Webology (ISSN: 1735-188X)
Volume 18, Number 6, 2021

A Comparative Evaluation Of Bio-Inspired Optimization Techniques For Feature Selection

A. Ameer Rashed Khan[1] , Dr. S.Shajun Nisha[2] , Dr. M.Mohamed Sathik[3]

[1]Ph.D Research Scholar, PG & Research Department of Computer Science, Sadakathullah Appa College, Rahmath Nagar, Tirunelveli, India. Affiliation of Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli 627012. Reg No.18211192161021,

[2]Assistant Professor & Head, PG & Research Department of Computer Science, Sadakathullah Appa College, Rahmath Nagar, Tirunelveli, India.

[3]Principal, Sadakathullah Appa College, Rahmath Nagar, Tirunelveli, India.

ABSTRACT

Feature selection is an effective approach to reduce the number of features of data, which enhances the performance of classification in machine learning. In this paper, we compare feature selection optimization algorithms to reduce the number of the selected features while enhancing the accuracy. Bio-Inspired Optimization Algorithm is based on biological evolution of nature or from inspirational biological environment. It is an emerging approach used to develop new and robust competing techniques. Optimization means making things better or most effective use of situation. For solving learning and data analysis problems this techniques are used for better performance. In Medical data analysis, optimization techniques and hybrid Bio- Inspired techniques are merged, it is used mainly in Machine learning and Artificial intelligence. In this paper various bio inspired optimization algorithms like Ant Colony Optimization (ACO), Bat Optimization (BAT), Particle Swarm Algorithm (PSO), Grey Wolf Optimization (GWO), Whale Optimization Algorithm (WOA), Slime Mould Optimization (SMA) has been applied to various classification datasets from the UCI machine learning repository.

Keywords: Bio-Inspired, Optimization, Feature Selection, Ant Colony Optimization (ACO), Bat Optimization (BAT), Particle Swarm Algorithm (PSO), Grey Wolf Optimization (GWO), Whale Optimization Algorithm (WOA), Slime Mould Optimization (SMA).

I INTRODUCTION

Machine learning has been widely applied in many practical applications such as data mining, text processing, pattern recognition and medical image analysis, and these fields often rely on the datasets with a large amount of data [1]. However, part of the features may be irrelevant or even misleading for the machine learning algorithms, which increase the computational overhead and reduce accuracy of classification especially for the high-dimensional datasets [2], [3]. Thus, it is necessary to conduct feature selections.

The main principle of feature selection is to find an optimal subset of features which is discriminating from the full dataset, and the selected subset should remain or even enhance the classification performance of the original dataset [4]. Feature selections are useful methods because they can eliminate redundant noise from the datasets so that making the machine learning algorithms perform to execute faster and more efficient. In other words, by using feature selection, the machine learning approaches may perform better while saving costs [5].

Feature Selection techniques can be classified into filter or wrapper. If the FS approach is independent of the learning algorithm then it is called filter approach otherwise wrapper approach. The filter approach is computationally more efficient than the wrapper approach. However, the major drawback of the filter approach is that it may not independent of inductive biases of the learning algorithms that are used for the construction of the classifier. The computational overhead of the wrapper approach is more due to the evaluation of subset of features by using learning algorithms. However, the wrapper approach can provide better results than filter approach in terms of accuracy.

Finding an optimal set of features is challenging and computationally expensive task. Recently, Bio-Inspired Optimization seems to be effective and reliable tools for solving several optimization problems (e.g., machine learning, data mining problems, engineering design, and feature selection) [6]. Bio-Inspired Optimization have been largely employed to solve feature selection problems including: GWO [7], [8], Genetic Algorithm (GA) [9], Ant Colony Optimization (ACO) [10], PSO [11], Differential Evolution (DE) [12], Dragon algorithm (DA) [13], to name a few.

II RELATED WORK

Recently, the area of optimization has gained much attention from researchers especially in hybrid metaheuristics field [14]. For instance, the first proposed feature selection method using hybrid metaheuristic was in 2004 [15] using local search methods and the GA algorithm.

In the literature, PSO has been hybridized with other metaheuristics for continuous search space problems. In [16], for instance, a hybrid PSO with GA (PSOGA) was proposed. Other similar works are: a PSO with DE (PSODE) [17], hybrid PSO and Gravitational Search Algorithm (GSA) (PSOGSA) [18]. Moreover, PSO was hybridized with Bacterial Foraging Optimization algorithm for power system stability enhancement in [19]. These hybrid approaches are aimed to share the strength of each other to expand the capability of exploitation and reducing the chances of dropping in local optimum.

Similarly, GWO has gained much attention in the hybrid metaheuristics field. For instance, in [20] and [21], the authors have hybridized GWO with DE for test scheduling and continuous optimization. Tawhid and Ali [22] have hybridized GWO with GA for minimizing potential energy functions. Gaidhane and Nigam [23] proposed a hybridized GWO and Artificial Bee Colony (ABC) to improve the complex systems performance. Another hybrid method is GWOSCA proposed in [24] using GWO and Sine Cosine Algorithm (SCA). These studies have shown that the hybrid methods performed much better compared to other global or local search methods.

Metaheuristics have been popular in the field of feature selection as well. For instance, a hybrid filter feature selection approach has been proposed in [25] using SA with GA to improve the search ability of GA, the performance was evaluated on eight datasets collected from UCI and obtained a good outcomes considering the selected number of attributes. Another study hybridized GA with SA and evaluated on the Farsi characters hand-printed [26].

Moreover, a hybrid PSO with novel local search strategy based on information correlation was proposed in [27]. A hybrid GA with PSO named GPSO for wrapper feature selection using SVM classifier for classifying microarray data [28]. In the same filed unreliable data, the authors proposed a hybrid mutation operator for an improved multi objective PSO [29]. For Digital Mammogram datasets, a hybrid GA with PSO to enhance the feature set was proposed in [30]. In [31] and [32], two hybrids were proposed using ACO and GA to perform feature selection. Another similar method can be found in [32]. In [33], a hybrid of DE and ABC was used as a feature selector. For the same purpose, Nekkaa and Boughaci [34], proposed a hybrid harmony search algorithm with a local stochastic search. Recently, in [35] a hybrid WOA and SA was proposed for wrapper feature selection. Besides, a hybrid between GWO and antlion optimization (ALO) for feature selection was proposed in [36].

In spite of the good performance of above-mentioned methods we can state that none of them is capable to solve all problems related to feature selection. As such, improvements can be made to the existing methods to enhance the solutions of feature selection problems. In the next section, the methodology is discussed, and the proposed binary hyper metaheuristic is clearly explained.

Table 1: Comparison of Various Optimization Algorithms

S. No	Optimization Algorithm	Source of Inspiration	Algorithm Advantages	Algorithm parameters	Algorithm Application
1	Ant Colony optimization (ACO) 2004	It is based on the way an ant colony will send out workers to hunt randomly for food, who leave pheromone trails behind them	1 - Inherent parallelism 2-Positive Feedback accounts for rapid discovery of good solutions 3- Efficient for Traveling Salesman Problem and similar problems 4-Can be used in dynamic applications (adapts to changes such as new distances, etc)	Ant Colony optimization has three parameters: 1 - No of ants 2- Pheromone intensity 3- Pheromone Evaporation	It has been applied to job shop scheduling, frequency assignment, network load balancing, graph coloring and to detect system faults, machine learning related optimization, and bioinformatics problems.
2	Bat Optimization (BAT) 2010	Inspired from Echolocation Features for hunting Process.	1-It is easy to successfully formulate for Continuous optimization problems. 2- It is easy to implement because it has a Good capability of parameters Adjustment. 3-it has quick convergence rate by Switching from intensification to Diversification to obtain optimal Solutions. 4- it is considered a promising algorithm.	Bat algorithm has Two parameters : 1- Loudness 2- pulse rate	Bat algorithms used to Solve problems in many Filed such as engineering, Computer science, Mathematics, energy, Materials science and Optimization, Classifications (Image Processing, Feature Selection, Scheduling, Data Mining) and so on.
3	Grey Wolf Optimization (GWO) 2014	Inspired from the Social hierarchy and Hunting mechanism of	1- It has the ability to avoid local optima. 2- it provides higher performance in Unknown and challenging search	Grey wolf Optimization has Only two Parameters: 1- A 2- C	Grey wolf optimization Algorithm used to solve problem in many filed Such as engineering, Optimization problems, Multi-objective Problems and so on.

		grey wolves in Nature.	Space. 3- it provides high performance on Constrained problems not only on Unconstrained problems.		
4	Particle Swarm Optimization (PSO) 1995	It is inspired by the social behavior of bird flocking and fish schooling.	1- Simple concept, 2-Easy implementation 3-Robustness to control parameters 4-Computational efficiency when compared with mathematical algorithm and other heuristic optimization techniques	Particle Swarm Optimization has two Parameters: 1- p best 2- g best	Electronics, Electromagnetic, signal, image and video processing, neural networks, communication networks
5	Whale Optimization (WOA) 2016	It mimics the foraging of humpback whales	The uniqueness of the whale algorithm is the ability to employ a random or best agent in the search space to chase the prey. It also has the ability to simulate the bubble-net attaching mechanisms of the humpback whale by using spirals	Whale Optimization has two Parameters: 1- b 2- p	Data mining, Machine learning, IOT, wireless sensor, cryptography, robotics,
6	Slime Mould Optimization (SMA) 2020	Inspired by the behaviors of the slime mold in obtaining the optimal path for connecting food	1-SMA enables to maintain a certain disturbance rate while guaranteeing fast convergence, thus avoiding local trapping during fast convergence 2-It ensures the efficiency of the early exploration and the accuracy of the later exploitation	Slime Mould Optimization has three Parameters: 1- Wrapping food 2- Approaching food 3-Finding food	in optimizing constrained problems and SMA is also applicable to engineering optimization problems in real life with satisfactory optimization results

			3-It can be used to optimize classifiers such as SVM and KNN		

III METHODOLOGY

In this work, we select 10 datasets from the widely used UCI irvine Machine Learning Repository and the main information of these selected datasets are shown in Table 2. The outline of the work is depicted in Fig.1

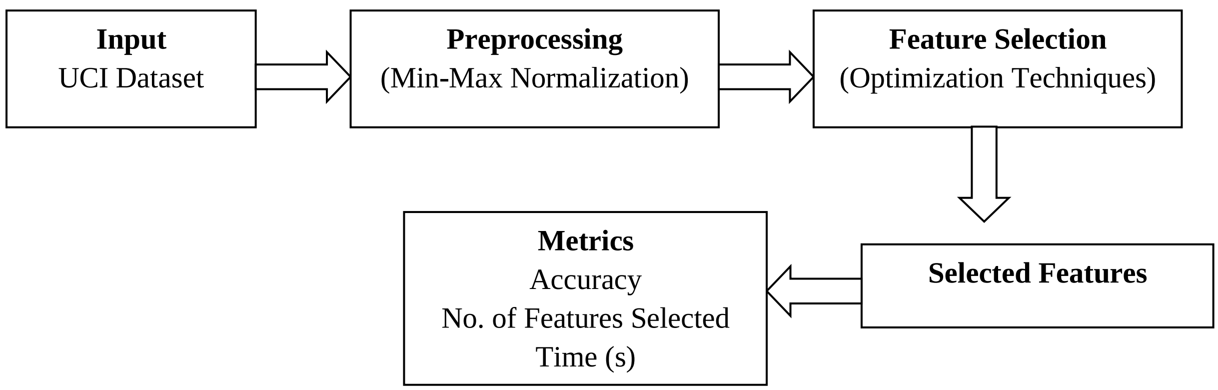

Fig.1 Outline of the Work

1) Preprocessing

Normalization is commonly used to maintain the balance of significance amongst the attributes, when attributes are on a diverse scale. When datasets are with diverse range of attributes, they are preprocessed by min–max normalization method. In this process all the values are transferred into same scale between 0 and 1, thus giving importance to the attribute even with the low range of value on scale. It is the method of scaling the given dataset within the specified range of values between 0 and 1. From the Eq. (1), the normalized feature is derived.

$$n' = \frac{n - min_{ds}}{max_{ds} - min_{ds}} (n\mathrm{fmax} - \mathrm{nfmin}) + \mathrm{nfmin} \qquad (1)$$

2) Feature Selection

a) Ant Colony Optimization (ACO)

ACO algorithms are based on the way an ant colony will send out workers to hunt randomly for food, who leave pheromone trails behind them. Where worker ants find food, they will then proceed back to the nest reinforcing their pheromone trail. If there are two paths to the food, initially ants will randomly pick a path, but over time the shorter path will be more often reinforced. Other ants will be attracted to follow stronger pheromone trail paths, and by this method the colony will swiftly find food sources, and get a large number of workers to them by the most efficient paths [37].

As ACO is employed as a meta-heuristic ACO has been applied in several important areas. The first problem the algorithm was tested upon was the travelling salesman problem, which is known to be NP-hard and extremely challenging depending on the number of cities included[37,38]. ACO has since then been applied to engineering problems[37], vehicle routing problems, machine learning related optimization, and bioinformatics problems, amongst others[37]. It has proven to be an effective and interesting meta-heuristic algorithm.

ACO suffers from many of the same issues as genetic algorithms. Its effectiveness is there, when enough computing power or time is available to be utilised in solving a given problem. A sufficient number of iterations need to be completed to achieve convergence and the algorithm can be extremely complex to execute, especially when the underlying search strategy is a more complex one[37].

ACO is very promising when applied to single objective optimization although it has been somewhat overtaken in application by more common genetic algorithms or simpler search strategies where compute power or time is limited. There have been some efforts to modify ACO to multi-objective paradigm and this is an interesting avenue that is perhaps deserving of further study[39].

b) Particle Swarm optimization (PSO)

Particle swarm optimization (PSO) is a computational intelligence oriented, stochastic, population-based global optimization technique proposed by Kennedy and Eberhart in 1995[9]. It is inspired by the social behavior of bird flocking and fish schooling. PSO has been applied to many engineering problems due to its unique searching Mechanism, simple concept, computational efficiency and easy implementation. It utilizes a “population” of particles that fly through the problem hyperspace with given velocities. At each iteration, the velocities of the individual particles are stochastically adjusted according to the historical best position for the particle itself and the neighborhood best position. Both the particle best and the neighborhood best are derived according to a user defined fitness function. The movement of each particle naturally evolves to an optimal or near-optimal solution.

PSO is not largely affected by the size and nonlinearity of the problem, and can converge to the optimal solution in many problems where most analytical methods fail to converge. Each particle (population member) in the swarm correspond to a solution in a high-dimensional space with four vectors, its current position, best position found so far, the best position found so far by its neighborhood and its velocity and adjusts its position in the search space based on the best position reached by itself (pbest) and its neighbor (gbest) during the search process. PSO is easier to implement and there are fewer parameters to adjust. PSO has a more effective memory capability than the GA. PSO maintains diversity as all the particles use the information related to the most successful particle in order to improve themselves, whereas in GA the worse solutions are removed and only the good ones are saved.

c) Bat Optimization (BAT)

The bat algorithm [40] was proposed by Xin-She Yang, based on the echolocation of micro bats. Bats usually use echolocation to find food. During removal, bats usually send out short pulses, however, when they encounter food, their pulse send out rates increase and the frequency goes up. The increase in frequency means frequency-tuning, which shortens the echolocations time and increases the location accuracy. In addition, this amazing orientation mechanism makes bats being able to distinguish the difference between an obstacle and a prey, allowing them to hunt even in complete darkness [41].

The main two parameters that are used in this algorithm are pulse rates and emission, and the values of these two parameters can be tweaked. The Bat Algorithm also utilizes the frequency-tuning method, to expand the variety of solutions that are present in the population, even though at the same time. It uses automatic zooming that attempts to adjust the exploration and exploitation throughout the process by imitating the variation of the heartbeat outflow and the loudness of the bats during the hunting of their preys. There is a unique ability of echolocation of micro-bats which can find their prey and separate different sorts of the insect in total darkness. These excellent characteristics make the algorithm more efficient with an excellent quick start.

d) Grey Wolf Optimization (GWO)

The GWO is initially proposed by Mirjalili et al. [42], and its algorithm is inspired by the democratic behavior and the hunting mechanism of grey wolves in the wild. In a pack, the grey wolves follow very firm social leadership hierarchy. The alpha(α) wolves are described as the leaders of the pack with male and female, which is considered as the fittest solution. The second level of grey wolves, which are subordinate wolves that help the leaders, is called beta (β) which is known as the second best solution.

Deltas (δ) are the third level of grey wolves which has to submit to alphas and betas, but dominate the omega, and this level of wolves is taken as a third best solution, respectively. The lowest rank of the grey wolves is omega (ω), which have to surrender to all the other governing wolves. The candidate solutions which are left over are taken as omega (ω). In the GWO, the optimization (hunting) is guided by alpha, beta, and delta. The omega wolves have to follow these, β and δ wolves.

e) Whale Optimization Algorithm (WOA)

WOA is a new metaheuristic algorithm proposed by Mirjalili and Lewis [43] and mimics the foraging of humpback whales. The humpback whales hunt school of krill or small fishes close to the surface by swimming around them within a shrinking circle and creating distinctive bubbles along a circle or ‘ 9’-shaped path. Encircling prey and spiral bubble-net attacking method were represented in the first phase of the algorithm; exploitation phase, the second phase where search randomly for a prey (exploration phase). The following subsections discuss the mathematical model of each phase in details. Note that in the equations, a uniform distribution will be used to generate random numbers.

WOA is a new optimization algorithm [43]. It is inspired to mimic the natural behavior of the humpback whales. These whales are usually depending on the hunting behavior as a way of survival. However hunting strategy have been previously introduced to address optimization problems, the uniqueness of the whale algorithm is the ability to employ a random or best agent in the search space to chase the prey. It also has the ability to simulate the bubble-net attaching mechanisms of the humpback whale by using spirals [43].The modelling of this algorithm includes three operators simulate the search for prey (exploration phase), the encircling prey, and the bubble-net foraging (exploitation phase) behavior of humpback whales.

f) Slime Mould Optimization Algorithm (SMA)

Chen [44] has recently been proposed a new optimization algorithm inspired by the behaviors of the slime mold in obtaining the optimal path for connecting food. This algorithm was known as the slime mold algorithm (SMA).

The mathematical model of the SMA based on Chen proposition [44] is described in the following. In the first stage, when SMA searches for the food, it uses its odor in the air as a means of reaching the food. Based on the behavior of the slime mold, it is formulated as follows to simulate the contraction mode.

3. Evaluation Measures

The datasets are randomly partitioned into three diverse equivalent portions (e.g., validation, training, and testing datasets). The dividing of the data is repeated for multiple times to guarantee strength and measurable noteworthiness of the outcomes. The following statistical measures are tested from the validation data in each run:

1) The Average of Classification Accuracy

It is an indicator depicts how precise is the classifier given the chosen set of features when algorithm run N times, and it is calculated as follows:

$$AvgAcc = \frac{1}{N}\sum_{k=1}^{N} AvgAcc^{k} \quad (2)$$

where $AvgAcc^{k}$ is the value of accuracy gained at run k.

2) The Average of Selected Feature

It is an indicator to the average selected features to the overall features when algorithm run N times, and it is calculated as follows:

$$AvgSelection = \frac{1}{N}\sum_{k=1}^{N} \frac{AvgSelection^{k}}{M} \quad (3)$$

where $AvgSelectionk$ is the selected features at run k, and M shows the dataset's total number of features.

3) Average Computational Time

It is an indicator to the average of computational time in seconds gained when algorithm run N times, and it is calculated as follows:

$$AvgCT = \frac{1}{N}\sum_{k=1}^{N} AvgCT^{k} \quad (4)$$

where $AvgCTk$ is the value of computational time gained at run k.

IV RESULTS AND ANALYSIS

In this section, we conduct tests to evaluate the performance of the different optimization algorithm for feature selections.

Table 2 shows the different UCI benchmark datasets and its features.

TABLE 2. UCI Benchmark datasets and its features

S.No	Dataset	Instances	No. Features	Classes
1	Breast Cancer	699	9	2
2	Breast EW	569	30	2
3	Heart EW	270	13	2
4	Lymphography	148	18	4
5	Lung Cancer	226	23	2
6	Diabetic	1151	19	2
7	Parkinsons	195	22	2
8	Vote	300	16	2
9	Tic-tac-toe	958	9	3
10	Zoo	101	16	7

Table 3 shows the different datasets and its extracted features.

TABLE 3: Classification Accuracy comparison between the optimization algorithms using various datasets

S.No	Dataset	Full Dataset	ANT	BAT	PSO	WOA	GWO	SMA
1	Breast Cancer	96.3	91	93	94	96	**97**	96.5
2	Breast EW	93	92	94	94	97.4	**98**	97.7
3	Heart EW	77	77	78	82	83.2	84	**84.6**
4	Lymphography	70	70.2	69.4	71	86	**92**	89
5	Lung Cancer	83	82	76.7	84	84	86	**88.2**
6	Diabetic	60.3	61.3	61.3	62.85	61.3	**66.2**	65.1
7	Parkinsons	78	89.2	90	90.55	90.7	**92**	**92**
8	Vote	87	89	89	96	96	**97**	96
9	Tic-tac-toe	72	72	71	74.5	79.8	**86**	81
10	Zoo	80	83	86	87	87.4	88	**88.3**

TABLE 4: Average selected features comparison using optimization algorithms.

S.No	Dataset	Full Dataset	ANT	BAT	PSO	WOA	GWO	SMO
1	Breast Cancer	9	5.2	5.7	5.09	4	**3.6**	4.2
2	Breast EW	30	19	16.5	13	13	14	**12**

3	Heart EW	13	9.49	7.42	6.18	5.9	**5.4**	5.8
4	Lymphography	18	11.05	9.2	8.3	8	**7.2**	7.6
5	Lung Cancer	23	15.6	12.4	10.4	12.4	8.4	**4.25**
6	Diabetic	19	13	10.4	9.5	8.5	7.44	**7**
7	Parkinsons	22	18	13.8	11.2	11	**8.4**	10.2
8	Vote	16	8.8	7.4	7.2	7.4	**6.5**	7.2
9	Tic-tac-toe	9	6.6	6.85	6	6	5.9	**5.2**
10	Zoo	16	8.3	9.7	8.4	8.1	6.3	**5.9**

TABLE 5: Computational time (seconds) comparison between the optimization algorithms.

S.No	Dataset	ANT	BAT	PSO	WOA	GWO	SMO
1	Breast Cancer	41.4	42.63	26.2	24.2	**10.8**	22.6
2	Breast EW	44.3	50.92	20.02	36.7	**12.6**	31.5
3	Heart EW	29.9	25.9	23.4	30	**9.6**	23.2
4	Lymphography	26.7	21.4	19.5	19	**8**	18.3
5	Lung Cancer	30.4	26.6	21.9	21.5	**12.8**	19.4
6	Diabetic	27.7	19.5	22.3	22.1	**16.4**	20.3
7	Parkinsons	31.3	27.1	26.8	25	**20.2**	25.3
8	Vote	30.9	24.6	23.4	23.1	19.6	**15.6**
9	Tic-tac-toe	56.89	40.4	35.9	34.4	**22.3**	33.2
10	Zoo	27.02	19	18.5	19.2	**14.8**	17.3

Conclusion

Bio inspired algorithms are creating a paradigm shift in the realm of Computer Science. These algorithms are inspired by nature and so their boundaries are boundless. By hybridizing the algorithms we can proceed to the next generation modeling and computing. This work renders a summary of various algorithms to afford optimization and enhanced feature selection. In this paper, performance analysis on various Bio inspired Optimization techniques for benchmark data set is analyzed. Among the many optimization techniques here only six are compared. Mainly: Ant Colony Optimization (ACO), Bat Optimization (BAT), Particle Swarm Algorithm (PSO), Grey Wolf Optimization (GWO), Whale Optimization Algorithm (WOA), Slime Mould Optimization (SMA). Accuracy, Features selected, computational time, are calculated for all the six models. We found that Grey Wolf Optimization (GWO) performs well with an Accuracy, Features selected and Computational Time. The Slime Mould Optimization (SMA) technique comes closer to the accuracy measure of Grey Wolf Optimization (GWO).

REFERENCES

[1] M. I. Jordan and T. M. Mitchell, ``Machine learning: Trends, perspectives, and prospects," Science, vol. 349, no. 6245, pp. 255_260, Jul. 2015.

[2] H. Faris, A. A. Heidari, A. M. Al-Zoubi, M. Mafarja, I. Aljarah, M. Eshtay, and S. Mirjalili, ``Time-varying hierarchical chains of salps with random weight networks for feature selection," Expert Syst. Appl., vol. 140, Feb. 2020, Art. no. 112898.

[3] J. Li, K. Cheng, S.Wang, F. Morstatter, R. P. Trevino, J. Tang, and H. Liu, ``Feature selection: Adata perspective," ACM Comput. Surv., vol. 50, no. 6, pp. 1_45, Jan. 2018.

[4] M. Mafarja, I. Aljarah, H. Faris, A. I. Hammouri, A. M. Al-Zoubi, and S. Mirjalili, ``Binary grasshopper optimisation algorithm approaches for feature selection problems," Expert Syst. Appl., vol. 117, pp. 267_286, Mar. 2019.

[5] Q. Al-Tashi, S. J. A. Kadir, H. M. Rais, S. Mirjalili, and H. Alhussian, ``Binary optimization using hybrid grey wolf optimization for feature selection," IEEE Access, vol. 7, pp. 39496_39508, 2019.

[6] E.-G. Talbi, Metaheuristics: From Design to Implementation, vol. 74, Hoboken, NJ, USA: Wiley, 2009.

[7] E. Emary, H. M. Zawbaa, and A. E. Hassanien, ``Binary grey wolf optimization approaches for feature selection," Neurocomputing, vol. 172, pp. 371_381, Jan. 2016.

[8] Q. Al-Tashi, H. Rais, and S. Jadid, ``Feature selection method based on grey wolf optimization for coronary artery disease classi_cation," in Proc. Int. Conf. Reliable Inf. Commun. Technol., 2018, pp. 257_266.

[9] M. M. Kabir, M. Shahjahan, and K. Murase, ``A new local search based hybrid genetic algorithm for feature selection," Neurocomputing, vol. 74, no. 17, pp. 2914_2928, Oct. 2011.

[10] S. Kashef and H. Nezamabadi-pour, ``An advanced ACO algorithm for feature subset selection," Neurocomputing, vol. 147, pp. 271_279, Jan. 2015.

[11] R. Bello,Y. Gomez, A. Nowe, and M. M. Garcia, ``Two-step particle swarm optimization to solve the feature selection problem," in Proc. 7th Int. Conf. Intell. Syst. Des. Appl. (ISDA), Oct. 2007, pp. 691_696.

[12] E. ZorarpacI and S. A. Özel, ``A hybrid approach of differential evolution and artificial bee colony for feature selection," Expert Syst. Appl., vol. 62,pp. 91_103, Nov. 2016.

[13] M. M. Mafarja, D. Eleyan, I. Jaber, A. Hammouri, and S. Mirjalili, ``Binary dragonfly algorithm for feature selection," in Proc. Int. Conf. New Trends Comput. Sci. (ICTCS), Oct. 2017, pp. 12_17.

[14] E.-G. Talbi, Metaheuristics: From Design to Implementation, vol. 74, Hoboken, NJ, USA: Wiley, 2009.

[15] X. Lai and M. Zhang, ``An ef_cient ensemble of GA and PSO for real function optimization," in Proc. 2nd IEEE Int. Conf. Comput. Sci. Inf. Technol., Aug. 2009, pp. 651_655.

[16] B. Niu and L. Li, ``A novel PSO-DE-based hybrid algorithm for global optimization," in Proc. Int. Conf. Intell. Comput., 2008, pp. 156_163.

[17] S. Mirjalili and S. Z. M. Hashim, ``A new hybrid PSOGSA algorithm for function optimization," in Proc. Int. Conf. Comput. Inf. Appl., Dec. 2010, pp. 374_377.
[18] S. M. Abd-Elazim and E. S. Ali, ``Ahybrid particle swarm optimization and bacterial foraging for power system stability enhancement," Complexity, vol. 21, no. 2, pp. 245_255, 2015.

[19] A. Zhu, C. Xu, Z. Li, J.Wu, and Z. Liu, ``Hybridizing grey wolf optimization with differential evolution for global optimization and test scheduling for 3D stacked SoC," J. Syst. Eng. Electron., vol. 26, no. 2, pp. 317_328, Apr. 2015.

[20] D. Jitkongchuen, ``A hybrid differential evolution with grey wolf optimizer for continuous global optimization," in Proc. 7th Int. Conf. Inf. Technol. Elect. Eng. (ICITEE), Oct. 2015, pp. 51_54.

[21] M. A. Tawhid and A. F. Ali, ``A Hybrid grey wolf optimizer and genetic algorithm for minimizing potential energy function," Memetic Comput., vol. 9, no. 4, pp. 347_359, 2017.

[22] P. J. Gaidhane and M. J. Nigam, ``A hybrid grey wolf optimizer and arti_cial bee colony algorithm for enhancing the performance of complex systems," J. Comput. Sci., vol. 27, pp. 284_302, Jul. 2018.

[23] N. Singh and S. B. Singh, ``A novel hybrid GWO-SCA approach for optimization problems," Eng. Sci. Technol. Int. J., vol. 20, no. 6, pp. 1586_1601, Dec. 2017.

[24] M. Mafarja and S. Abdullah, ``Investigating memetic algorithm in solving rough set attribute reduction,'' Int. J. Comput. Appl. Technol., vol. 48, no. 3, pp. 195_202, Oct. 2013.

[25] R. Azmi, B. Pishgoo, N. Norozi, M. Koohzadi, and F. Baesi, ``A hybrid GA and SA algorithms for feature selection in recognition of hand-printed Farsi characters,'' in Proc. IEEE Int. Conf. Intell. Comput. Intell. Syst., Oct. 2010, pp. 384_387.

[26] P. Moradi and M. Gholampour, ``A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy,'' Appl. Soft Comput., vol. 43, pp. 117_130, Jul. 2016.

[27] E.-G. Talbi, L. Jourdan, J. Garcia-Nieto, and E. Alba, ``Comparison of population based metaheuristics for feature selection: Application to microarray data classification,'' in Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl., Mar./Apr. 2008, pp. 45_52.

[28] Z. Yong, G. Dun-wei, and Z. Wan-qiu, ``Feature selection of unreliable data using an improved multi-objective PSO algorithm,'' Neurocomputing, vol. 171, pp. 1281_1290, Jan. 2016.

[29] J. Jona and N. Nagaveni, ``A hybrid swarm optimization approach for feature set reduction in digital mammograms,'' WSEAS Trans. Inf. Sci. Appl., vol. 9, pp. 340_349, 2012.

[30] M. E. Basiri and S. Nemati, ``A novel hybrid ACO-GA algorithm for text feature selection,'' in Proc. IEEE Congr. Evol. Comput., May 2009, pp. 2561_2568.

[31] R. S. Babatunde, S. O. Olabiyisi, and E. O. Omidiora, ``Feature dimensionality reduction using a dual level metaheuristic algorithm,'' Optimization, vol. 7, no. 1, pp. 49_52, 2014.

[32] J. B. Jona and N. Nagaveni, ``Ant-cuckoo colony optimization for feature selection in digital mammogram,'' Pakistan J. Biol. Sci., vol. 17, no. 2, pp. 266_271, 2014.

[33] E. ZorarpacI and S. A. Özel, ``A hybrid approach of differential evolution and artificial bee colony for feature selection,'' Expert Syst. Appl., vol. 62, pp. 91_103, Nov. 2016.

[34] M. Nekkaa and D. Boughaci, ``Hybrid harmony search combined with stochastic local search for feature selection,'' Neural Process. Lett., vol. 44, no. 1, pp. 199_220, 2016.

[35] M. M. Mafarja and S. Mirjalili, ``Hybrid whale optimization algorithm with simulated annealing for feature selection,'' Neurocomputing, vol. 260, pp. 302_312, Oct. 2017.

[36] H. M. Zawbaa, E. Emary, C. Grosan, and V. Snasel, ``Large dimensionality small-instance set feature selection: A hybrid bioinspired heuristic approach," Swarm Evol. Comput., vol. 42, pp. 29_42, Oct. 2018.

[37] Dorigo M, Stützle T. Ant colony optimization: Overview and recent advances. Handbook of Metaheuristics, 2010, 146, 227–263.

[38] Dorigo M. Optimization, Learning and Natural Algorithms. PhD thesis, Politecno di Milano, Milan, Italy, 1992.

[39] López-Ibáñez M, Stutzle T. Automatic configuration of multi-objective ant colony optimization algorithms. Lecture Notes in Computer Science, 2010, 6234, 95–106.

[40] Zamuda, A.; Brest, J.; Mezura-Montes, E. Structured Population Size Reduction Differential Evolution with Multiple Mutation Strategies on CEC 2013 real parameter optimization. In Proceedings of the IEEE Congress on Evolutionary Computation, Cancun, Mexico, 20–23 June 2013; pp. 1925–1931.

[41] H.-U. Schnitzler and E. K. V. Kalko, "Echolocation by insect-eating bats," Bio Science, vol. 51, no. 7, pp. 557–569, July 2001.

[42] Mirjalili, S., Mirjalili, S.M., Lewis, A.: Grey wolf optimization. Adv. Eng. Softw. **69**, 46–61 (2014)

[43] M. Seyedali and A. Lewis, "The Whale Optimization Algorithm," Adv. Eng. Soft., vol. 95, pp. 51-67, 2016.

[44] M. Chouksey, R.K. Jha, R. Sharma, A fast technique for image segmentation based on two meta-heuristic algorithms, Multimedia Tools Appl. (2020) 1–53.