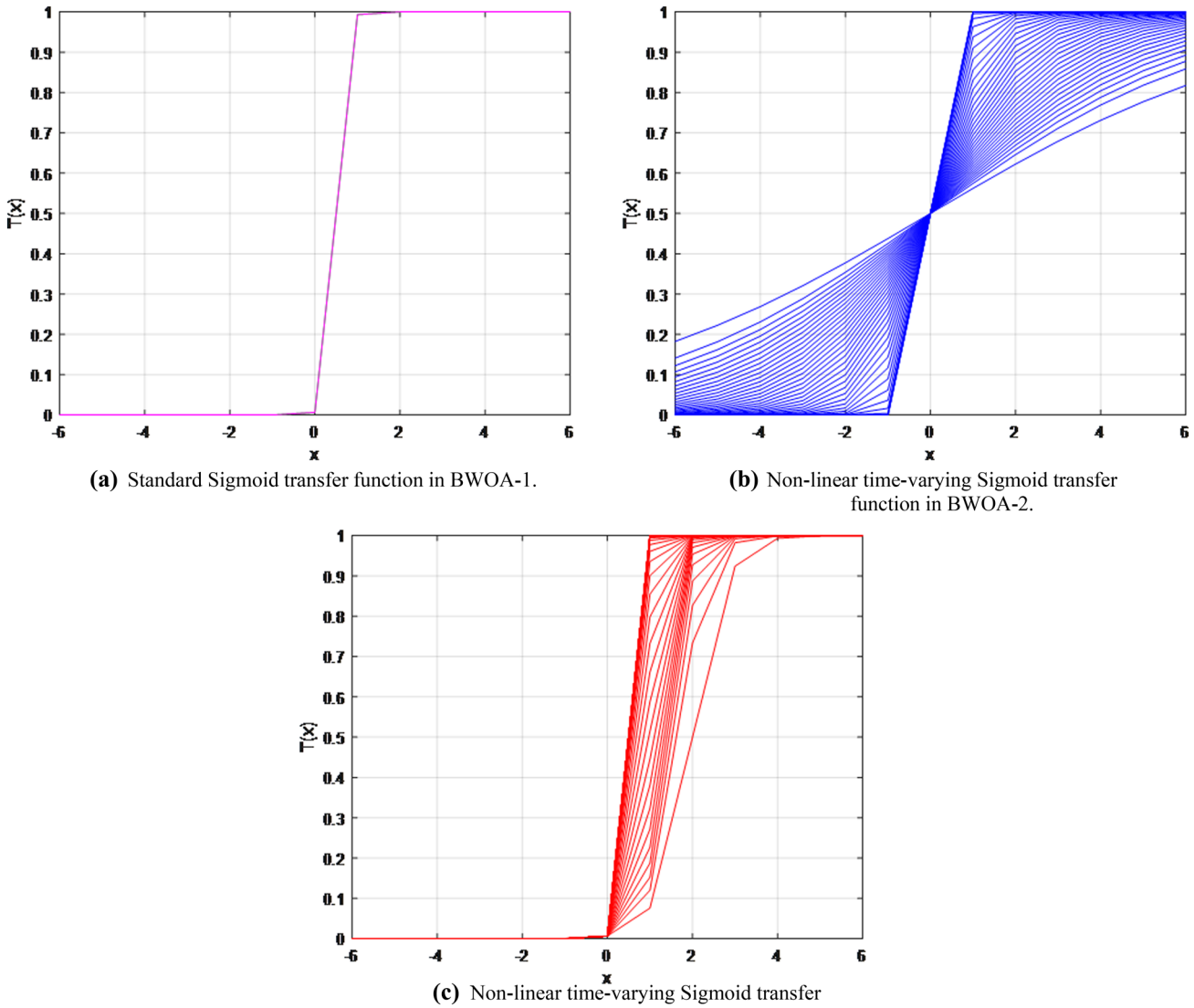
Method and material

Dataset

This study uses the Three-Dimensional Exact Legendre Moment Invariants (3D ELMI) molecular descriptors of Amphetamine-type Stimulants (ATS) and non-ATS drugs as the dataset [12]. The dataset contains 1187 attributes (including the class label) and 7190 drug instances with an equal number of ATS drug molecules (from the pihkal.info database) and non-ATS drug molecules (from ChemSpider database). Table 2 displays the dataset description. In addition, works in [23] described the procedures involved in constructing 3D ELMI molecular descriptors. Besides, research in [24] has employed another type of moment invariants to generate the molecular descriptors for the same ATS and non-ATS drugs.

Methodology

Figure 2 illustrates the architecture of the binary whale optimization algorithm (BWOA) for solving the molecular descriptor selection problem. The proposed approach is expected to accelerate and ease the learning process of the selected classifier after eliminating irrelevant molecular descriptors in the original dataset. This process can improve the classification accuracy efficiently. A new proposed version of the nonlinear time-varying sigmoid transfer function or BWOA-3 is integrated into the wrapper feature selection algorithm. BWOA-3 is responsible for searching and selecting the relevant molecular descriptors. The selected descriptors are then evaluated by the k-NN classification algorithm (with Euclidian distance and $k = 5$) [25] with the implementation of a holdout validation strategy with stratified random sampling of 80% train set and 20% test set [26, 27]. Stratification split is used to reduce the extent of overfitting in the classification model [28, 29]. The classification error rate [18, 26, 30, 31] is calculated from the test set



(a) Standard Sigmoid transfer function in BWOA-1.

(b) Non-linear time-varying Sigmoid transfer function in BWOA-2.

(c) Non-linear time-varying Sigmoid transfer function in BWOA-3.

Fig. 1 Demonstration of the standard Sigmoid transfer function and nonlinear time-varying Sigmoid transfer functions when $Tv_{max} = 4$ and $Tv_{min} = 0.01$ during 70 iterations with time step 2. Note that more vertical curves belong to the lower values of Tv and increasing values of t

Table 2 Dataset description

Attribute	No. of attribute	Attribute type	Description
Molecule id	1	String	The unique reference id of each drug
Feature *n*	1185	Numeric (real numbers)	3D ELMI molecular descriptors. *n* is the descriptor index starting from 1 to 1185
Class	1	Nominal {0,1}	Binary class labels 0 (non-ATS) and 1 (ATS)

as input to BWOA-3 to compute the fitness of each search agent (whale) to discover the global solution. The low fitness value signifies the best whale with an optimal solution. The searching of optimal descriptors is repeated and terminated when the specified maximum iteration is reached. The optimal descriptor subset is then input to the k-NN classifier to get the final classification result.

Parameter settings

All algorithms are repeated ten times in the experiments with different random seeds. The random seed is used to ensure all algorithms started with the same random numbers, evaluated with the same data partitions, and reproducible results. The experimental results are viewed as the

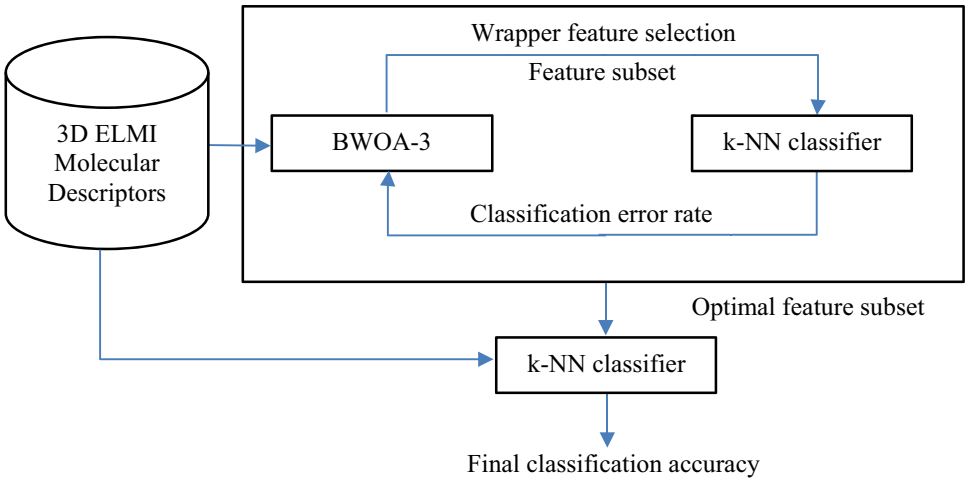


Fig. 2 BWOA Wrapper Feature Selection Architecture

Table 3 Parameter settings

Parameter	Value
Search agent (whale) size, N	8
Iteration length, T	70
No. of runs, M	10
Problem dimension	Same as number of features
Search domain	[0, 1]
$\vec{a}$ in BWOA	Linearly decreased from 2 to 0
Maximum frequency, *fmax* in BBA	2
Minimum frequency, *fmin* in BBA	0
Maximum loudness, A in BBA	2
Maximum pulse rate, r in BBA	1
α and γ in BBA	0.9
Somersault factor, s in BMRFO	2
α in the fitness function	0.99
β in the fitness function	0.01

mean of metrics obtained from ten independent runs to achieve statistically valid results. All algorithms are developed using MATLAB R2021a, performed on PC with an Intel Core i7-6700 machine, 3.40 GHz CPU with Windows 10 operating system, and 16 GB of RAM. Table 3 outlines the parameter settings.

Performance metrics

Various metrics are applied to compare the proposed BWOA-3 with the comparative algorithms, including the average fitness, the standard deviation of fitness, the average accuracy, the standard deviation of accuracy, and the average length of the selected feature subset. In addition, the k-NN classifier performance was measured using sensitivity, specificity, and f-measure derived from the confusion matrix. The sensitivity metric evaluates the ability of the model to predict true positives. The metric specificity measures the ability of the model to predict true negatives, while the f-measure metric measures the level of accuracy that refers to the balance existing between sensitivity and precision. A precision metric is the ratio of the number of true positives that are predicted correctly. Furthermore, this study performed two nonparametric statistical tests: the Friedman test [32] and Wilcoxon signed-rank test [33] on BWOA-3 and the comparative algorithms to check whether their performance differences are significant or not. Also, a qualitative assessment is performed by analyzing the convergence behavior of BWOA-3 and comparative algorithms based on convergence curves visualized in Fig. 3.

Results and discussion

Two experiments were conducted to evaluate the proposed algorithm. The first experiment in Sect. 5.1 aims to validate and compare the performance of the proposed BWOA-3 algorithm with the two binary whale optimization algorithm (BWOA) variants: BWOA-1 [8, 18] and BWOA-2 [34]. The second experiment in Sect. 5.2 aims to compare the performance of BWOA-3 with the other three swarm-intelligence (SI) algorithms: Binary bat algorithm (BBA) [35], Binary

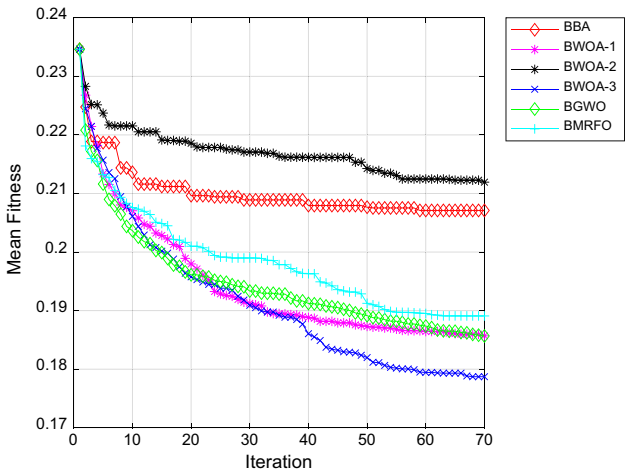


Fig. 3 The convergence curves of BWOA-3 and comparative algorithms

Table 4 Results of average fitness, fitness standard deviation (Std), and average computational time (CT) (seconds) of BWOA-1, BWOA-2, and BWOA-3 algorithms

Algorithm	Avg Fitness	Std	Avg. CT (seconds)
BWOA-1	0.18570	0.01170	271.70
BWOA-2	0.21193	***0.00625***	721.20
BWOA-3	***0.17872***	0.00639	113.50

Table 5 Results of average accuracy, accuracy standard deviation, and the average number of selected descriptors of BWOA-1, BWOA-2, and BWOA-3 algorithms

Algorithm	Avg Accuracy	Accuracy Std	Avg No. of Descriptors
BWOA-1	81.47	1.10489	272
BWOA-2	79.21	***0.59741***	721
BWOA-3	***82.04***	0.66106	***114***

gray wolf algorithm (BGWO) [30], and Binary manta-ray foraging algorithm (BMRFO) [36].

Assessment with other variants of BWOA

In this section, the experimental results achieved by BWOA-1, BWOA-2, and BWOA-3 are presented. Please note that bolded and italic text highlighted the best results. Table 4 reports the average fitness and the standard deviation of fitness. The average accuracy, the standard deviation of accuracy, and the average number of features are indicated in Table 5. BWOA-3 has shown the smallest average fitness that indicates the algorithm has a good ability to avoid trapping in local optima, thus obtaining the optimum solution. Associated with that BWOA-3 is seen has outperformed the BWOA-1 and BWOA-2 in terms of decreasing the number of selected features and increasing the classification accuracy. Table 6 outlines the resultant average of accuracy, specificity, sensitivity, and f-measure by the k-NN classifier employing descriptors selected by BWOA-1, BWOA-2, and BWOA-3 together with the Friedman test ranking value stated in the bracket. According to the Friedman test analysis, the small rank value denotes the best result. By examining the results in Table 6, it witnessed that the BWOA-3 algorithm is superior against the other BWOA variants when attaining the final Friedman rank 1.

Assessment with other SI algorithms

This section displays the comparative results of BWOA-3 with BBA, BGWOA, and BMRFO in Tables 7, 8, and 9. As exhibited in Tables 7 and 8, the BWOA-3 algorithm is seen to overtake others with the lowest average fitness and was capable of selecting the smallest number of descriptors with the fastest computational (convergence) time. Moreover, it is also seen to achieve the highest average classification accuracy. The results in Table 9 show BWOA-3 are also prominent in sensitivity, specificity, and f-measure by obtaining the smallest Friedman rank compared to BBA, BGWOA, and BMRFO algorithms.

Figure 3 demonstrates the convergence curve of all algorithms. From observation, BWOA-3 converged faster and deeper than other algorithms to find the global optimum. This verified that the proposed transfer function had provided the fit balance between exploitation and exploration phases in BWOA-3 that leads to better convergence and high capability in avoiding the local optima. These factors have influenced a satisfactory performance in the BWOA-3 algorithm.

Furthermore, a Wilcoxon signed-rank test based on the mean fitness is applied. The test is to validate whether there are significant differences in average classification accuracy between BWOA-3 and comparative algorithms with a significance level of 0.05. The null hypothesis states that no significant difference between the two algorithms is accepted

Table 6 Comparison of classification performances of BWOA-1, BWOA-2, and BWOA-3 algorithms using Friedman ranks

Algorithm	Avg Accuracy	Avg Sensitivity	Avg Specificity	Avg F-Measure	Avg rank	Final rank
BWOA-1	81.47 (2)	82.49 (1)	80.71 (3)	82.27 (2)	2	2
BWOA-2	79.21 (3)	76.23 (3)	82.98 (2)	80.33 (3)	2.75	3
BWOA-3	82.04 (1)	79.40 (2)	85.22 (1)	82.82 (1)	1.25	***1***

Table 7 Results of average fitness, fitness standard deviation (Std), and average computational time (CT) (seconds) of BWOA-3 with other SI algorithms

Algorithm	Avg Fitness	Fitness Std	Avg. CT (seconds)
BWOA-3	***0.17872***	0.00639	***113.50***
BBA	0.20710	0.01244	283.15
BGWO	0.18567	***0.00471***	494.80
BMRFO	0.18909	0.00723	596.78

Table 8 Performance comparisons between BWOA-3 and other SI algorithms based on the average accuracy, the standard deviation of accuracy, and the average number of selected features

Algorithm	Avg Accuracy	Std Accuracy	Avg No. of Descriptors
BWOA-3	***82.04***	0.66106	***114***
BBA	79.30	1.18100	255
BGWO	81.43	***0.46631***	319
BMRFO	81.02	0.73987	144

when the p-value is greater or equal to 0.05. Otherwise, the null hypothesis is rejected when the p-value is lower than 0.05. The p-value value greater or equal to 0.05 is underlined in Table 10. As seen in Table 10, all the comparative algorithms state a p-value less than 0.05 when compared with BWOA-3, which proved there are significant differences in mean classification accuracy.

As can be observed in Table 11, the descriptors subset selected by the proposed BWOA-3 has improved the performance of the k-NN classifier by effectively reducing 90.71% of the average computational time and increased the average accuracy by 30.44%, average sensitivity by 26.71%, average specificity by 34.99%, and average f-measure by 31.02%. Furthermore, the application of stronger classification algorithms such as support vector machines and random forest

Table 10 P-values of Wilcoxon signed-rank test based on average classification accuracy between BWOA-3 with comparative algorithms

Algorithm	p-value
BWOA-1	0.037
BWOA-2	0.002
BBA	0.002
BGWO	0.002
BMRFO	0.002

are believed can improve the classification accuracy together with the selected descriptors.

Conclusions and future work

In this paper, the new variant of binary whale optimization algorithm (BWOA) with the employment of the new version of the nonlinear time-varying Sigmoid transfer function, namely BWOA-3, was proposed. BWOA-3 is applied to the descriptors selection problem for Amphetamine-type Stimulants (ATS) drug classification. The performance of BWOA-3 is compared with two variants of BWOA, BBA, BGWO, and BMRFO. The experimental results showed that BWOA-3 overtook the comparative optimization algorithms that evidenced that the transfer function selection has significantly enhanced the BWOA algorithm in terms of convergence speed, classification performance, number of selected features, and stability quality. Additionally, it exhibited that the proposed model was efficient, fast, and coherent. This research opens several research directions for future work in the fields of optimization, metaheuristics, feature selection, and applications of these disciplines. As future directions, we think that proposing several new time-dependent transfer functions is highly beneficial to develop enhanced binary optimizers and may change the direction of research in the binary optimization field. As a next step,

Table 9 Comparison of classification performances of BWOA-3 with other SI algorithms using Friedman ranks

Algorithm	Avg Accuracy	Avg Sensitivity	Avg Specificity	Avg F-Measure	Avg rank	Final rank
BWOA-3	82.04 (1)	79.40 (1)	85.22 (1)	82.82 (1)	1	***1***
BBA	79.30 (4)	76.45 (4)	82.87 (4)	79.44 (4)	4	4
BGWO	81.43 (2)	78.68 (3)	84.99 (2)	82.39 (2)	2.25	2
BMRFO	81.02 (3)	78.91 (2)	83.83 (3)	81.87 (3)	2.25	2

Table 11 Comparison of classification performance using original descriptors and descriptors selected by BWOA-3 and without descriptors selection technique

Dataset	Avg CT (seconds)	Avg Accuracy	Avg Sensitivity	Avg Specificity	Avg F-Measure
Original descriptors	3.23	62.89	62.66	63.13	63.21
Descriptors subset by BWOA-3	***0.30***	***82.04***	***79.40***	***85.22***	***82.82***
% Improvement	↓ 90.71%	↑ 30.44%	↑ 26.71%	↑ 34.99%	↑ 31.02%

employing the nonlinear time-varying update formulation to other families of time-dependent transfer functions is an interesting research direction. Furthermore, the efficacy of the proposed BWOA-3 can be further examined by application on the established molecular descriptors. Finally, the proposed BWOA-3 approach can be applied as preprocessing step of many pattern recognition, machine learning, and feature selection tasks.

Acknowledgements This work was supported by Universiti Teknikal Malaysia Melaka through the Fundamental Research Grant Scheme [FRGS/1/2020/FTMK-CACT/F00461] from the Ministry of Higher Education, Malaysia.

Declarations

Conflict of interest The authors do not have any conflict of interest.

References

1. World Drug Report (2021) United Nations publication
2. Chung H, Choe S (2019) Amphetamine-type stimulants in drug testing. Mass Spectrom Lett 10:1–10. https://doi.org/10.5478/MSL.2019.10.1.1
3. Grisoni F, Consonni V, Todeschini R (2018) Impact of molecular descriptors on computational models. In: Brown JB (ed) Computational chemogenomics, 1st edn. Humana Press, New York, pp 171–209
4. Hussien AG, Hassanien AE, Houssein EH (2017) Swarming behaviour of salps algorithm for predicting chemical compound activities. In: 2017 IEEE 8th international conference on intelligent computing and information systems, ICICIS 2017. pp 315–320
5. Sayed GI, Tharwat A, Hassanien AE (2019) Chaotic dragonfly algorithm: an improved metaheuristic algorithm for feature selection. Appl Intell 49:188–205. https://doi.org/10.1007/s10489-018-1261-8
6. Algamal ZY, Qasim MK, Lee MH, Ali HTM (2020) QSAR model for predicting neuraminidase inhibitors of influenza A viruses (H1N1) based on adaptive grasshopper optimization algorithm. SAR QSAR Environ Res 31:803–814. https://doi.org/10.1080/1062936X.2020.1818616
7. Algamal ZY, Qasim MK, Lee MH, Mohammad Ali HT (2020) High-dimensional QSAR/QSPR classification modeling based on improving pigeon optimization algorithm. Chemom Intell Lab Syst 206:104170. https://doi.org/10.1016/j.chemolab.2020.104170
8. Yusof NM, Muda AK, Pratama SF (2021) Swarm intelligence-based feature selection for amphetamine-type stimulants (ATS) drug 3D molecular structure classification. Appl Artif Intell 35:914–932. https://doi.org/10.1080/08839514.2021.1966882
9. Houssein EH, Hosney ME, Oliva D et al (2020) A novel hybrid Harris hawks optimization and support vector machines for drug design and discovery. Comput Chem Eng 133:106656. https://doi.org/10.1016/j.compchemeng.2019.106656
10. Houssein EH, Hosney ME, Elhoseny M et al (2020) Hybrid Harris hawks optimization with cuckoo search for drug design and discovery in chemoinformatics. Sci Rep 10:1–23. https://doi.org/10.1038/s41598-020-71502-z
11. Houssein EH, Hosney ME, Oliva D (2021) A hybrid seagull optimization algorithm for chemical descriptors classification. In: 2021 international mobile, intelligent, and ubiquitous computing conference, MIUCC 2021. pp 381–386
12. Pratama SF (2017) Three-dimensional exact legendre moment invariants for amphetamine-type stimulants molecular structure representation. Universiti Teknikal Malaysia Melaka (UTeM)
13. Mirjalili S, Lewis A (2016) The whale optimization algorithm. Adv Eng Softw 95:51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008
14. Saremi S, Mirjalili S, Lewis A (2015) How important is a transfer function in discrete heuristic algorithms. Neural Comput Appl 26:625–640. https://doi.org/10.1007/s00521-014-1743-5
15. Islam MJ, Li X, Mei Y (2017) A time-varying transfer function for balancing the exploration and exploitation ability of a binary PSO. Appl Soft Comput J 59:182–196. https://doi.org/10.1016/j.asoc.2017.04.050
16. Kennedy J, Eberhart RC (1997) A discrete binary version of the particle swarm algorithm. Proc IEEE Int Conf Syst Man Cybern 5:4104–4108. https://doi.org/10.1109/icsmc.1997.637339
17. Kahya MA, Altamir SA, Algamal ZY (2021) Improving whale optimization algorithm for feature selection with a time-varying transfer function. Numer Algebr Control Optim 11:87–98. https://doi.org/10.3934/naco.2020017
18. Eid HF (2018) Binary whale optimisation: an effective swarm algorithm for feature selection. Int J Metaheuristics 7:67. https://doi.org/10.1504/ijmheur.2018.10012912
19. Mafarja M, Aljarah I, Heidari AA et al (2018) Binary dragonfly optimization for feature selection using time-varying transfer functions. Knowl Based Syst 161:185–204. https://doi.org/10.1016/j.knosys.2018.08.003
20. Chantar H, Thaher T, Turabieh H et al (2021) BHHO-TVS: a binary harris hawks optimizer with time-varying scheme for solving data classification problems. Appl Sci. https://doi.org/10.3390/app11146516
21. Emary E, Zawbaa HM, Hassanien AE (2016) Binary ant lion approaches for feature selection. Neurocomputing 213:54–65. https://doi.org/10.1016/j.neucom.2016.03.101
22. Sharawi M, M. Zawbaa H, Emary E, Zawbaa HM (2017) Feature selection approach based on whale optimization algorithm. In: 2017 ninth international conference on advanced computational intelligence (ICACI), Doha. Institute of Electrical and Electronics Engineers Inc., pp 163–168

23. Pratama SF, Muda AK, Choo YH, Abraham A (2018) Preparation of ATS drugs 3D molecular structure for 3D moment invariants-based molecular descriptors. Advances in Intelligent Systems and Computing. Springe, New York, pp 252–261
24. Pratama SF, Muda AK, Choo YH et al (2017) ATS drugs molecular structure representation using refined 3D geometric moment invariants. J Math Chem 55:1951–1963. https://doi.org/10.1007/s10910-017-0775-3
25. Altman NS (1992) An introduction to kernel and nearest-neighbor nonparametric regression. Am Stat 46:175–185. https://doi.org/10.1080/00031305.1992.10475879
26. Mafarja M, Aljarah I, Faris H et al (2019) Binary grasshopper optimisation algorithm approaches for feature selection problems. Expert Syst Appl 117:267–286. https://doi.org/10.1016/j.eswa.2018.09.015
27. Mafarja M, Mirjalili S (2018) Whale optimization approaches for wrapper feature selection. Appl Soft Comput 62:441–453. https://doi.org/10.1016/j.asoc.2017.11.006
28. Sun Y (2021) Mammograms classification using ELM based on improved sunflower optimization algorithm. J Phys Conf Ser. https://doi.org/10.1088/1742-6596/1739/1/012047
29. Alhudhaif A, Polat K, Karaman O (2021) Determination of COVID-19 pneumonia based on generalized convolutional neural network model from chest X-ray images. Exp Syst Appl. https://doi.org/10.1016/j.eswa.2021.115141
30. Emary E, Zawbaa HM, Hassanien AE (2016) Binary grey wolf optimization approaches for feature selection. Neurocomputing 172:371–381. https://doi.org/10.1016/j.neucom.2015.06.083
31. Emary E, Zawbaa HM, Ghany KKA, et al (2015) Firefly optimization algorithm for feature selection. ACM Int Conf Proceeding Ser 02–04-Sept: https://doi.org/10.1145/2801081.2801091
32. Nadimi-Shahraki MH, Banaie-dezfouli M, Zamani H, Taghian S (2021) B-MFO : a binary moth-flame optimization for feature selection from medical datasets. Computers 10:1–18. https://doi.org/10.3390/computers10110136
33. Wilcoxon F (1945) Individual comparisons by ranking methods. Biomet Bull 1:80–83. https://doi.org/10.2307/3001968
34. Improving-bridge-expansion-and-contraction-installation-replacement-decision-system-using-hybrid-chaotic-whale-optimization-algorithmApplied-Sciences-Switzerland.pdf
35. Mirjalili S, Mirjalili SM, Yang XS (2014) Binary bat algorithm. Neural Comput Appl 25:663–681. https://doi.org/10.1007/s00521-013-1525-5
36. Ghosh KK, Guha R, Bera SK et al (2021) S-shaped versus V-shaped transfer functions for binary Manta ray foraging optimization in feature selection problem. Neural Comput Appl 33:11027–11041. https://doi.org/10.1007/s00521-020-05560-9

Authors and Affiliations

Norfadzlia Mohd Yusof[1] · Azah Kamilah Muda[2] · Satrya Fajri Pratama[2] · Ajith Abraham[3]

[1] Fakulti Teknologi Kejuruteraan Elektrik dan Elektronik, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

[2] Fakulti Teknologi Maklumat dan Komunikasi, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

[3] Machine Intelligence Research Labs (MIR Labs) Scientific Network for Innovation and Research Excellence, Auburn, WA, USA

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).
Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.
These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.
We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.
While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.
These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.
To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.
Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.
If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com