Int. j. inf. tecnol.
https://doi.org/10.1007/s41870-024-02030-6

ORIGINAL RESEARCH

A proactive grey wolf optimization for improving bioinformatic systems with high dimensional data

Ali Hakem Alsaeedi[1,2] · Dhiah Al-Shammary[1] · Suha Mohammed Hadi[2] · Khandakar Ahmed[3] · Ayman Ibaida[3,5] · Nooruldeen AlKhazraji[4]

Received: 28 February 2024 / Accepted: 31 May 2024

Abstract This paper introduces a new methodology for optimization problems, combining the Grey Wolf Optimizer (GWO) with Simi-stochastic search processes. Intelligent optimizations represent an advanced approach in machine learning and computer applications, aiming to reduce the number of features used in the classification process. Optimizing bioinformatics datasets is crucial for information systems that classify data for intelligent tasks. The proposed A-Proactive Grey Wolf Optimization (A-GWO) solves stagnation in GWO by applying a dual search with a Simi-stochastic search. This target is achieved by distributing the population into two groups using a different search technique. The model's performance is evaluated using two benchmarks: the Evolutionary Computation Benchmark (CEC 2005) and seven popular biological datasets. A-GWO demonstrates highly improved efficiency in comparision to the original GWO and Particle Swarm Optimization (PSO). Specifically, it enhances exploration in 66% of CEC functions and achieves high accuracy in 70% of biological datasets.

Keywords Grey wolf optimization · Features selection · Bioinformatic systems · Data mining

✉ Ayman Ibaida
Ayman.Ibaida@vu.edu.au

Ali Hakem Alsaeedi
ali.alsaeedi@qu.edu.iq

Dhiah Al-Shammary
d.alshammary@qu.edu.iq

Suha Mohammed Hadi
dr.suhahadi@gmail.com

Khandakar Ahmed
Khandakar.Ahmed@vu.edu.au

Nooruldeen AlKhazraji
Khazraji@slb.com

1 College of Computer Science and Information Technology, University of Al-Qadisiyah, Diwaniyah, Iraq

2 Informatics Institute for Postgraduate Studies, Iraqi Commission for Computer and Informatics, Baghdad, Iraq

3 Intelligent Technology Innovation Lab, Victoria University, Melbourne, VIC, Australia

4 Digital Department of America's Land, Schlumberger Technology Corporation, Sugar Land, TX 77478, USA

5 Footscray, Australia

1 Introduction

Today, biomedical data is generated continuously from various biomedical equipment and experiments due to technological improvements in medical sciences. Optimizing bioinformation datasets plays a vital role in the information system for data classification and other intellectual functions [1]. Intelligent optimizers represent an advanced stage of machine learning and computer applications by reducing the total number of features for the process of classification function. Figure 1 illustrates the scenario of the intelligent bioinformatic system.Please confirm if the author names are presented accurately and in the correct sequence (given name, middle name/initial, family name). Author 1 Given name: [specify authors given name] Last name [specify authors last name]. Also, kindly confirm the details in the metadata are correct.Thank you they are correct.However, we kindly request to have the author Ayman Ibaida as a second corresponding author or as a main corresponding author if it is not possibleThe labels in figure (fig 10) is not readable. Please provide a new figure with legible labels in Vector EPS or tiff / jpeg format with 600 dpi resolution.

Published online: 31 August 2024

Thanks, we have uploaded new modified high quality jpeg format images for Fig.10

The metaheuristic algorithm has successfully optimized several bioinformatic systems approaches [2–6]. The concept of search in metaheuristic algorithms depends mainly on randomness. That is a systematic form to simulate the neuter behavior of animals or neutral phenomena. Developing search operations for metaheuristic algorithms depend on improving two factors exploration and exploitation[7, 35, 37, 38]. It is impossible to improve one factor and leave the other. Improving the exploration without exploitation will increase the scattering search and the instability of results. Also, improving the exploitation without exploration will increase the probability of falling into stagnation. Therefore, the robust optimization models have high exploration and exploitation [8]. The main challenge in searching with metaheuristic algorithms is that the algorithm potentially enters stagnation; therefore, more algorithm exploration is required to overcome the stagnation in the local optima. The Grey Wolf Optimization algorithm (GWO) is one of the recent metaheuristic algorithms that have been used in [9, 36, 39]. It has several characteristics that distinguish it from other algorithms. It is fast, easy to implement, and has a simple mathematics model [10]. To improve the search capabilities of the algorithm, we attributed exploration and exploitation by adding the dual search method when the algorithm is prone to stagnation during the search process. In dual search, the population is split into several groups, each searching individually for the best solution [11, 12]. The GWO has reduced computational costs by replacing time-consuming exhaustive searches. On the other hand, most of these approaches suffer from a lack of diversity, local optima, and an imbalance between the exploratory and exploitative capabilities of the algorithm. in other word, these problems can be a milestone on the way to the best solution in any research.Kindly check and confirm the edit made in the title.Thanks, we checked and a pleasure to confirm the edit made in the title.Journal instruction requires a city for affiliations; however, these are missing in affiliation [1]. Please verify if the provided city is correct and amend if necessary.City has been corrected.please, also I mistakenly added a new affiliation to the author Ayman Ibaida and failed to remove it. Can you please remove it

Many problems can arise when searching for the best subset of features [13]. Random search, depth search, breadth search, or hybrid approaches are some of the search methods

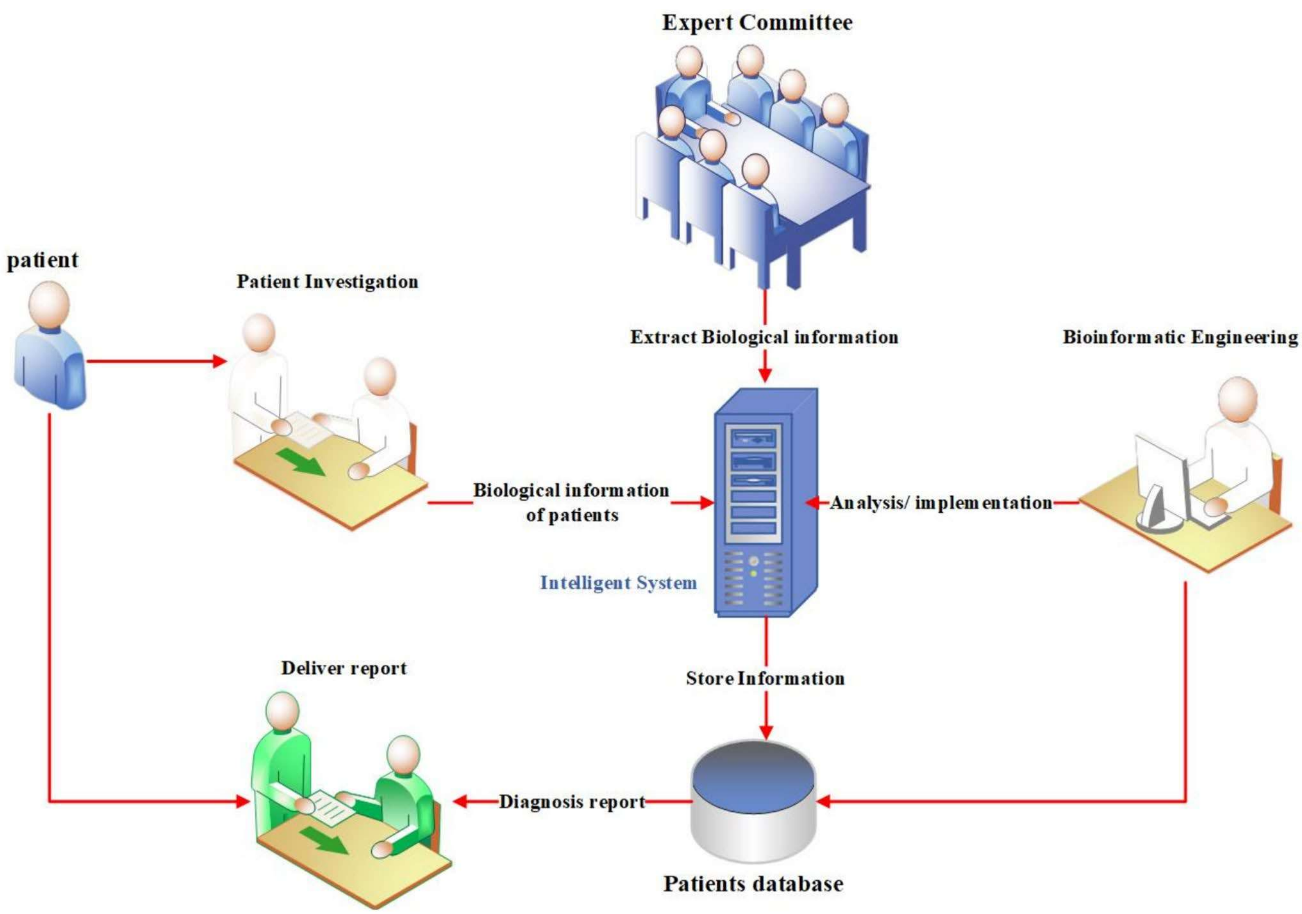


Fig. 1 Scenario of intelligent bioinformatic system

for determining relevant subsets of features. Breadth-first search is not suitable for large datasets because it is impossible to select the best subset of components in a 2d-solution, where d is the size of the features. In optimization, the algorithm is a big part of how quickly reasonable solutions can be found.

1.1 Motivation

Generally, The devices that process data has grown at an exponential rate, so computer servers that get a lot of data may have some features that aren't ideal [8]. The noise, device errors, or reputation in the elements potentially increase the number of unrelated features in the data objects. Data processing can be improved by extracting and selecting important data types and features. In addition, machine learning algorithms have significant problems with high-dimensional data that contain many undesired features [14, 15]. The improvement in feature selection. Improving the feature selection improves the systems' performance and reduces the complexities of time and storage. The successes of GWO [16] and the importance of feature selection have motivated us to develop a new version of GWO (A-GWO) for optimizing features for bioinformatic systems and machine learning approaches. We found that original version of GWO suffers from several drawbacks; some of them are:

1. Stuck in local optima: The GWO quickly gets into local optima, so the algorithm is limited in the advanced stages of the search [17]. Therefore, this negatively affects the algorithm's performance and causes it to fail to help change the path for the better. This hypocritical behavior of the algorithm results from the dependence on a single induction method and the presence of various solutions proposed by the algorithm [8].
2. Low convergence: GWO's search process suffers from a lack of convergence in climbing towards the global solution or the best possible local optima [18]. The algorithm updates the proposed solutions based on the previous best position of the grey wolf alpha, beta, and delta.
3. Imperarical results: Generally, GWO yields highly variant results, especially for high-dimensional objective functions. The instability of the result after each iteration detracts from the algorithm's reality.

1.2 Contribution

The contribution of this paper lies in proposing search framework that transfers or substantially helps the algorithm overcome the stagnation interval during the search process. The proposed method switches the search engine to randomness in the case of stagnation. It returns to the ongoing search in the systematic process (regular search) to contribute to the results' stability. In the proposed model, if the GWO does not achieve any increase in progress toward a global solution, it will increase stochastic search to improve the exploration of a search engine. Stochasticity leads to scattering in search processing. Therefore, the search procedure must rely on the current local optima as the central part of the optimization equation to advance to the global

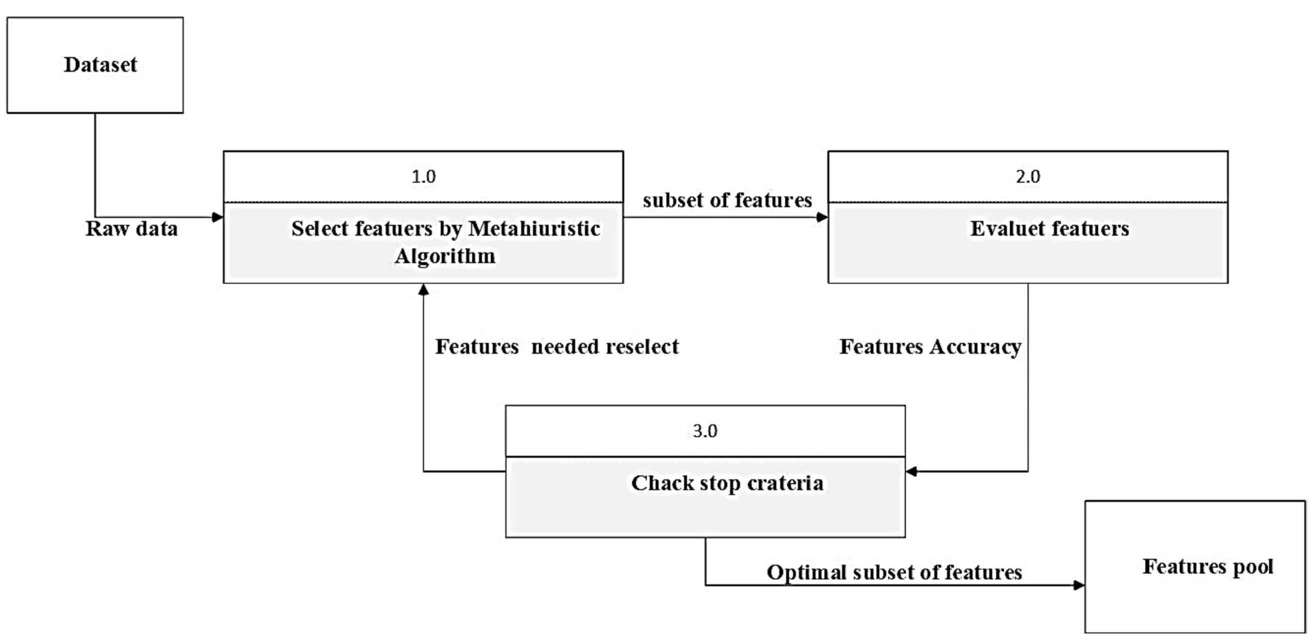


Fig. 2 Feature selection by wrapper model and optimization algorithm

Fig. 3 A feature selection example with a threshold is 0.5, The dataset's features that correspond to the index of selected features will be chosen, while the others will be omitted

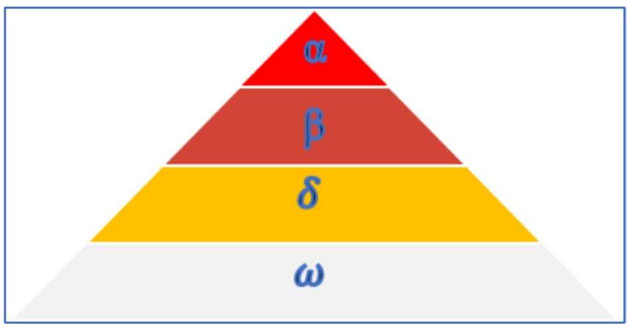


Fig. 4 Hierarchy of grey wolf (dominance decreases from the top down)

solution. Two goals are achieved by gradually increasing the stochastic search: improving the exploration and reducing the scatter of the search process. However, the proposed A-GWO provided several promising achievements:

1. Low stuck in local optima: The proposed model (A-GWO) margins are two techniques that give a proactive way during search operations. Therefore, the probability of falling into stagnation is low depending on our tests.
2. High convergence: The proposed A-GWO's search process converges highly during search progress. It is based on two fundamentally different methods of behavior and techniques used. Diversity in generating solutions greatly enhances search convergence [19]. The presence of this feature in the proposed method enables the algorithm to increase the convergence rate.
3. Stability results: The experiment results show that the proposed A-GWO gives pretty close results when run several times. The results of the proposed method make it seem like the proposed techniques are very stable and reliable.

1.3 Evaluation strategies

The proposed model is tested and evaluated by Congress on Evolutionary Computation 2005 (CEC'2005) as a benchmark Function. The proposed A-GWO compared with the original GWO and Practical Swarm Optimization (PSO) in feature selection and finding optimal solutions in CEC benchmark Function. Moreover proposed model was compared with recent features selection models.

1.4 Paper organization

The rest of the paper is organizaed into 6 sections (2.7). Section 2 describes related work. Section 3 illustrates Wrapper Feature selection models. Section 4 states the Grey Wolf Optimization (GOW) model. Section 5 explains the Proposed A proactive Grey Wolf Optimization (A- GOW) model. Dataset and result discussion are discussed in Sect. 6. Finally, conclussions are discussed in Sect. 7.

2 Related works

Several studies have proposed different optimization models to develop the metaheuristic model. Most of the methods are based on the combination of two algorithms. However, both algorithms potentially suffer from technical problems in search. Suppose the fundamental issues of the algorithm are not addressed, such as discovering and exploiting optimal solutions, dealing with slack, and reducing the dispersion of results. In that case, good results will not be obtained.

Chakraborty et al. [19] proposed a hybrid machine learning based on an ensemble technique with A Grey Wolf Optimization(GWO) for feature selection. They select features in the proposed model based on a binary representation. The binary model in featuer selection has limited search space by 0 or 1; therefore, It's easy to get into the stagnation area [20].

In [21], The authors propose a new hybrid filter and wrapper approach that uses Minimum Redundancy Maximum Relevancy (rMRMR) as a filtering approach to select the best-ranked genes. The modified Grey Wolf Optimizer (MGWO) is used as a wrapper approach to search for other small groups of genes. The modified Grey Wolf Optimizer (MGWO) is used as a wrapper approach to search for different small groups of genes. To maintain diversity in searching for promising genes in the search space, MGWO incorporates TRIZ principles. After each GWO improvement loop, the main task of these Triz operators is to increase the diversity of the whole updated GWO population iteratively. Thus, this can be considered a hybrid mechanism with creative operators. Running the algorithm iteratively does not mean improving the algorithm technically. This work increases the time complexity of the algorithm. The technique depends entirely on the number of runs.

Wen et al. [22] proposed a new position update equation presented by using a random individual in the population to guide the search for unique candidate individuals. They

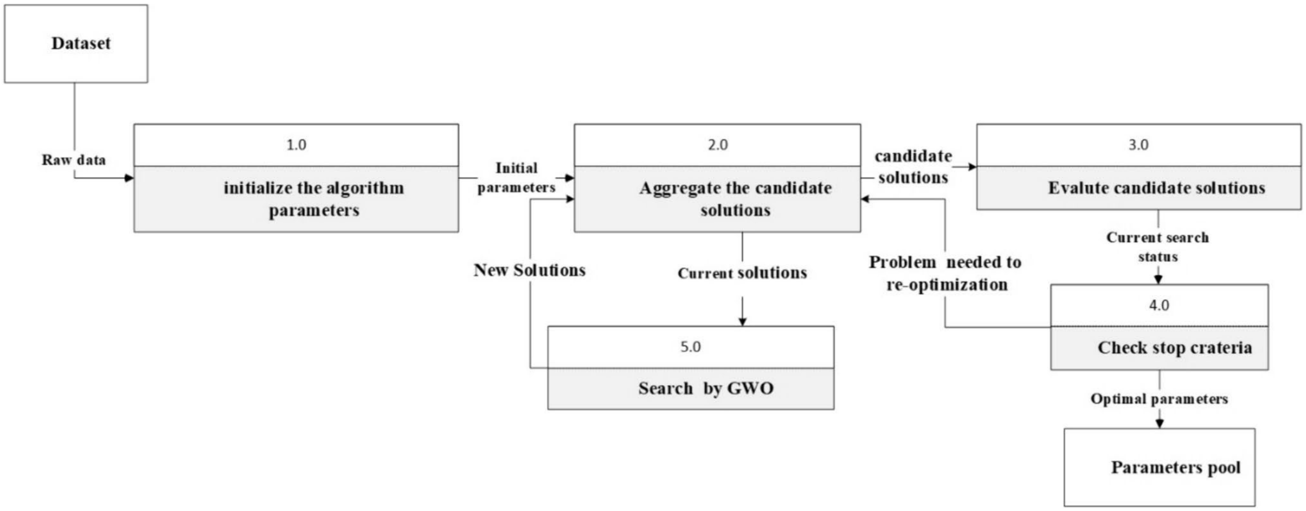


Fig. 5 Grey wolf optimization (GWO)

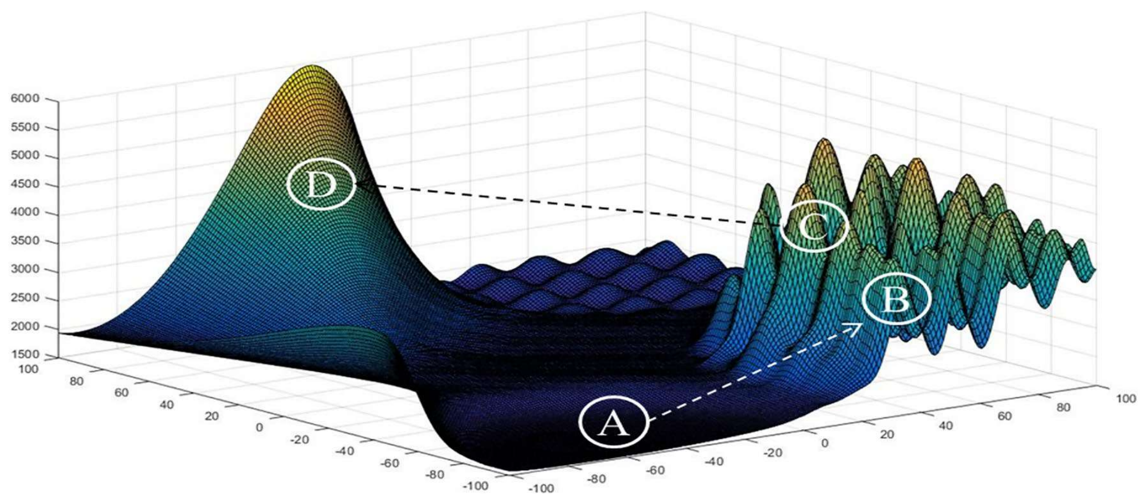


Fig. 6 Example of search processes of proposed A-GWO

introduced a strategy with nonlinear control parameters to fully utilize the GWO algorithm and balance exploration and exploitation. The proposed model updates the position by adding two random variables to the final equation of GWO that responds to generate new candidate solutions. Adding these variables to the basic form of the algorithm converts the algorithm ultimately into a simi-stochastic. Therefore, the proposed model potentially suffers from scattering search and instability results.

Alkafagi et al. [14] proposed a new framework to optimize search progress in PSO by restarting the entire population when the PSO algorithm stagnates. The authors highlighted the population baseline as the cause of the decline in research quality. The proactive model would reset the population when the system detects stagnation in the search process. This method increases exploration and decreases the likelihood that the local optimum will stagnate. Researching from the beginning gives a logical reason that the beginning will provide better solutions than the current situation.

Al-Shammary et al. [8] proposed a new version of Particle Swarm Optimization (PSO) to search for the best possible solution to a high-complexity problem. The model randomly divides the population into two groups to foster exploration and prevent stagnation. Dual search helps algorithms search for high-dimensional issues. The random searches yield inconsistent and untrustworthy results.

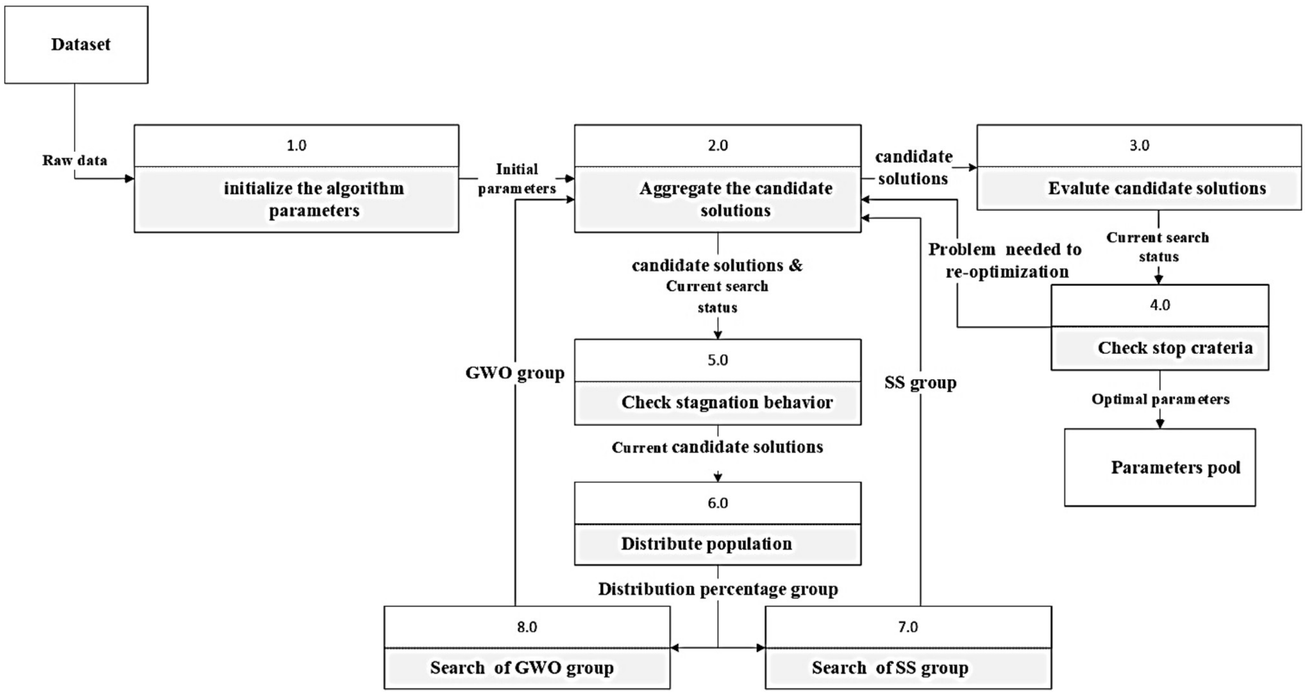


Fig. 7 Proposed a proactive grey wolf optimization (A-GWO)

3 Wrapper feature selection models

The wrapper features selection models are excellent for selecting high correlation features from high-dimensional data [23]. It picks features by trial and error and may choose a different group of features after each cycle. The wrapper model randomly picks characteristics; therefore, the system relies on unpredictable and variant results [24]. Metaheuristics choose wrapper model features. This study uses the proposed A-GWO to represent the features model in the wrapper model. Figure 2 shows how the wrapper model uses the optimization algorithm to choose the best features.

Figure 3 illustrates the mechnisim of feature selection scenario of the proposed approach.

4 Grey wolf optimization (GOW) model

Grey Wolf Optimization (GOW) is a metaheuristic proposed by Mirjalili et al. [9], inspired by the haunting of Grey Wolves. Technically, it consists of four groups in its community divided by their capabilities. These four groups are called Alpha, Beta, Delta, and Omega. These are the solutions of the GWO, with Alpha, Beta, and Delta being the first, second, and third best solutions, respectively. The Omega group is the weakest solution in the GWO. Figure 4 shows the hierarchy dominant strict society of wolves from strong (best) to worst.

The mathematical representation of GWO is shown by the Eqs. 1..4 [9]:

$$X^i_{t+1} = \left| X^i_t - C^i.A^i(X^i_p - X^i_t) \right| \tag{1}$$

where t is the current iteration, A and C are coefficient vectors, X^i_p is the position vector of the prey in ith dimension, and X^i_t is positions of prey and grey wolves in ith dimension. Equations 2 and 3 calculate the coefficient A, C, respectively.

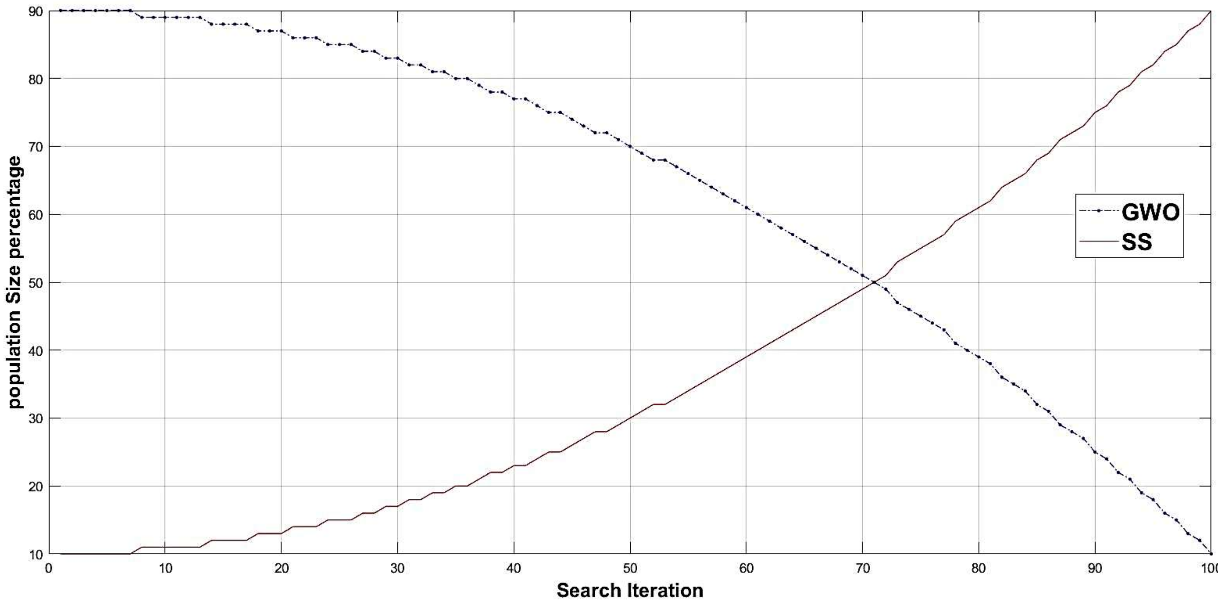
$$A^i = |2a^i.r_1 - a^i| \tag{2}$$

$$C^i = 2r_2 \tag{3}$$

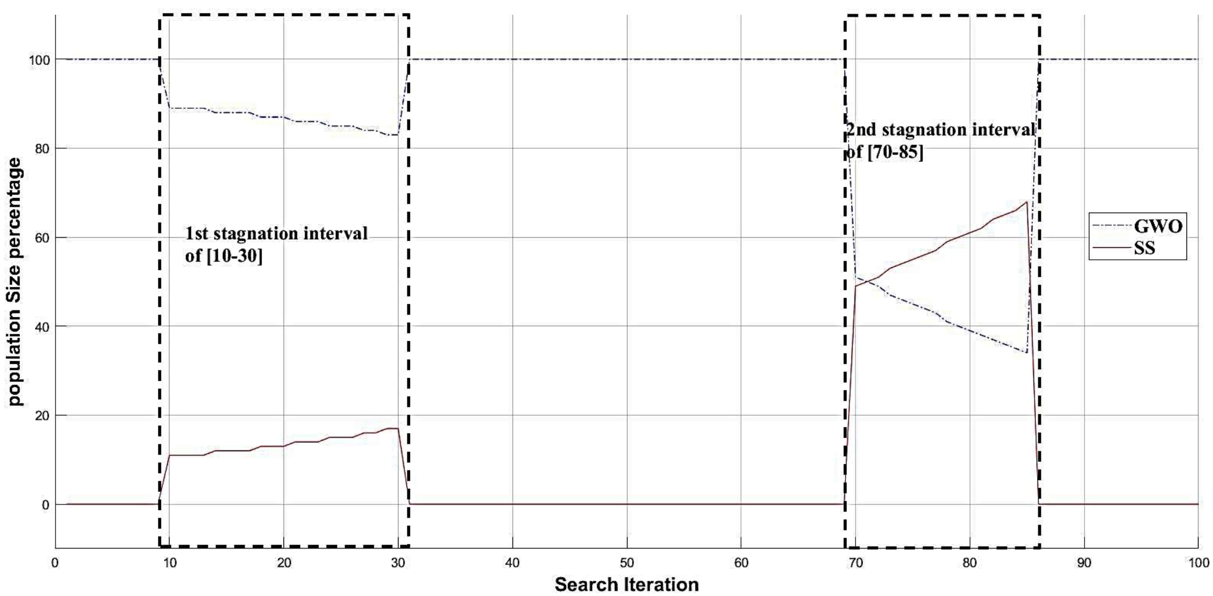
r_1, r_2 are random vectors within intervals [0, 1], and a is linearly decreasing number from 2 to 0 at each iteration: Eq. 4 calculates a:

$$a = 2 - \frac{2\ t}{iter_{max}} \tag{4}$$

where $iter_{max}$ determine the maximum number of iterations. Position of wolves $[X, Y]$ changes concerning prey position $[X^*, Y^*]$.



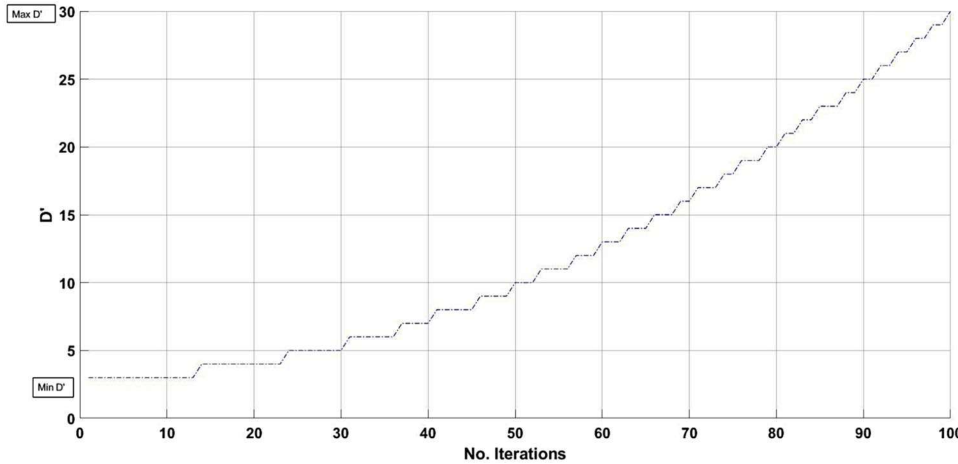
a- The first scenario (the proposed A-GWO senses stagnation in the search process)



b- the second scenario (the proposed A-GWO stagnation during the [10-30] and [70-85])

Fig. 8 Possible scenarios for a portion of each group of A-WO (GWO and SS)

Fig. 9 Number of the updating dimensions in the 2nd group (SS_g) (where $iter_{maz}$=100, $Max_{D\prime}$ =30, $Max_{D\prime}$=3)



Updating formulas for the three best wolves ($AlphaX_1^i, BetaX_2^i, andDeltaX_3^i$) are as shown in Eqs. 5, 6, and 7:

$$X_{1,t}^i = X_\alpha^i - A_1^i.D_\alpha^i, D_\alpha^i = \left|C_\alpha^i.X_\alpha^i - X^i\right| \tag{5}$$

$$X_{2,t}^i = X_\beta^i - A_2^i.D_\beta^i, D_\beta^i = |C_2^i.X_\beta^i - X\prime^i| \tag{6}$$

$$X_{3,t}^i = X_\delta^i - A_2^i.D_\delta^i, D_\delta^i = |C_3^i.X_\delta^i - X^i| \tag{7}$$

Equation 8 calculates the final position of candidate solution X

$$X_{t+1}^i = \frac{X_{1,t}^i + X_{2,t}^i + X_{2,t}^i}{3} \tag{8}$$

Figure 5 illustrates the search methodology of the GWO algorithm.

5 Proposed a proactive grey wolf optimization (A-GOW) model

Evolutionary Algorithms (EAs) generally consist of three steps: initialization, update position, and termination. The development in A- GOW is through adding a new framework search methodology, which is the sense of the stagnation and distribution of the population into two search groups. Each group searches with a different technique. This step aims to enhance the exploration of the proposed model. The proposed system handles the scattering in the searching process by maintaining the limited number of wolf dimensions without modifying the earlier search. The wolf's dimensions are subject to updating gradually during increases as the research stages progress. Figure 6 illustrates the scenario of search progress in the proposed A-GWO. This example shows three local optima (A, B, and C) and one global solution (D). Assume the algorithm only in location B suffers from stagnation, so it applied dual search only in location B to overcome stuck ($B \rightarrow C$) and return to principles of GWO to jump from location ($C \rightarrow D$).

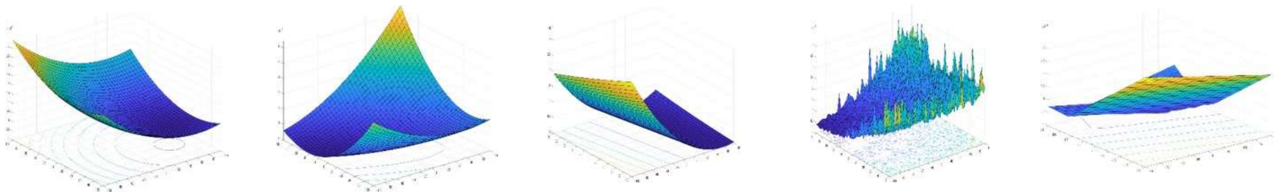
Figure 7 illustrates the scenario of the searching process of the proposed A- GOW.

5.1 Initialization

The proposed model's initial parameters are set in this step (process 1.0). The percentage of each group's EGWO is used to determine the proportion of each group's tendency in the final population. Minimum and maximum dimensions are subject to updating in the second group. Moreover, the essential parameters of the GWO are also determined to create an integrated environment to work as an algorithm.

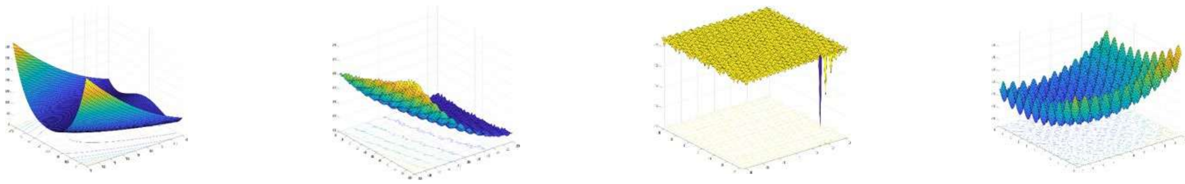
5.2 Aggregate the candidate solutions and system behavior

This phase in the proposed system consists of four processes (2.0, 3.0, 4.0, and 5.0). The general purpose of those steps is to aggregate and check system behavior. After aggregating the candidate solutions (process 2.0), the system has to evaluate solutions (process 3.0) and check the stop criteria of the search process(process 4.0). If the system satisfies the stop criteria, return the available best solution as an outcome of the search process; otherwise, the system goes on searching for other accepted solutions. The proposed system

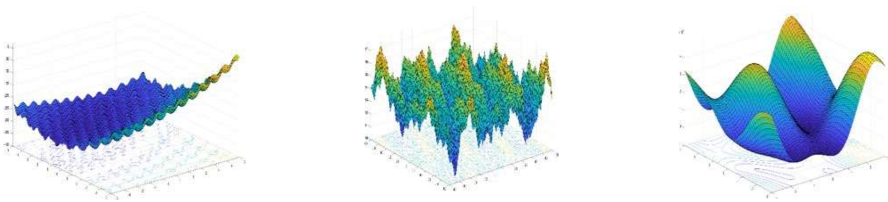


F1 F2 F3 F4 F5

a- 2-D versions of unimodal benchmark functions.

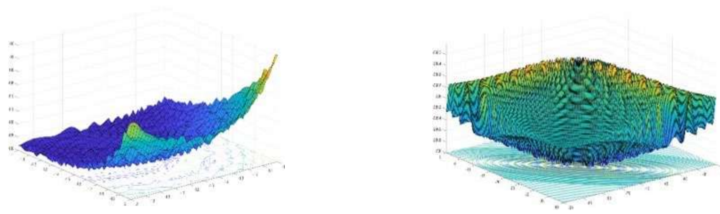


F6 F7 F8 F9



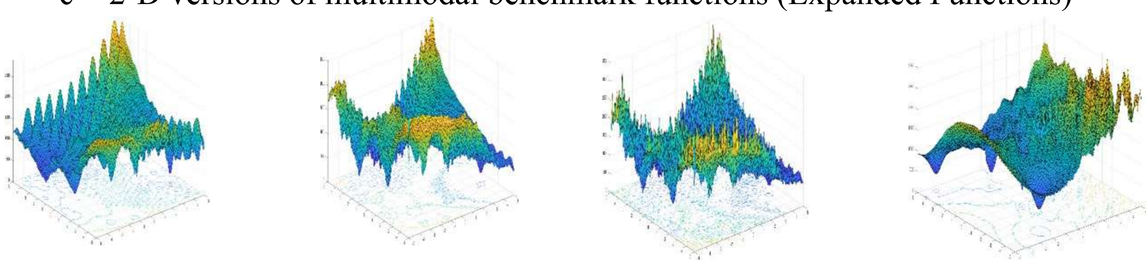
F10 F11 F12

b- 2-D versions of multimodal benchmark functions (Basic Functions).



F13 F14

c- 2-D versions of multimodal benchmark functions (Expanded Functions)



F15 F16 F17 F18

d- 2-D versions of multimodal benchmark functions(Hybrid Composition Functions)

Fig. 10: 2D versions of the CEC'05 benchmark functions

Table 1 Details of biological datasets

Data set	Size			Data field
	Samples	Features	Classes	
ALLAML	72	7129	2	Bioinformatic data
CLL_SUB_111	111	11,340	3	
GLI_85	85	22,283	2	
Leukemia	72	7070	2	
Lung	203	3312	5	
lung_discrete	73	325	7	
TOX_171	171	5748	4	

Table 2 Parameters of A-GWO, GWO, and PSO

Parameter	A-GWO	GWO	PSO
Population size	30×30	30×30	30×30
Upper/lower boundary	±100	±100	±100
P_{max}	0.9	–	–
P_{min}	0.1	–	–
$\max_{D'}$	1.0	–	–
$\min_{D'}$	0.4	–	–
Max iterations	50	50	50
Vmax	–	–	6
wMax	–	–	0.9
wMin	–	–	0.2
c1, c2	–	–	2

applied the dual search after failing to figure out a new best candidate solution during n-iterations (process 5.0).

5.3 Population distributed

The A- GOW is inspired by the principle of sharing life matters between two partners. Mathematically, the population is divided into two groups with inverse relationships in terms of population size in each group – Search Agents number of each group. The first group GOWg consists of some wolves searching based on the principle of grey wolf optimization techniques. While the second group SSg, the wolfs searching in simi-stochastic(as shown in Fig. 7). They are selected randomly from the entire of wolves population. Equations 9 and 10 calculate the number of wolves in each group [8].

$$P_1 = Pop_{size} - round\left(\left(P_{max} - \left(\frac{t}{iter_{maz}}\right)^2 (P_{max} - P_{min})\right) Pop_{size}\right) \tag{9}$$

$$P_2 = Pop_{size} - P_1 \tag{10}$$

where P_1, P_2 are the size of GOW_g and SS_g, respectively. P_{max}, P_{min} are the percentage of GOW_g and SS_g, while the pop_{size} is the value of the total population of a new generation, and t is the current iteration.

The split population is only if the algorithm falls in local optima -using a dual search process. If the search progress continues, the algorithm searches using the traditional GWO principles method. Therefore, three scenarios are expected: the stagnation period is continuous, the stagnation period is not continuous, or the stagnation does not happen during search progress. In the first scenario, the proposed system increases the random search if the algorithm senses stagnation in the search process. It uses this feature to increase exploration and try to jump to the best local optima from the current stage to improve the search method. In the second scenario, suppose the system fells into stagnation during the [10–30] and [70–85] in the search iterations only in these intervals applied dual search operation. Figure 8 illustrates the portion of each group of A-WO (GWO and SS).

5.4 Search methodology

The search methodology updates position step represent in processes (7.0 and 8.0). As long as the algorithm does not suffer from stagnation, the search is based on the GWO algorithm. The population is split into two groups if the condition does not improve the search result during the n iterations. Each group searches in parallel with the second group in a double search method. The dual search is abandoned when the stagnation of the algorithm is eliminated. The proposed model used updating to specific dimensions for the second group to get a greater chance of reducing the dispersion by the proposed algorithm. So that this modernization gradually increases because the nature of metaheuristic algorithms in exploration factor decreases in the advanced stages of research. According to [8, 25–27], the exploration alone does not satisfy the preferred requirements; therefore, the exploitation is necessary to support the stability and orient the algorithm to the correct local optima. Equation 11 calculates the size of diminution at each stagnation period for the proposed model's second group (SS) [28].

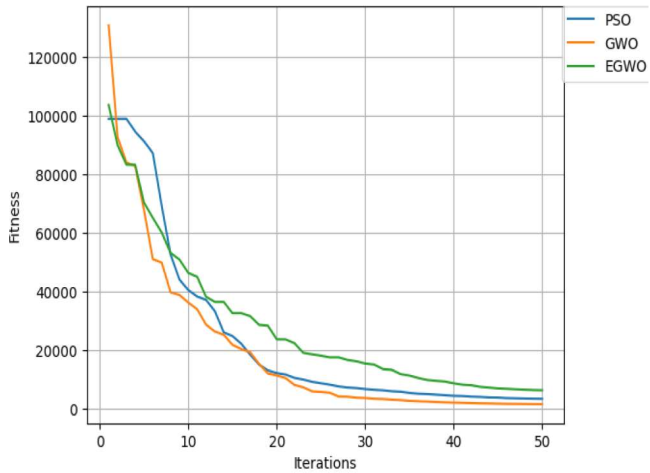
$$SS_{D'} =\sim \left(\left(M_{D'} + \left(\frac{t}{iter_{maz}}\right)^2 * (max_{D'} - min_{D'})\right) * D\right) \tag{11}$$

where: $\max_{D'}$, $\min_{D'}$ are maximum and minimum diminution of a new wolf in the SS group. D is the problem dimension, and D' is a new dimension applied on their simi-stochastic operation. Figure 9 illustrates the relation between the number of mutation gens in SS and search iteration.

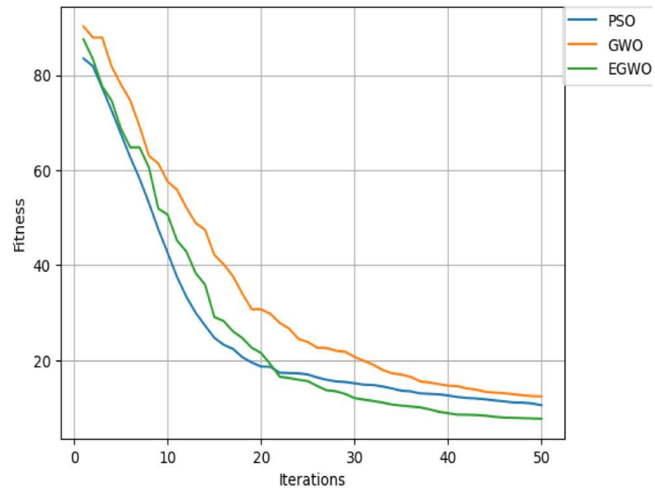
Equation 12 calculates the new position of candidate solutions belonging to the SS group

Table 3 1 Comparison result of PSO, GWO, and A-GWO of 18 functions of CEC'05 over 30 independent runs

Obj. function	PSO [8]					GWO [21]					A-GWO				
	Best	Worst	Average	Std	time	Best	worst	Average	std	time	Best	Worst	Average	std	time
F1	58.531	79.908	67.87	10.941	0.83	2.417	7.864	**5.77**	2.935	0.81	10.257	12.047	11.21	1.057	0.81
F2	16.463	36.576	26.62	10.058	0.77	0.520	0.781	0.65	0.131	0.85	0.856	1.676	1.19	0.129	0.84
F3	3209.992	3672.547	3420.85	233.966	0.97	1236.388	2157.346	1600.2	489.973	1.06	2789.960	8287.032	6349.67	3086.822	1.03
F4	8.843	13.917	10.54	2.927	0.8	10.925	14.526	12.39	1.890	0.81	6.615	8.482	7.69	0.965	0.83
F5	5042.495	30,945.187	20,355.74	13,582.085	0.81	389.080	663.239	521.55	137.312	0.92	131.580	318.296	241.93	97.887	0.85
F6	43.408	97.159	67.15	27.419	0.77	11.363	14.641	12.51	1.847	0.83	12.469	19.058	16.02	4.098	0.83
F7	5.386	6.877	6.37	0.849	0.79	0.021	0.063	0.04	0.022	0.85	0.037	0.084	0.06	0.021	0.84
F8	– 2940.22	– 2127.55	– 2623	434.959	0.8	– 5267.03	– 2684.03	– 3712	1369.611	0.82	– 6937.11	– 3437.35	– 5617	1201.941	0.81
F9	244.285	306.418	268.40	33.320	0.94	50.494	223.516	132.52	86.858	0.82	89.787	103.261	97.02	6.792	0.85
F10	4.652	5.580	5.03	0.488	0.81	1.000	2.981	2.15	1.029	0.85	2.328	3.244	2.87	0.479	0.89
F11	193.245	208.179	201.01	7.484	0.82	1.037	1.084	1.07	0.027	0.87	1.043	1.051	**1.05**	0.004	0.85
F12	2.589	3.673	3.01	0.579	0.86	2.415	5.473	4.04	1.538	0.88	2.012	2.439	2.28	0.237	0.94
F13	17.601	99.744	48.37	44.778	0.84	6.619	25.299	13.59	10.204	0.94	6.786	12.339	8.66	3.183	0.95
F14	0.998	8.841	5.26	3.964	0.66	1.992	8.841	5.91	3.530	0.68	2.982	7.874	4.94	2.587	0.72
F15	0.001	0.020	0.01	0.011	0.14	0.001	0.020	0.01	0.011	0.14	0	0	0	0.000	0.15
F16	– 1.032	– 1.032	– 1.03	0.000	0.06	– 1.032	– 1.032	– 1.032	0.000	0.06	– 1.032	– 1.032	– 1.032	0.000	0.06
F17	0.398	0.398	0.398	0.000	0.06	0.398	0.401	0.401	0.000	0.06	0.398	0.398	0.398	0.000	0.06
F18	3.000	3.000	3.00	0.000	0.06	3.000	3.030	3.020	0.002	0.07	3.002	3.010	3.005	0.001	0.07

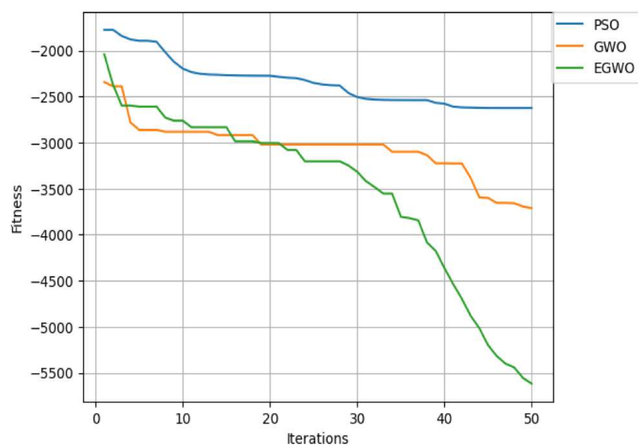


a- F3

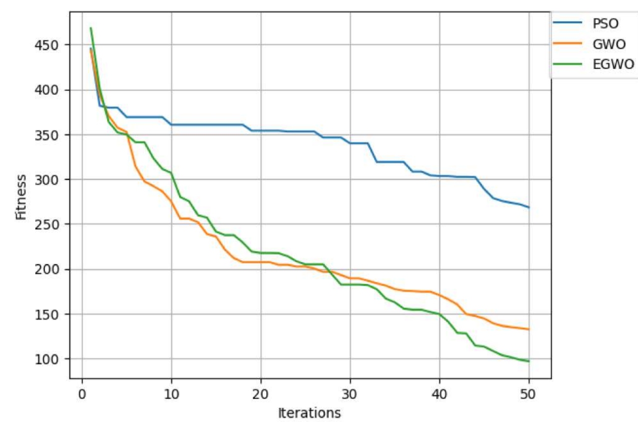


b- F4

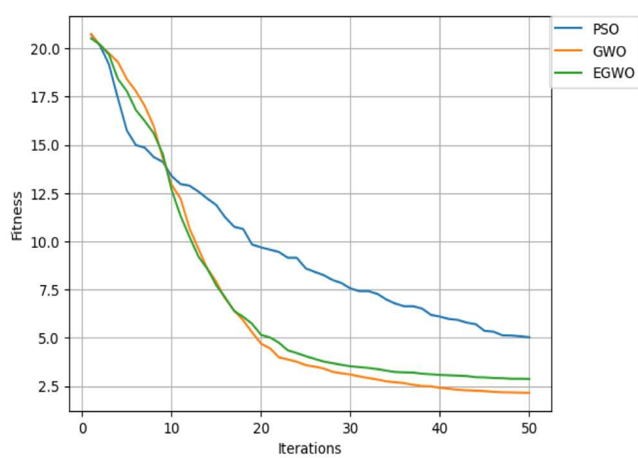
Fig. 11 Comparison of PSO [8], GWO [21], and A-GWO on Functions (F1, F4)



F8

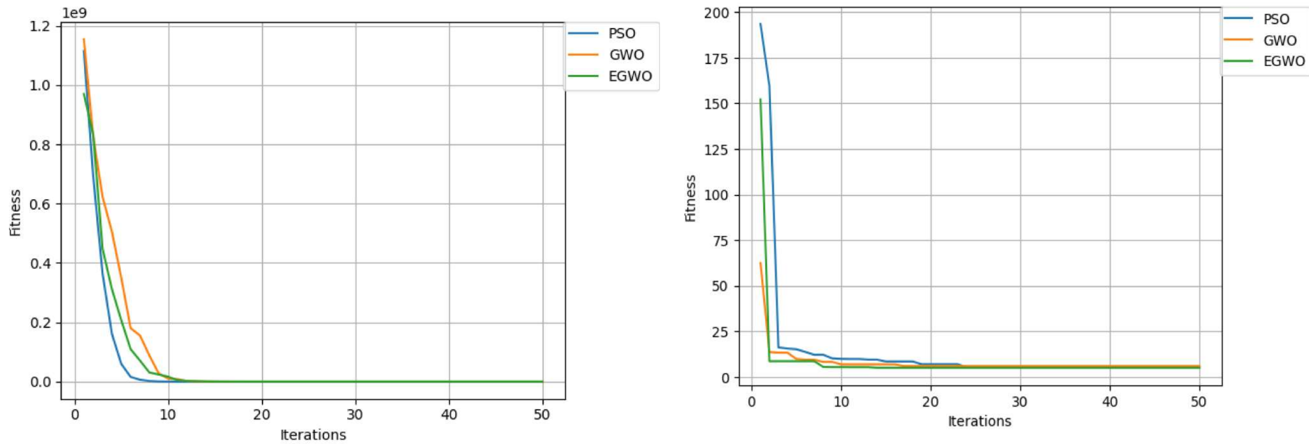


F9



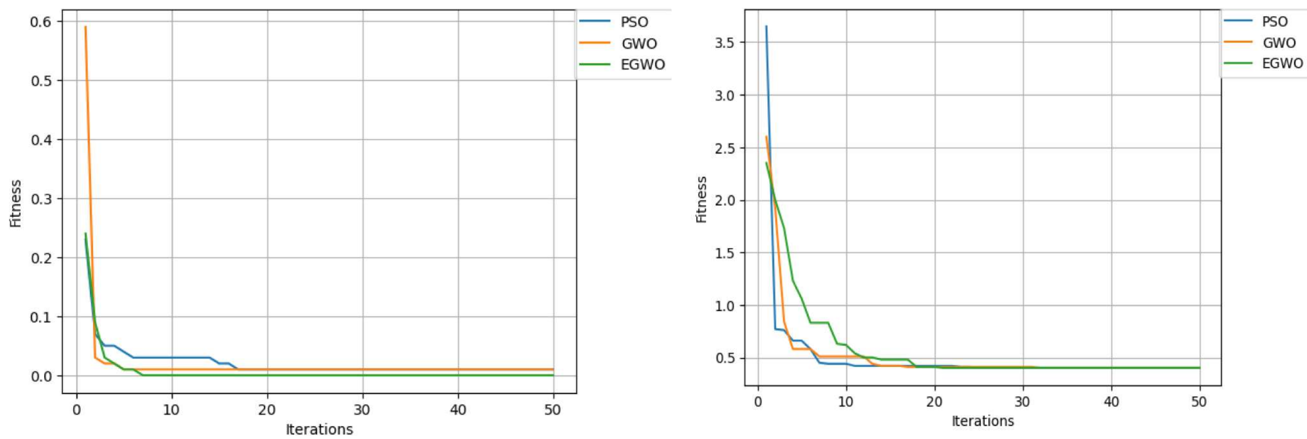
F10

Fig. 12 Comparison of PSO [8], GWO [21], and A-GWO on Functions (F8, F9, and F10)



F13 **F14**

Fig. 13 Comparison of PSO [8], GWO [21], and A-GWO on Functions (F13, F14)



F15 **F17**

Fig. 14 Comparison of PSO [8], GWO [21], and A-GWO on Functions (F15, F17)

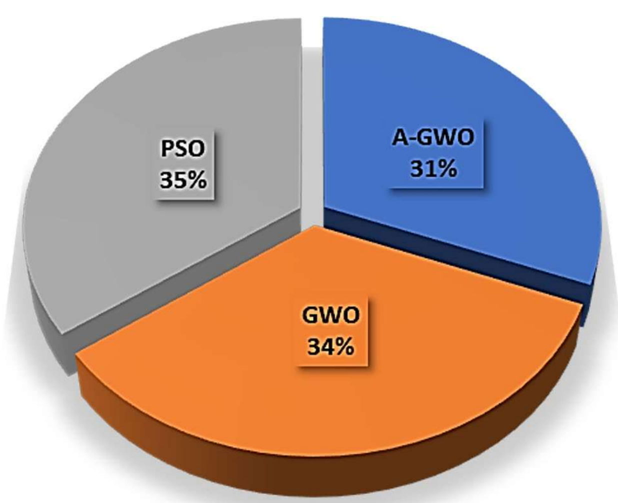


Fig. 15 Comparison of PSO [8], GWO [21], and A-GWO in execution time

$$X(t+1) = \frac{c_1}{2} + \left(1 - X_t\right)best_t + c_2 best_t \tag{12}$$

where: c_1, c_2 the random value within the interval [-1,1].

5.5 Algorithmic methodology

The proposed model is iluustrated algorithmically in Algorithm 1. This algorithm illustrates the technical requirements to implement the new proposed version of Grey Wolf Optimization like buffers and variables. Moreover, the procedural sequence of the proposed model processes is calrified and explaination comments are added.

6 Dataset and result discussion

The evaluation of the new metaheuristic model is necessary to analyse its exploration and stability and to compare its performance with recent methodologies. We use the Evolutionary Computation Benchmark (CEC) (the CEC 2005 is used as a benchmark function) to test the exploration of the proposed (A-GWO) model [29]. It is suitable for evaluating metaheuristic performance [26, 30, 31]. In features selection, used bioinformatics datasets with different dimensions are collected from the website (https://jundongl.github.io/scikit-feature/datasets.html).

Algorithm 1 A-GWO

Input: Problem search space (p) // Represents the initial population
Output: Optimal population P_{best} // Solution with the best fitness value
1. Initialization:
   - GWO parameters (including inertia weight, acceleration coefficients)
   - Stopping criteria (maximum iterations, stagnation threshold)
2. Generate Population (P):
   - Initialize a population of candidate solutions.
3. Evaluate Fitness:
   - Calculate the objective function for each particle in P.
4. Check Stopping Criteria:
   - If the stopping criteria are met (e.g., maximum iterations reached), terminate the algorithm.
   - Return the best solution found in the last generation of P as the output.
5. Check Stagnation:
   - If the search progress stagnates after n iterations (convergence slows down):
   - Divide P into two groups based on equations 9 and 10 (details likely explained elsewhere).
   - Group 1 (SS): Random search operations.
   - Group 2 (GWO): Solutions requiring further exploration.
   - Adjust group sizes based on stagnation level and search progress.
   - Otherwise, continue with the standard GWO update rule (regular GWO).
6. Repeat:
- Go back to step 3 and continue iterating until the optimal solution is found.

Table 4 Comparison of accuracy and percentage of feature selection of standalone RF, GWO, PSO, and A-GWO

Dataset name	Standalone RF	PSO [8]		GWO [21]		A-GWO	
	Accuracy	Accuracy	Percentage of features %	Accuracy	Percentage of features %	Accuracy	Percentage of features %
ALLAML	95.45	97.08	21.5	97.61	21	100	16
CLL_SUB_111	82.35	82.35	12	85.29	16	88.24	26
GLI_85	82.62	88	31	92	26	96.15	22
Leukemia	65.52	62.52	19	68.97	16	65.52	11
Lung	81.96	81.97	15	81.97	15	81.97	15
Lung_discrete	68.18	77.37	20	72.73	16	81.82	16
TOX_171	69.32	72	27	71.15	20	75	14

Table 5 Compare the proposed A-GWO with other studies

References	ALLAML	CLL_SUB_111	GLI_85	TOX_171
[32]	98.91	–	–	–
[33]	98.86	85.95	95.02	80.98
[34]	94.57	81.89	–	81.90
Proposed A-GWO	100	88.24	96.15	75

6.1 Congress on evolutionary computation benchmark functions (CEC)

The CEC'05 is used to test the exploration of evolutionary algorithms. It comprises eighteen benchmark functions divided into two categories (Unimodal and multimodal). These benchmark functions consisted of the shifted, rotated, extended, and combined variants of the classical methods, with the most significant complexity among the current benchmark functions [9]. The CEC functions have many local optima and help examine exploration and local optima avoidance of algorithms. Figure 10 illustrates the 2D versions of the benchmark functions used.

6.2 Biological datasets

Molecular Biology Datasets are data used for genotypic and phenotypic studies. This paper analyzes and implements the proposed model on high-dimensional biological data. Sven of molecular biology datasets is used for testing the proposed framework for feature selection. Table 1 illustrates the details of the biological datasets used in this paper.

6.3 Simulation properties and algorithm parameters

The algorithms A-GWO, GWO, and PSO, will be searching for minimum value in the search space of benchmark functions; therefore, each algorithm requires the minimum average. Table 2 shows the main Parameters of algorithms: proposed A-GWO, GWO, and PSO:

The A-GWO, GWO, and PSO algorithms were coded in Python 3.9, and all of the experiments were performed on a computer with intel(R) Core (TM) i7-8750H CPU @ 2.20GHz and 16.00 GB RAM in the Windows 11 environment.

6.4 Results discussion

This section illustrates and discusses evaluation strategies, describes experiments generally, experiments for feature selection methods, and comparisons with other studies.

6.4.1 Exploration experiment

In this section, the GWO algorithm is benchmarked on 18 benchmark functions. The A-GWO algorithm was run 30 times over each benchmark function of the CEC'05. We analyze empirical statistical results depending on the extracted average and standard deviation of run 30 times, which are shown in Table 3.

The standard deviation (Std.) is calculated for each benchmark function over 30 runs. When the value of Std. is higher, the standard deviation values reflect that the results differ significantly from one run to the next and are impractical. In this scenario, the higher standard deviation values reflect a limitation in the algorithm's performance. That means the results differ significantly from run to run and are not stable. The algorithms in this benchmark function search for minimum value; therefore, the algorithm that archives minimum value performs better. According to Table 3, A-GWO can provide very competitive results.

The results on the Unimodal Functions (F1 to F5) results show that GWO is best in functions (F1, F2, and F3) according to an average final of 30 iterations, while the A-GWO gets a better result in functions (F4 and F5).

Although the GWO algorithm obtained favorable results for up to 60% of the total unimodal benchmark functions, it failed to stabilize the results, getting a relatively high standard deviation for most functions. However, the A-GWO algorithm received the best results at 40%. Moreover, the algorithm required less time to execute. Therefore, it is better than the rest of the algorithms, so it can be said that the algorithm is the best among those that run over the unimodal benchmark functions. Figure 11 shows the performance of algorithms PSO, GWO and A-GWO.

The result in multimodal functions (basic functions F7-F12) shows that A-GWO achieves optimal performance in all functions according to time and search stability. In addition, the algorithm gets the best minimum solutions in 67% of multimodal-basic functions benchmark functions. Those benchmark functions have several local optimums, so the optimization algorithm requires high exploration to satisfy relevant results. Because the proposed algorithm explored a lot, it could get around the global search problems in these functions and do better than the other algorithms.

Figure 12 shows the ability of the A-GWO to search in multimodal functions (basic functions F8, F9 and F10).

As for the multimodal-expanded functions, the proposed algorithm achieves an absolute preference (as shown in Fig. 13, despite the complexities of searching in those functions.

In the last set of CEC'05, the performance of the algorithms was relatively close in terms of the best result, the standard deviation of the results obtained, and the time taken to end the search. Figure 14 shows the search behavior of algorithms on functions F15, F18.

The time factor is significant in evaluating search algorithms. This factor reflects the relationship of the algorithm's behavior with the search space and considers time as a function of the accuracy factor. The time taken to experiment with Congress on Evolutionary Computation Benchmark functions (CEC) is 37.13 Sec (11.49 for A-GWA, 12.71 for GWO, and 12.93 for PSO). Figure 15 shows the ratios of the time distributions for implementing the algorithms A-GWO, GWO, and PSO.

6.4.2 *Featuers selection experiment*

In the feature selection experiment, we implemented the algorithm over 1000 iterations and calculated the accuracy by Eq. 13, the percentage of features selected by each algorithm PSO, GWO, and A-GWO.

$$Accuracy = \frac{correct\ predation\ cases}{total\ cases} \tag{13}$$

The Random Frost (RF) is used as an objective function to examine the features selected by each model (PSO, GWO, and A-GWO). In comparison, the performance of the proposed algorithm with standalone RF, GWO, and PSO over high accuracy of features selected by each model. The proposed model (A-GWO) gets the best five over seven (71%) accuracy in seven datasets. Table 4 illustrates the performance of the algorithms (standalone RF, GWO, and PSO) compared with the proposed A-GWO.

6.5 Comparative with other studies

In this section, The proposed model is compared with previous studies on the same dataset. Table 5 compares the proposed A-GWO with three different feature selection models. The proposed A-GWO has achieved higher classification accuracy rates than others. In other words, the proposed feature selection framework has consistently accurate classification results and demonstrated its value as a potentially suitable model for feature selection.

Clearly, the proposed A-GWO has achieved higher classification accuracy rates than others. In other words, the proposed feature selection has consistently improved classification results and demonstrated its value as a suitable mode for the optimizer method.

7 Conclusion

Intelligent optimization approaches, exemplified by metaheuristic algorithms, have proven effective in reducing features in high-dimensional data. However, they face the challenge of balancing exploration and exploitation, where focusing solely on one can lead to drawbacks like instability or stagnation. This paper addresses this challenge by proposing the A-Proactive Grey Wolf Optimization (A-GWO), which employs a dual search process. A-GWO divides the population into two groups, each utilizing distinct search techniques: one leveraging GWO and the other leveraging Simi-stochastic methods. This approach promotes a better balance between exploration and exploitation, leading to improved performance. While the framework effectiveness depends on sufficient change in search locations, even with slight changes, it can still offer benefits. However, future work could explore a proactive random model to dynamically adapt to search space behavior and identify instances of stagnation for even more optimized performance.Please specify the significance of the symbol [bold] reflected inside Table [3] by providing a description in the form of a table footnote. Otherwise, kindly amend if deemed necessary. Thanks, all bold numbers in Table 3 are corrected into normal form

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. Victoria University, Australian Gov, Khandakar Ahmed.

Declarations

Conflict of interest No conflict of interest exists. We wish to confirm that there are no known conflicts of interest associated with this publication.

References

1. Calvet L, Benito S, Juan AA, Prados F (2022) On the role of metaheuristic optimization in bioinformatics. Int Trans Oper Res. https://doi.org/10.1111/itor.13164
2. González P, Prado-Rodriguez R, Gábor A, Saez-Rodriguez J, Banga JR, Doallo R (2022) Parallel ant colony optimization for the training of cell signaling networks. Expert Syst Appl 208:118199. https://doi.org/10.1016/j.eswa.2022.118199
3. Zhao M et al (2022) Intelligent healthcare system using patients confidential data communication in electrocardiogram signals. Front Aging Neurosci 14:870844. https://doi.org/10.3389/fnagi.2022.870844

4. Nematzadeh S, Kiani F, Torkamanian-Afshar M, Aydin N (2022) Tuning hyperparameters of machine learning algorithms and deep neural networks using metaheuristics: a bioinformatics study on biomedical and biological cases. Comput Biol Chem 97:107619. https://doi.org/10.1016/j.compbiolchem.2021.107619
5. Bhattacharyya A, Chakraborty R, Saha S, Sen S, Sarkar R, Roy K (2022) A two-stage deep feature selection method for online handwritten bangla and devanagari basic character recognition. SN Comput Sci 3(4):260. https://doi.org/10.1007/s42979-022-01157-2
6. Al-Janabi S, Alkaim A (2022) A novel optimization algorithm (Lion-AYAD) to find optimal DNA protein synthesis. Egypt Informatics J 23(2):271–290. https://doi.org/10.1016/j.eij.2022.01.004
7. Alsaeedi AH, Aljanabi AH, Manna ME, Albukhnefis AL (2020) A proactive metaheuristic model for optimizing weights of artificial neural network. Indones J Electr Eng Comput Sci 20(2):976–984. https://doi.org/10.11591/ijeecs.v20.i2.pp976-984
8. Al-Shammary D, Albukhnefis AL, Alsaeedi AH, Al-Asfoor M, Dhiah Al-Shammary MA, Albukhnefis AL, Alsaeedi AH (2022) Extended particle swarm optimization for feature selection of high-dimensional biomedical data. Concurr Comput Pract Exp. https://doi.org/10.1002/cpe.6776
9. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. Adv Eng Softw 69:46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007
10. Geetha TV, Deepa AJ (2022) A FKPCA-GWO WDBiLSTM classifier for intrusion detection system in cloud environments. Knowl-Based Syst 253:109557. https://doi.org/10.1016/j.knosys.2022.109557
11. Hussain SM (2021) “Multi-metaheuristic features selection model for high-dimensional biomedical data BT. In: Peng S-L, Hao R-X, Pal S (eds) Proceedings of First International Conference on Mathematical Modeling and Computational Science. Springer Singapore, Singapore, pp 309–321
12. Jabor AH, Ali AH (2019) Dual heuristic feature selection based on genetic algorithm and binary particle swarm optimization. J Univ BABYLON Pure Appl Sci 27(1):171–183. https://doi.org/10.29196/jubpas.v27i1.2106
13. Kareem SS, Mostafa RR, Hashim FA, El-Bakry HM (2022) “An effective feature selection model using hybrid metaheuristic algorithms for iot intrusion detection. Sensors (Basel) 22(4):1396. https://doi.org/10.3390/s22041396
14. Alkafagi SS, Almuttairi RM (1818) A proactive model for optimizing swarm search algorithms for intrusion detection system. J Phys Conf Ser 1:2021. https://doi.org/10.1088/1742-6596/1818/1/012053
15. Shukla AK, Singh P, Vardhan M (2019) Gene selection for cancer types classification using novel hybrid metaheuristics approach. Swarm Evol Comput. https://doi.org/10.1016/j.swevo.2020.100661
16. Sharma I, Kumar V, Sharma S (2022) A comprehensive survey on grey wolf optimization. Recent Adv Comput Sci Commun 15(3):323–333. https://doi.org/10.2174/266625581399920100716 5454
17. Zhu A, Xu C, Li Z, Wu J, Liu Z (2015) Hybridizing grey wolf optimization with differential evolution for global optimization and test scheduling for 3D stacked SoC. J Syst Eng Electron 26(2):317–328. https://doi.org/10.1109/JSEE.2015.00037
18. Kaveh A, Zakian P (2018) Improved GWO algorithm for optimal design of truss structures. Eng Comput 34(4):685–707. https://doi.org/10.1007/s00366-017-0567-1
19. Chakraborty C, Kishor A, Rodrigues JJPC (2022) Novel enhanced-grey wolf optimization hybrid machine learning technique for biomedical data computation. Comput Electr Eng 99:107778. https://doi.org/10.1016/j.compeleceng.2022.107778
20. Nuiaa RR, Manickam S, Alsaeedi AH, Alomari ES (2022) A new proactive feature selection model based on the enhanced optimization algorithms to detect DRDoS attacks. Int J Electr Comput Eng 12(2):1869–1880. https://doi.org/10.11591/ijece.v12i2.pp1869-1880
21. Alomari OA et al (2021) Gene selection for microarray data classification based on gray wolf optimizer enhanced with TRIZ-inspired operators. Knowl-Based Syst 223:107034. https://doi.org/10.1016/j.knosys.2021.107034
22. Long W, Jiao J, Liang X, Tang M (2018) An exploration-enhanced grey wolf optimizer to solve high-dimensional numerical optimization. Eng Appl Artif Intell 68:63–80. https://doi.org/10.1016/j.engappai.2017.10.024
23. Alfoudi AS, Alsaeedi AH, Abed MH, Otebolaku AM, Razooqi YS (2021) Palm vein identification based on hybrid feature selection model. Int J Intell Eng Syst 14(5):469–478. https://doi.org/10.22266/ijies2021.1031.41
24. Sharda S et al (2022) A hybrid machine learning technique for feature optimization in object-based classification of debris-covered glaciers. Ain Shams Eng J 13(6):101809. https://doi.org/10.1016/j.asej.2022.101809
25. Zhong C, Li G, Meng Z (2022) Beluga whale optimization: a novel nature-inspired metaheuristic algorithm. Knowl-Based Syst 251:109215. https://doi.org/10.1016/j.knosys.2022.109215
26. Hashim FA, Houssein EH, Hussain K, Mabrouk MS, Al-Atabany W (2022) Honey Badger Algorithm: new metaheuristic algorithm for solving optimization problems. Math Comput Simul 192:84–110. https://doi.org/10.1016/j.matcom.2021.08.013
27. Osuna-Enciso V, Cuevas E, Morales Castañeda B (2022) A diversity metric for population-based metaheuristic algorithms. Inf Sci (Ny) 586:192–208. https://doi.org/10.1016/j.ins.2021.11.073
28. Al-saeedi AH (2016) Binary mean-variance mapping optimization algorithm (BMVMO). J Appl Phys Sci 2(2):42–47. https://doi.org/10.20474/japs-2.2.3
29. Suganthan PN et al. (2005) “Problem definitions and evaluation criteria for the CEC 2005 special session on real-parameter optimization,” Tech. Report, Nanyang Technol. Univ. Singapore, May 2005 KanGAL Rep. 2005005, IIT Kanpur, India
30. Ahmed AM, Rashid TA, Saeed SAM (2020) Cat swarm optimization algorithm: a survey and performance evaluation. Comput Intell Neurosci. https://doi.org/10.1155/2020/4854895
31. Mirjalili S, Lewis A (2013) S-shaped versus V-shaped transfer functions for binary Particle Swarm Optimization. Swarm Evol Comput 9:1–14. https://doi.org/10.1016/j.swevo.2012.09.002
32. Debata PP, Mohapatra P (2021) Identification of significant biomarkers from high-dimensional cancerous data employing a modified multi-objective meta-heuristic algorithm. J King Saud Univ Comput Inf Sci. https://doi.org/10.1016/j.jksuci.2020.12.014
33. Wei G, Zhao J, Feng Y, He A, Yu J (2020) A novel hybrid feature selection method based on dynamic feature importance. Appl Soft Comput J 93:106337. https://doi.org/10.1016/j.asoc.2020.106337
34. Milletich R, Asmar A, Milletich R (2020) “Stochastic optimization in machine learning pipelines : selecting features and hyperparameters stochastic optimization in machine learning pipelines : selecting features and hyperparameters,” no. 24, pp. 1–10
35. Kumar A, Maan VS, Saini M (2024) Performance optimization of hydroelectric power-plants using computational intelligence techniques. Int J Inform Tecnol. https://doi.org/10.1007/s41870-024-01771-8
36. Singh PN, Mahapatra RP (2024) A novel deep learning approach for arrhythmia prediction on ECG classification using recurrent CNN with GWO. Int J Inform Tecnol 16:577–585. https://doi.org/10.1007/s41870-023-01611-1
37. Sharma S, Yadav NS (2023) A multilayer stacking classifier based on nature-inspired optimization for detecting cross-site scripting

attack. Int J Inform Tecnol 15:4283–4290. https://doi.org/10.1007/s41870-023-01459-5

38. Angaitkar P, Janghel RR, Sahu TP (2023) gHPCSO: gaussian distribution based hybrid particle cat swarm optimization for linear B-cell epitope prediction. Int J Inform Tecnol 15:2805–2818. https://doi.org/10.1007/s41870-023-01294-8
39. Sonkar N, Pandey S, Kumar S (2023) Bundle relaying scheme for network deployed using grey wolf optimization in delay tolerant networks. Int J Inform Tecnol 15:345–353. https://doi.org/10.1007/s41870-022-01135-0