Efficient Improved Henry Gas Solubility Optimization and Its Application in Feature Selection Problems

Jiayin Wang, Ronghe Zhou, Yukun Wang*, and Zhongfeng Li

Abstract—**The henry gas solubility optimization algorithm is a meta-heuristic algorithm inspired by Henry's law. While it has demonstrated effectiveness in solving various optimization problems, it does face certain limitations such as insufficient population diversity, and slow convergence speed when dealing with complex problems. In this paper, we propose an enhanced version of henry gas solubility optimization algorithm, known as E_HGSO. First, we introduce a new group search formula to improve the ability of avoiding easy to fall into local optimum and searching in a single range, while we introduce the concept of a search factor to strike a balance between exploration and exploitation. Second, we introduce a position update formula to enhance the diversity and randomness of the search process. Finally, we propose a new worst gas position update formula with a Lévy flight mechanism. This mechanism enhances the gas search's ability to adapt to different distance requirements within the search space, leading to improved search efficiency and accuracy. To evaluate the effectiveness of the E_HGSO algorithm, we conducted a comparison with eight algorithms on the CEC2017 benchmark functions. The results of the Friedman test and Wilcoxon rank sum test indicate that the proposed E_HGSO outperformed the comparison algorithms. Furthermore, we applied E_HGSO to the feature selection problem. The results indicate that E_HGSO performs competitively across various metrics, including, classification accuracy, convergence speed, and convergence precision.**

Index Terms— **Henry gas solubility optimizer, search factor, lévy flight, feature selection**

I. INTRODUCTION

Meta-heuristic optimization algorithms are widely employed for solving global optimization problems. These algorithms simulate nature and human intelligence to search for optimal solutions. These algorithms exhibit several key characteristics: (1) Suitable for solving large scale optimization problems involving multiple variables and constraints. These algorithms are able to search in high-dimensional spaces and find optimal solutions. (2) Highly flexible, with appropriate meta-algorithms selectable based on the specific characteristics of the problem. (3) Faster convergence and shorter solution times compared to alternative optimization techniques. (4) Meta-heuristic optimization algorithms perform well in dealing with non-linear problems. They are able to find globally optimal or near-optimal solutions by using diverse search strategies. Additionally, the flexibility and independence from gradients of meta-heuristic algorithms provide them with an advantage in solving global optimization problems. When compared to traditional optimization methods like simulated annealing algorithm, meta-heuristic algorithms excel in finding optimal solutions by simulating nature and human intelligence, which makes them particularly effective in tackling complex optimization problems [1].

Meta-heuristic optimization algorithms can be classified into four main categories: Evolution-based algorithms, Group intelligence based algorithms, Human based algorithms, Physics and chemistry based algorithms.

Evolution-based algorithms are mainly designed to achieve the overall progress of the group and ultimately complete the optimal solution by simulating the evolutionary law of superiority and inferiority in nature (Darwin's law). The Genetic Algorithm (GA) [2] and Differential Evolution (DE) [3] are the main representatives. With the continuous exploration of natural evolution-based algorithms by scientists, various evolutionary optimization algorithms have been proposed, among which some of the more popular algorithms are Evolutionary Strategies (ES) [4], Evolutionary Programming (EP) [5], Gene Expression Programming (GEP) [6], Covariance Matrix Adaptive Evolutionary Strategies (CMA-ES) [7], Biogeography Based Optimization (BBO) [8] and so on.

Group intelligence optimization algorithms use the intelligence of the group to achieve the global optimal solution. In such algorithms, each group is considered as a population of organisms that performs tasks that cannot be performed by individuals through collaborative behaviour among individuals. Some of the optimization algorithms based on group intelligence are listed below: Particle Swarm Optimization (PSO) [9], Beluga Whale Optimization (BWO) [10], Bee Colony Algorithm (BCA) [11], Artificial Hummingbird Algorithm (AHA) [12], Grey Wolf Optimization (GWO) [13], Snake Algorithm (SO) [14], Butterfly Optimization Algorithm (BOA) [15], Honey Badger Algorithm (HBA) [16].

Manuscript received April 22, 2024; revised July 26, 2024. This work was supported by Liaoning Provincial Joint Funds Project of China (Grant No. 2023-MSLH-323).

Jiayin Wang is a postgraduate student of School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, Liaoning 114051, PR China; (e-mail: wangjy_97@163.com)

Ronghe Zhou is a PhD Candidate of School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, Liaoning 114051, PR China; (e-mail: ronghezhou@ustl.edu.cn）

Yukun Wang is an associate professor of School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, Liaoning 114051, PR China; (*Corresponding author to provide phone: 86-0412-5929699； e-mail: wyk410@163.com)

Zhongfeng Li is an associate professor of School of Electrical Engineering, Yingkou Institute of Technology, Yingkou, Liaoning, PR China; (e-mail: afeng0601@163.com).

Algorithms based on human behaviour typically use patterns of human behavior and decision-making processes to achieve optimization and solve problems. For example, human teaching behaviour, social behaviour, learning behaviour,management behaviour and so on. The following is a list of intelligent optimization algorithms proposed based on human behaviour: Teaching-Learning-Based Optimization (TLBO) [17], Tabu Search Algorithm (TS) [18], League Championship Algorithm (LCA) [19], Seeker Optimization Algorithm (SOA) [20], Exchange Market Algorithm (EMA) [21], Group Counselling Optimization Algorithm (GCO) [22], Social Learning optimization (SLO) [23], Cultural Evolution Algorithm (CEA) [24], Volleyball Premier League Algorithm (VPL) [25].

Physics and chemistry based algorithms focus on solving problems by using physical and chemical principles to simulate the behaviour and change of a material system through computer programming. These algorithms are usually used to compute problems in the fields of physics, chemistry, engineering and some of the more popular algorithms are: Simulated Annealing (SA) [26], Gravitational Local Search (GLSA) [27], Big-bang Big-Crunch (BBBC) [28], Gravitational Search Algorithm (GSA) [29], Central Force Optimization (CFO) [30], Galaxy-based-Search Algorithm (GbSA) [31], Flow Direction Algorithm [32].

It is widely acknowledged that there is no universal optimization algorithm capable of perfectly solving all types of optimization problems. Because the characteristics and complexity of different problems vary significantly. With the rapid advancement of technology has led to an exponential growth in the volume of data generated across various domains, accompanied by increasing complexity and diversity of the data. However, challenges such as data redundancy and excessively long modeling times have become significant obstacles to effective data analysis. To address these pressing issues more efficiently, there is an increasing need for optimization algorithms that can handle both continuous and discrete optimization problems simultaneously. The HGSO algorithm is an attractive algorithm owing to the fact that equilibrium exploration and exploitation play pivotal role in the algorithm, a property that makes HGSO suitable for solving complex optimization problems with many locally optimal solutions.

Henry's law is a fundamental principle of physical chemistry, proposed in 1803 during the study of gas solubility in liquids. It can be expressed as: at constant temperature and pressure, the solubility of a volatile solute in a solution is proportional to the equilibrium partial pressure of that solute above the liquid surface. The novel meta-heuristic algorithm for Henry's Gas Solubility Optimization (HGSO) is inspired by the principles of Henry's law, and it mimics the behavior governed by this fundamental physical chemistry concept to solve complex optimization problems [33]. The HGSO algorithm mimics the crowding behavior of the gas to balance mining and exploration in the search space and avoid local optimizations. A number of researchers have conducted some studies on the HGSO problem from both practical problems and theoretical studies, which in turn leads to better use of the HGSO algorithm to solve practical problems. Betül Sultan Yıldız et al. [34] proposed a meta-heuristic algorithm based on the integration of chaotic mapping into HGSO, named Chaotic Henry's Gas Solubility optimization Algorithm (CHGSO) , which improves the speed of convergence, solves real engineering optimization problems, and obtains the optimal variable in mechanical design and manufacturing optimization problems. Fatma A. Hashim et al. [35] proposed an improved Henry's Gas Solubility optimization (MHGSO) algorithm for the discovery of functional motifs in DNA genome sequences, which is capable of accurately detecting the target modality. Serdar Ekinci et al. [36] introduced a new the Henry's Gas Solubility Optimisation (OBL/HGSO) based on inverse learning. Cao et al. [37] proposed a SVR-based prediction method, Henry Gas Solubility optimization Algorithm, by randomly generating support vector regression machine parameters in a certain range to form a parameter population, the prediction accuracy (PA) to get the population and SVR used, and at the same time updating the population and the optimal SVR parameters by PAs and HGSO to get the best overall performance. Davood Mohammadi et al. [38,39] borrowed a new scheme from quantum theory to update the position of each solution, improving the original algorithm to improve the exploration performance to explore the search space, named QHGSO.

HGSO algorithm exhibits a wide range of applications across various fields; however, it still encounters certain challenges including susceptibility to local optima, and sluggish convergence speed. During the initial stage of iteration, the interaction among gases in the HGSO algorithm often leads to gas cluster aggregation, resulting in premature convergence. However, as the iteration progresses, this interaction becomes less effective in facilitating individuals to escape from local optima, thereby impacting the algorithm's accuracy. Therefore, this paper proposes three strategies to address the shortcomings of HGSO by creating a new grouping search formula, a new position update formula, and the introduction of a grouping mechanism for Lévy flights.

The paper is structured as follows: Section I gives a review of the existing literature; Section II describes the base HGSO algorithm along with the shortcomings and motivation of this study. Section III presents the improved algorithm E_HGSO. Section IV gives the experimental setup of the E_HGSO and analyses the experimental results and compares them with other algorithms. Section V applies our modified algorithm to feature selection and finally, Section VI discusses and concludes the study.

II. BASIC PRINCIPLE OF HENRY’S GAS SOLUBILITY OPTIMIZATION ALGORITHM

A. Henry’s law

In this section, the concept of the Henry's Gas Solubility Optimization Algorithm (HGSO) will be introduced. The algorithm is inspired by the famous Henry's law. Simulating the cumulative behaviour of natural gas, the HGSO algorithm balances exploration and extraction in the search space and avoids local optimization.

Henry's law discovered by Henry of England in 1803 while studying the law of solubility of gases in liquids. This can be expressed as follows: at a given temperature and equilibrium, the solubility of a gas in a liquid (expressed as a mole fraction) is directly proportional to the equilibrium partial pressure of that gas. As shown in the following relationship:

$$S_g = H \times P_g \quad (1)$$

Where H is Henry's constant and P_g is the partial pressure of the gas. H gives a good indication of the amount of gas dissolved, strictly speaking, Henry's law is only an approximate law and cannot be applied to systems with higher pressures. In this sense, the Henry constant is simply a function of temperature and has nothing to do with pressure.

The Henry's coefficient varies with temperature, and as the temperature increases, the volatility of the volatile solute increases and the Henry's coefficient increases, which can be described by the van't Hoff equation as follows:

$$\frac{d \ln H}{d(1/T)} = \frac{-\nabla_{sol} E}{R} \quad (2)$$

$$H(T) = \exp(B/T) \times A \quad (3)$$

Where ∇_{sol}E is the enthalpy of dissolution, the gas constant R is the gas constant, and A and B are the two parameters of the relationship between H and T. H is a function of parameters A and B.

Van's Hoff equation is valid when ∇_{sol}E is a constants:

$$H(T) = \exp(-C \times (1/T - 1/T^{\theta})) \times H^{\theta} \quad (4)$$

B. HGSO mathematical model

HGSO is characterized by several fundamental structural components, including the initialization of candidate solutions, the iterative refinement of those solutions, the evaluation of their fitness, and the selection of the optimal solution. It maintains a population of candidate solutions in the form of gas particles dissolved in a given liquid. The properties of these gas particles are updated during the exploration and development phases of the HGSO in order to find the best positions in the search space.

This section describes the mathematical model of the HGSO algorithm in the following steps:

Initialization process: The population of candidate solutions with N gas particles is initialised with the relationship between the number of gases and the positions of the gases, as well as the number of gases i, the value of Henry's constant $j(H_j(t))$, the partial pressure $P_{i,j}$ of the gas i in the cluster j, and the value of the ∇_{sol}E/R constant $j(C_i)$, respectively, by using the following two equations:

$$X_i(t+1) = X_{\min} + r \cdot (X_{\max} - X_{\min}) \quad (5)$$

$$H_j(t) = l_1 \times rand(0,1), P_{i,j} = l_2 \times rand(0,1), C_j = l_3 \times rand(0,1) \quad (6)$$

Where $X_{(i)}$ denotes the position of the ith gas among all gases N, r is a random number between 0 and 1, X_{min} is the minimum boundary, X_{max} is the upper boundary, and t denotes the iteration time. In Eq.(6), l_1, l_2, l_3 represent constants of 5E-02, 100, 1E-02, respectively.

Aggregation and evaluation: The population agent is divided into an equal number of clusters based on the type of gas. Each cluster contains similar gases and the same value of Henry's constant (H_j). At the same time, each cluster j is evaluated to determine the best gas to obtain the highest equilibrium from other gases of the same type. The gases are then ranked to find the best gas in the entire group.

Updating of Henry's coefficient: The Henry coefficients are updated applying the following equation:

$$H_j(t+1) = H_j(t) \times \exp(-C_j \times (\frac{1}{T(t)} - \frac{1}{T^{\theta}})) \quad (7)$$

$$T(t) = \exp(-t / iter) \quad (8)$$

Where H_j is the Henry's coefficient of the jth cluster, T is the temperature, T^{θ} is a constant with a constant value of 289.15 and *iter* is the maximum number of iterations.

Solubility update: At the tth iteration, the solubility of the ith gas particle in the jth cluster is updated using the following equation:

$$S_{i,j}^t = K \times H_j^{t+1} \times P_{i,j}^t \quad (9)$$

Where $S_{i,j}$ represents the solubility of the gas, $P_{i,j}$ is the partial pressure of the gas in the jth cluster of the gas, and K is a constant therein.

Position update: The way in which the position of the jth cluster gas is updated:

$$\begin{aligned} X_{i,j}(t+1) = {} & X_{i,j}(t) + F \times r \times \gamma \times (X_{i,best}(t) - X_{i,j}(t)) \\ & + F \times r \times \alpha \times (S_{i,j}(t) \times X_{best}(t) - X_{i,j}(t)) \end{aligned} \quad (10)$$

$$\gamma = \beta \times \exp(-\frac{F_{best}(t) + \varepsilon}{F_{i,j}(t) + \varepsilon}), \varepsilon = 0.05 \quad (11)$$

In Eqs. 10 and 11, $X_{(i,j)}$ is used to denote the position of gas i in cluster j, r is a random number, t represents the current number of iterations, $X_{(i,best)}$ is the best position of gas i in cluster j, and X_{best} is the best position of gas i in the whole population, which is mainly used for balanced exploration and exploitation. γ represents the interaction force between gases in cluster i, α denotes the effect of other gas particles on the ith particle, and β is a constant. $F_{(i,j)}$ and F_{best} denote the fitness of gas i in cluster j and the fitness of the best gas in the whole population, respectively, and the value of F is used to guide the direction of gas movement.

Escape from local optimum: In order to solve the problem of falling into a local optimum during the search for the best gas, the HSGO algorithm uses Eq. 12 to update the worst agents for sorting and selection.

$$Nw = N \times (rand(c_2 - c_1) + c_1), c_1 = 0.1, c_2 = 0.2 \quad (12)$$

Where N is the population size, *rand* is a random number between [0, 1], c_1, c_2 are constants with values 0.1, 0.2 respectively.

Update the position of the worst individual: Position update for the worst agent:

$$G_{(i,j)} = G_{\min(i,j)} + r \times (G_{\max(i,j)} - G_{\min(i,j)}) \quad (13)$$

Where $G_{(i,j)}$ is the position of gas i in cluster j and r is a random number. G_{min} and G_{max} are the maximum and minimum values of the whole range.

III. The proposed E_HGSO algorithm

The basic HGSO has the disadvantages of slowly convergence and falling into local optimal solutions. In order to solve this drawback, this paper carries out three improvements on the basis of the original algorithm, which are as follows:

A. Creating new group search formulas

In order to avoid the dilemma of easy to fall into the local optimal and search a single range of problems, design a formula for the combination of intra-group and inter-group search, in order to expand the scope of the search, to improve the efficiency of the work at the same time, to avoid the local optimal. Meanwhile, we introduce the search factor α, β, to improve the accuracy:

$$r < -\tanh(\lambda \times \frac{T_{\max} - t}{T_{\max}}) \tag{14}$$

$$r < \mu e^{-\frac{t}{T_{\max}}} \tag{15}$$

If the gas i satisfies Eqs. 14 and 15 at the same time, the intra-group search will be performed, otherwise the inter-group search will be carried out, in which λ and μ are constants that need to be set manually. *tan*h is a hyperbolic tangent function, which is characterised by the fact that it takes the value of 0 at the origin, while it tends to 1 and -1 at the positive infinity and negative infinity, respectively.

B. New position update formula

The position renewal process in the HGSO algorithm, while speeding up convergence, also leads to a rapid loss of population diversity in the algorithm, and the algorithm tends to fall into a local optimum. For this disadvantage, this paper designs a new prey encircling formula that introduces two stochastic gases, adding diversity and randomness to the search, as shown in equation (16).

$$X_{i,j}(t+1) = X_{rr1}(t) + [(\frac{T_{\max} - 0.5t}{T_{\max}})^2 \cos\left(2\pi r\right)\gamma \times X_{best}(t) - (1 + S_{i,j}(t) \times X_{rr2}(t))] \tag{16}$$

Where rr_1, rr_2 are three unequal random numbers.

C. New formula for updating the worst gas

The search pattern of the HGSO algorithm is characterized by its simplicity, as it follows a fixed path towards the target during each search. However, this simplicity can lead to a susceptibility to local optima, thereby diminishing its local search capability. To tackle this challenge, we have introduced a novel formula for updating the position of the worst gas agent and incorporated the Lévy flight mechanism. By updating the position of the worst agent, the algorithm can converge more rapidly towards better solutions. This optimization not only saves search time but also improves the overall efficiency of the algorithm.

$$X_{i,j}(t) = \frac{1}{3}[X_{rr1}(t) + X_{rr2}(t) + X_{rr3}(t) + le'vy] \tag{17}$$

Where Lévy is a D-dimensional vector generated by the Lévy flight operator, rr_1, rr_2 and rr_3 are three unequal random numbers.

IV. Results of Experiment and statistical analysis

A. Benchmark functions

In this paper, CEC2017 benchmark functions [41] are used to verify and evaluate the performance of E_HGSO and its comparison algorithms. Different types of functions can effectively check the optimization ability of the algorithms. The CEC2017 benchmark functions include 30 different functions, which can be classified into 4 categories, The related information of the CEC2017 benchmark functions can be referred to in Table I below.

B. Sensitivity analysis of E_HGSO

From the description of the E_HGSO, α and β affect both within-group and between-group search. In this subsection, the sensitivity of these two parameters of E_HGSO is analysed. In performing the sensitivity analysis, we select different types of functions from CEC2017 as evaluation metrics and run the algorithm several times with different parameter combinations. The combination of α and β parameters with the best overall performance was selected as the initial parameters of E_HGSO. This method is often used in many studies.

From Fig.1, it can be seen that the comprehensive performance of E_HGSO under the P31 parameter combination is the best. Therefore, the values of α and β in this paper are set to 4 and 0.4, respectively.

C. Qualitative comparison between HGSO and E_HGSO

To compare the HGSO and E_HGSO, we will perform a qualitative analysis of their population diversity and convergence characteristics using the unimodal function f_3 and the composite function f_{10} from the CEC2017 test function suite. Population diversity is a crucial factor in the performance of optimization algorithms, as it reflects the algorithm's ability to explore the search space and avoid premature convergence. The population diversity can be quantified using the following formula:

$$Diversity(t) = \frac{1}{N} sqrt(\sum_{i=1}^{N} \|X_i(t) - X^g\|) \tag{18}$$

Where N represents the size of the population, $X_i(t)$ denotes the current position of the ith gas, and X^g denotes the best position among the gases in this cluster.

The rules for exploration and exploitation are defined as:

$$\begin{cases} Exploration(t) = \dfrac{div(t)}{\max(div)} \times 100 \\ Exploitation(t) = \left| \dfrac{div(t) - \max(div)}{\max(div)} \right| \times 100 \end{cases} \tag{19}$$

TABLE I
DIMENSION OF FUNCTIONS USED FOR PARAMETER SELECTION

Type	$f(x)$	Dimension
Unimodal functions	f_1	20
	f_2	10
multimodal functions	f_3	20
	f_4	10
	f_5	10
Hybrid functions	f_6	20
	f_7	20
	f_8	20
	f_9	10
Composition functions	f_{10}	10
	f_{11}	20
	f_{12}	20

TABLE II
DIFFERENT PARAMETER COMBINATION OF E_HGSO

	Different parameter combinations								
β \ α	1	2	3	4	5	6	7	8	9
0.1	P1	P2	P3	P4	P5	P6	P7	P8	P9
0.2	P10	P11	P12	P13	P14	P15	P16	P17	P18
0.3	P19	P20	P21	P22	P23	P24	P25	P26	P27
0.4	P28	P29	P30	P31	P32	P33	P34	P35	P36
0.5	P37	P38	P39	P40	P41	P42	P43	P44	P45
0.6	P46	P47	P48	P49	P50	P51	P52	P53	P54
0.7	P55	P56	P57	P58	P59	P60	P61	P62	P63
0.8	P64	P65	P66	P67	P68	P69	P70	P71	P72
0.9	P73	P74	P75	P76	P77	P78	P79	P80	P81

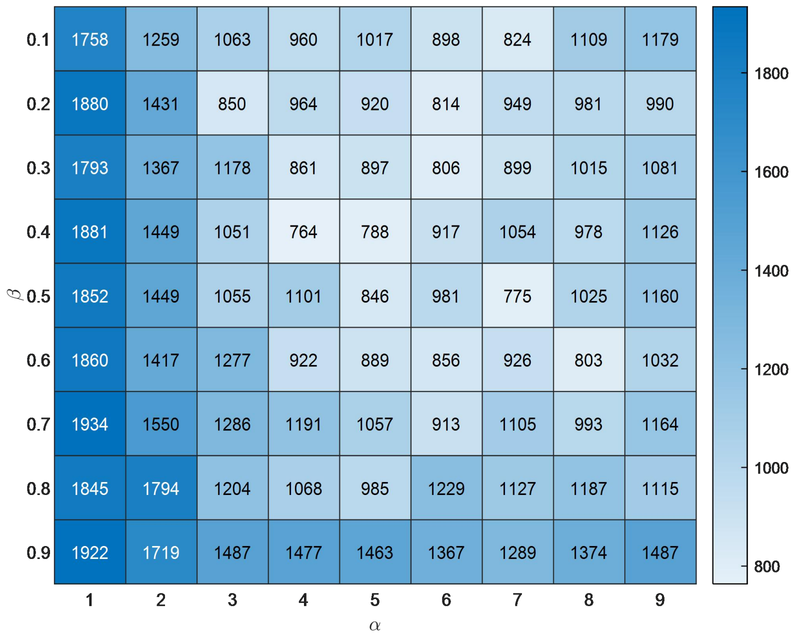

Fig.1. The sensitivity analysis results of E_HGSO for different types of function

The observations made from Fig.2 and 3 highlight notable differences between the E_HGSO algorithm and the original HGSO algorithm. The diversity curve analysis reveals that the E_HGSO algorithm experiences a significant decline in population diversity as the iteration progresses, leading to a progressively regionally stable population in the later stages. Additionally, the exploration and convergence curves illustrate that the enhanced algorithm excels in exploration during the early stages of iteration. As the iteration progresses, the algorithm's exploitation capabilities become more dominant, which contributes to improved convergence accuracy. Based on in-depth analysis of the convergence curves, we found that the E_HGSO algorithm exhibits superior performance in terms of fitness. When solving minimization problems, the E_HGSO algorithm demonstrates better performance and is more likely to find the global optimal solution. These results indicate that the E_HGSO algorithm is a highly efficient optimization tool worthy of attention, and it has shown remarkable advantages in solving complex problems.

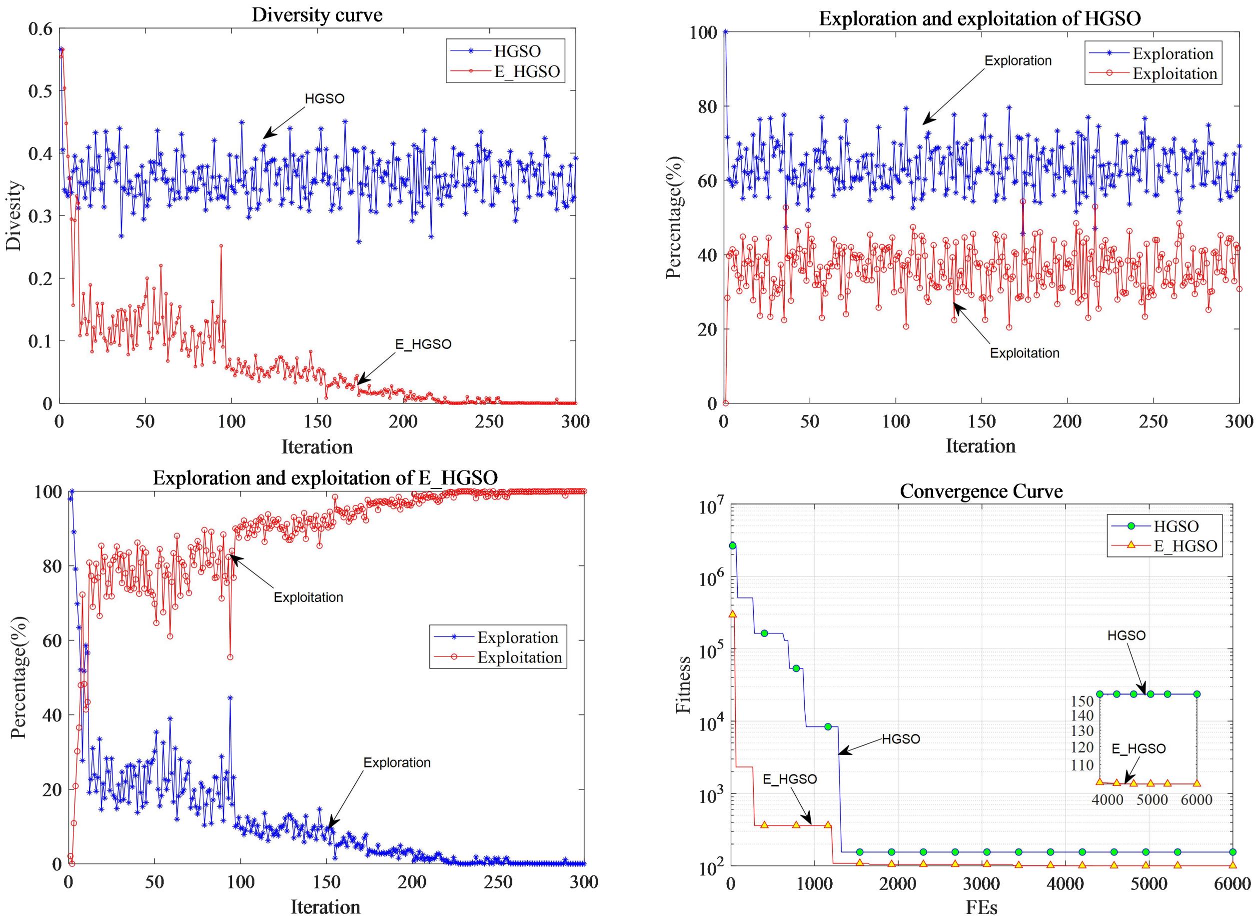

Fig.2. Qualitative comparison results on f_3

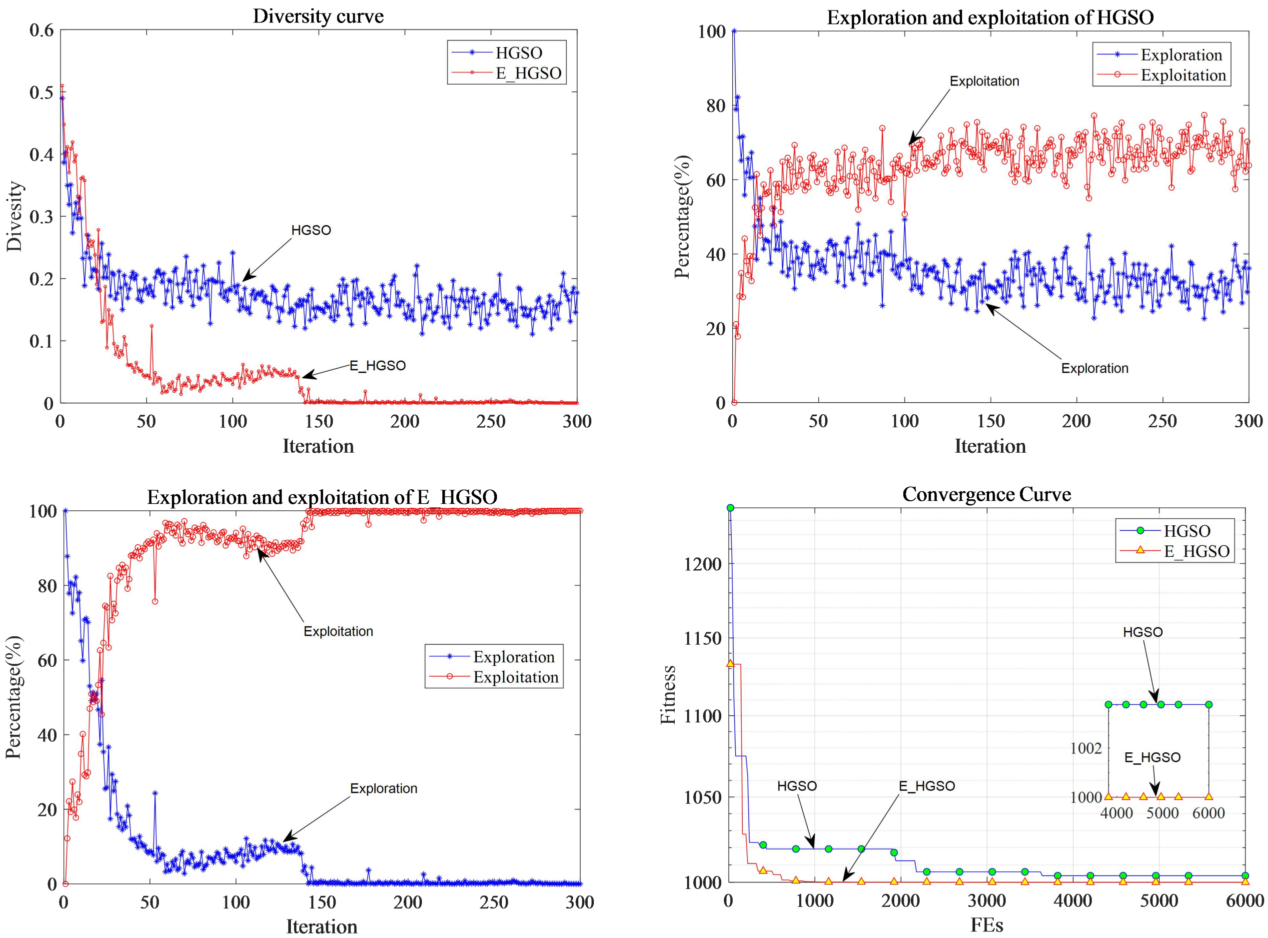

Fig.3. Qualitative comparison results on f_{10}

TABLE III
PARAMETERS SETTINGS OF COMPARISON ALGORITHMS

No.	Name	Parameter settings	Literature	Publication year
1	TS22	*ns* = 5 (Number of Stars), *SN* = 10 (Signal to Noise Ratio)	[18]	2022
2	HBA	β=6，*c*=2	[16]	2022
3	FDA	α=*N* (Popsize), β=8	[32]	2021
4	AHA	m=2n	[12]	2022
5	SO	Nm=Nf=0.5N,c1=0.5, Thresholds(food)=0.25,Threshold(temp)=0.25	[14]	2022
6	BWO	W_f decreases linearly from 0.1 to 0.05	[10]	2022
7	HGSO	c_1 = 0.1, c_2 = 0.2, ε = 0.05, α = 1, $T\theta$ = 289.15;	[33]	2019
8	QHGSO	m_1=0.1,m_2=0.2;$T\theta$ = 298.15;	[38]	2021
9	E_HGSO	α=4, β=0.4	-	-

TABLE VI
FORMANCE OF COMPARISON ALGORITHMS ON CEC2017

Name	Comparative results under different indicators					
	D=50			*D*=100		
	Mean	Std	Best	Mean	Std	Best
	Best/Second/Worst	Best/Second/Worst	Best/Second/Worst	Best/Second/Worst	Best/Second/Worst	Best/Second/Worst
TS22	2/0/0	4/1/0	0/1/0	1/1/0	3/3/0	0/1/0
HBA	0/0/0	0/0/6	0/1/0	0/1/0	0/1/5	0/1/0
FDA	1/1/0	1/0/1	2/0/0	0/1/0	0/1/0	1/1/0
AHA	2/2/0	1/3/3	3/3/0	1/3/0	0/1/2	2/2/0
SO	2/6/0	1/10/0	0/8/0	0/6/0	1/6//2	0/7/0
BWO	0/0/29	2/1/13	0/0/29	0/0/29	3/1/13	0/0/29
HGSO	0/0/0	4/1/3	0/0/0	0/0/0	1/2/4	0/0/0
QHGSO	6/12/0	3/8/3	11/9/0	5/11/0	5/9/3	7/11/0
E_HGSO	16/8/0	13/5/0	13/7/0	22/6/0	16/5/0	19/6/0

TABLE VII
AVERAGE RANKINGS OF DIFFERENT ALGORITHMS BY FRIEDMAN'S TEST AT CEC2017

Name	Ranking in different dimension	
	D=50	D=100
TS22	5.0345	5.0345
HBA	5.1034	5.1379
FDA	5.8966	5.8966
AHA	4.3103	3.7586
SO	3.2759	3.7241
BWO	9.0000	9.0000
HGSO	8.0000	7.8276
GHGSO	2.6552	3.2069
E_HGSO	1.7241	1.4138
p-value	7.4399e-34	1.0728e-32

The analysis of the convergence accuracy curve shows that although the final outcomes are generally comparable, a closer examination through the zoomed-in graph reveals that the results of the E_HGSO exhibit greater precision. By comparing the performance of E_HGSO to the original algorithm, we can conclude that the improvement strategy we implemented has been highly effective in enhancing the algorithm's efficiency.

D.Compare to other algorithms

To verify the performance of the E_HGSO algorithm, we compared it with the HGSO and QHGSO algorithms, as well as six other popular algorithms on the CEC2017 benchmark functions. The parameters for each algorithm were set as shown in III. The experiments were conducted on a range of benchmark problems to comprehensively evaluate the comparative performance of the algorithms.

TABLE IV
RESULTS OF COMPARING ALGORITHMS ON THE CEC2017 BENCHMARK FUNCTION (D=50)

f(x)	Index	TS22	HBA	FDA	AHA	SO	BWO	HGSO	QHGSO	E_HGSO
f_1	Mean	5.0467E+07	4.0720E+06	1.3641E+09	1.9665E+05	8.9469E+05	1.0045E+11	4.2407E+10	4.3082E+03	3.8899E+03
	Std	1.3110E+07	2.3629E+06	5.3791E+08	2.6273E+05	1.1299E+06	3.8826E+09	7.1478E+09	6.3049E+03	3.5388E+03
	Best	3.1725E+07	1.5259E+06	5.7839E+08	4.4219E+04	4.0069E+04	9.1458E+10	2.7898E+10	1.0051E+02	5.6110E+02
	Rank	6	5	7	3	4	9	8	2	**1**
f_3	Mean	1.0903E+05	1.2964E+05	6.7132E+04	3.8944E+04	1.2831E+05	1.9604E+05	1.4848E+05	4.3568E+04	4.4604E+04
	Std	1.3495E+04	1.6664E+04	1.3571E+04	8.9288E+03	1.2966E+04	1.7387E+04	8.2908E+03	1.1641E+04	1.1351E+04
	Best	8.2552E+04	1.0461E+05	3.9045E+04	2.1757E+04	9.6936E+04	1.6306E+05	1.2890E+05	2.5138E+04	2.6585E+04
	Rank	5	7	4	**1**	6	9	8	2	3
f_4	Mean	5.7336E+02	5.5927E+02	6.7770E+02	5.5229E+02	5.6634E+02	2.9846E+04	8.9138E+03	5.2109E+02	5.0436E+02
	Std	2.2978E+01	5.1383E+01	4.7657E+01	5.1601E+01	5.2665E+01	2.4544E+03	1.9489E+03	3.7593E+01	5.4080E+01
	Best	5.0061E+02	4.7269E+02	5.7918E+02	4.4995E+02	4.2917E+02	2.6155E+04	5.5490E+03	4.7525E+02	4.2277E+02
	Rank	6	4	7	3	5	9	8	2	**1**
f_5	Mean	8.0126E+02	7.1736E+02	8.0796E+02	7.9673E+02	6.3588E+02	1.1693E+03	1.0697E+03	6.3275E+02	5.5655E+02
	Std	2.7971E+01	3.0554E+01	5.4872E+01	3.1221E+01	2.0849E+01	2.2951E+01	2.1705E+01	5.6397E+01	1.4449E+01
	Best	7.3971E+02	6.4211E+02	7.0205E+02	7.3979E+02	5.7611E+02	1.1064E+03	1.0172E+03	5.9651E+02	5.3283E+02
	Rank	6	4	7	5	3	9	8	2	**1**
f_6	Mean	6.2790E+02	6.1063E+02	6.6336E+02	6.1031E+02	6.0578E+02	6.9883E+02	6.8211E+02	6.0069E+02	6.0007E+02
	Std	6.5889E+00	6.5621E+00	5.7539E+00	9.9865E+00	3.1244E+00	3.3280E+00	6.7293E+00	5.7897E-01	1.1680E-01
	Best	6.1620E+02	6.0124E+02	6.5001E+02	6.0044E+02	6.0127E+02	6.8612E+02	6.6606E+02	6.0016E+02	6.0001E+02
	Rank	6	5	7	4	3	9	8	2	**1**
f_7	Mean	1.1945E+03	1.0522E+03	1.1964E+03	1.2991E+03	9.0583E+02	1.9156E+03	1.7023E+03	9.3761E+02	8.0530E+02
	Std	5.6633E+01	5.6900E+01	7.0695E+01	1.3064E+02	3.6210E+01	4.5012E+01	9.4896E+01	6.4798E+01	1.2994E+01
	Best	1.0868E+03	9.8468E+02	1.0877E+03	1.1607E+03	8.4510E+02	1.8338E+03	1.4963E+03	8.4211E+02	7.8365E+02
	Rank	5	4	6	7	2	9	8	3	**1**
f_8	Mean	1.1237E+03	1.0441E+03	1.1086E+03	1.1267E+03	9.4128E+02	1.4837E+03	1.4109E+03	9.3797E+02	8.5976E+02
	Std	2.9783E+01	6.0920E+01	4.9144E+01	3.9253E+01	1.9572E+01	2.4079E+01	2.2455E+01	2.5914E+01	1.1804E+01
	Best	1.0718E+03	9.5292E+02	1.0348E+03	1.0516E+03	8.9883E+02	1.4228E+03	1.3630E+03	9.0008E+02	8.3781E+02
	Rank	6	4	5	7	3	9	8	2	**1**
f_9	Mean	1.5959E+04	9.5070E+03	1.0826E+04	1.0608E+04	1.9005E+03	3.6806E+04	2.8414E+04	1.3676E+03	9.1268E+02
	Std	1.8104E+03	3.2895E+03	4.0149E+03	2.4563E+03	6.8929E+02	2.0277E+03	2.6945E+03	5.1135E+02	1.5790E+01
	Best	1.1115E+04	4.5223E+03	5.9037E+03	6.1102E+03	1.3813E+03	3.1611E+04	2.4197E+04	9.8130E+02	9.0054E+02
	Rank	7	4	6	5	3	9	8	2	**1**
f_{10}	Mean	7.2256E+03	7.4491E+03	9.4600E+03	6.2099E+03	5.4231E+03	1.4494E+04	1.3883E+04	1.3844E+04	6.7160E+03
	Std	5.9467E+02	2.4416E+03	8.5528E+02	8.7344E+02	1.6821E+03	4.4110E+02	5.3178E+02	7.5078E+02	1.0808E+03
	Best	6.1812E+03	5.4022E+03	7.7380E+03	5.0095E+03	4.0029E+03	1.3116E+04	1.2337E+04	1.1158E+04	3.9975E+03
	Rank	4	5	6	2	1	9	8	7	3
f_{11}	Mean	1.6119E+03	1.4165E+03	1.5641E+03	1.2671E+03	1.3785E+03	2.0002E+04	7.1146E+03	1.1948E+03	1.1717E+03
	Std	1.9972E+02	7.6492E+01	1.1658E+02	4.1564E+01	1.0099E+02	2.1222E+03	1.1032E+03	4.6367E+01	2.4690E+01
	Best	1.3658E+03	1.2991E+03	1.3134E+03	1.2102E+03	1.2314E+03	1.4481E+04	5.1589E+03	1.1514E+03	1.1352E+03
	Rank	7	5	6	3	4	9	8	2	**1**
f_{12}	Mean	1.6116E+07	7.5670E+06	4.6460E+07	4.6598E+06	4.1380E+06	5.1212E+10	1.3160E+10	1.4403E+06	3.2523E+06
	Std	6.1552E+06	4.6226E+06	2.0397E+07	2.0030E+06	2.6387E+06	8.2247E+09	3.7722E+09	9.0879E+05	1.5757E+06
	Best	5.0264E+06	2.2004E+06	1.3216E+07	1.3912E+06	1.4595E+06	3.1126E+10	9.0912E+09	4.5131E+05	1.0010E+06
	Rank	6	5	7	4	3	9	8	1	2
f_{13}	Mean	2.5835E+04	2.8813E+04	4.6295E+04	1.0763E+04	3.0548E+04	2.8835E+10	3.2579E+09	4.9576E+03	5.7924E+03
	Std	6.5788E+03	2.2694E+04	2.8280E+04	1.0165E+04	2.3626E+04	6.0997E+09	7.9992E+08	4.3371E+03	2.5089E+03
	Best	1.4169E+04	3.7233E+03	1.4894E+04	2.5306E+03	8.2222E+03	1.3706E+10	1.8047E+09	1.4265E+03	3.2132E+03
	Rank	4	5	7	3	6	9	8	1	2
f_{14}	Mean	8.4709E+05	1.4273E+05	6.3328E+04	1.3600E+05	6.3258E+04	3.2999E+07	4.4556E+06	8.9367E+04	9.2519E+04
	Std	5.3562E+05	8.6058E+04	6.1835E+04	1.0280E+05	5.5343E+04	1.4235E+07	1.1044E+06	4.6108E+04	6.3578E+04
	Best	4.2088E+04	2.7848E+04	3.0987E+03	7.0458E+03	5.0904E+03	1.0471E+07	2.4029E+06	2.1208E+04	1.3495E+04
	Rank	7	6	2	5	1	9	8	3	4
f_{15}	Mean	8.3284E+03	1.6451E+04	1.2670E+04	1.3815E+04	1.0058E+04	4.5098E+09	4.4602E+08	5.3096E+03	7.7054E+03
	Std	4.5195E+03	1.5976E+04	6.3548E+03	6.4991E+03	5.6581E+03	1.0091E+09	1.6977E+08	4.2685E+03	3.9065E+03
	Best	2.8049E+03	2.9758E+03	2.7942E+03	2.1588E+03	3.7998E+03	2.8857E+09	1.6782E+08	1.5836E+03	2.0484E+03
	Rank	3	7	5	6	4	9	8	1	2
f_{16}	Mean	3.1906E+03	3.8253E+03	3.5006E+03	3.2990E+03	2.8448E+03	7.9349E+03	5.5097E+03	2.9038E+03	2.6354E+03
	Std	3.1995E+02	1.0916E+03	4.7662E+02	3.6932E+02	2.6655E+02	5.5200E+02	1.9917E+02	6.3660E+02	3.6722E+02
	Best	2.4603E+03	2.6894E+03	2.3230E+03	2.6676E+03	2.4691E+03	6.1835E+03	4.8928E+03	2.0434E+03	2.1466E+03

CONTINUED TABLE IV

f(x)	Index	TS22	HBA	FDA	AHA	SO	BWO	HGSO	QHGSO	E_HGSO
	Rank	4	7	6	5	2	9	8	3	**1**
f_{17}	Mean	3.0495E+03	2.8624E+03	3.2692E+03	3.0469E+03	2.7875E+03	6.4585E+03	4.2023E+03	2.8718E+03	2.5536E+03
	Std	2.2310E+02	3.0824E+02	3.7483E+02	3.3566E+02	2.6265E+02	8.9976E+02	2.1853E+02	5.8821E+02	2.3676E+02
	Best	2.4456E+03	2.3555E+03	2.5656E+03	2.5084E+03	2.2784E+03	4.9800E+03	3.6563E+03	1.9436E+03	2.1257E+03
	Rank	6	3	7	5	2	9	8	4	**1**
f_{18}	Mean	2.2132E+06	1.2029E+06	9.5095E+05	1.0108E+06	1.0142E+06	8.7331E+07	1.9173E+07	1.4743E+06	1.4039E+06
	Std	1.3360E+06	1.0278E+06	5.9740E+05	7.9327E+05	8.6533E+05	2.9390E+07	8.5574E+06	8.5454E+05	7.5519E+05
	Best	5.0565E+05	2.8802E+05	2.1752E+05	2.4760E+05	2.4156E+05	3.4373E+07	8.5363E+06	3.0749E+05	3.0248E+05
	Rank	7	4	1	2	3	9	8	6	5
f_{19}	Mean	1.1528E+04	2.3383E+04	1.8753E+04	1.6992E+04	1.5841E+04	2.4997E+09	2.2793E+08	1.7643E+04	1.4289E+04
	Std	2.9814E+03	1.4364E+04	1.1501E+04	9.8349E+03	1.5957E+04	7.2561E+08	9.3550E+07	1.1760E+04	6.5288E+03
	Best	4.2014E+03	2.6092E+03	2.6696E+03	2.1213E+03	2.9528E+03	1.1926E+09	8.6019E+07	2.0071E+03	4.0733E+03
	Rank	1	7	6	4	3	9	8	5	2
f_{20}	Mean	2.9869E+03	2.9890E+03	3.5321E+03	3.1207E+03	2.7037E+03	3.9342E+03	3.6540E+03	2.9608E+03	2.5778E+03
	Std	2.5001E+02	3.1864E+02	3.7418E+02	3.2832E+02	2.8703E+02	1.6928E+02	1.4943E+02	5.4669E+02	2.7560E+02
	Best	2.5153E+03	2.3050E+03	2.7759E+03	2.5320E+03	2.1832E+03	3.6085E+03	3.2723E+03	2.2289E+03	2.0796E+03
	Rank	4	5	7	6	2	9	8	3	**1**
f_{21}	Mean	2.6180E+03	2.5207E+03	2.5829E+03	2.5493E+03	2.4342E+03	3.1205E+03	2.9263E+03	2.4116E+03	2.3594E+03
	Std	3.1330E+01	4.0912E+01	5.4237E+01	4.5316E+01	2.1891E+01	5.7694E+01	4.1259E+01	2.2539E+01	1.0081E+01
	Best	2.5494E+03	2.4208E+03	2.4859E+03	2.4402E+03	2.3912E+03	2.9650E+03	2.7915E+03	2.3737E+03	2.3450E+03
	Rank	7	4	6	5	3	9	8	2	**1**
f_{22}	Mean	9.3129E+03	8.6458E+03	1.0596E+04	7.8216E+03	7.1586E+03	1.6296E+04	1.2945E+04	9.4631E+03	7.0911E+03
	Std	1.4795E+03	3.0252E+03	1.9010E+03	2.6627E+03	1.4696E+03	3.7052E+02	2.6327E+03	6.4236E+03	1.9536E+03
	Best	2.3540E+03	2.3199E+03	2.4252E+03	2.3017E+03	5.5213E+03	1.5691E+04	8.3125E+03	2.3000E+03	2.3001E+03
	Rank	5	4	7	3	2	9	8	6	**1**
f_{23}	Mean	3.1938E+03	2.9620E+03	3.0675E+03	3.0490E+03	2.9350E+03	4.0204E+03	3.8985E+03	2.8382E+03	2.7931E+03
	Std	7.0900E+01	4.1482E+01	9.1987E+01	7.4212E+01	3.0679E+01	7.9712E+01	1.5272E+02	4.0545E+01	1.8156E+01
	Best	3.0774E+03	2.8762E+03	2.9254E+03	2.9098E+03	2.8803E+03	3.8146E+03	3.4982E+03	2.7831E+03	2.7617E+03
	Rank	7	4	6	5	3	9	8	2	**1**
f_{24}	Mean	3.5627E+03	3.4472E+03	3.2206E+03	3.2941E+03	3.0675E+03	4.3521E+03	4.1219E+03	3.0622E+03	2.9644E+03
	Std	1.3761E+02	5.9303E+02	6.3170E+01	7.5164E+01	2.9529E+01	8.7614E+01	6.0039E+01	5.0885E+01	2.2422E+01
	Best	3.2147E+03	3.0312E+03	3.1160E+03	3.1319E+03	3.0221E+03	4.1185E+03	3.9832E+03	2.9910E+03	2.9112E+03
	Rank	7	6	4	5	3	9	8	2	**1**
f_{25}	Mean	3.1057E+03	3.0896E+03	3.1764E+03	3.1120E+03	3.0620E+03	1.3505E+04	6.9611E+03	3.0501E+03	3.0601E+03
	Std	2.4189E+01	2.7816E+01	5.3282E+01	2.5608E+01	2.5707E+01	6.7129E+02	7.9469E+02	2.8214E+01	2.6931E+01
	Best	3.0577E+03	3.0420E+03	3.0831E+03	3.0501E+03	3.0193E+03	1.1304E+04	6.1236E+03	2.9297E+03	3.0209E+03
	Rank	5	4	7	6	3	9	8	1	2
f_{26}	Mean	3.1577E+03	4.7228E+03	7.1967E+03	5.8833E+03	5.8565E+03	1.6161E+04	1.0898E+04	5.1183E+03	4.5526E+03
	Std	9.4806E+01	1.3621E+03	6.6441E+02	3.0580E+03	3.2590E+02	4.1904E+02	1.0232E+03	2.2644E+02	2.4157E+02
	Best	3.0293E+03	3.1195E+03	6.1208E+03	2.9081E+03	5.3409E+03	1.5265E+04	1.0020E+04	4.7040E+03	4.1860E+03
	Rank	1	3	7	6	5	9	8	4	2
f_{27}	Mean	3.5027E+03	3.9003E+03	3.6590E+03	3.6014E+03	3.6401E+03	5.6367E+03	5.3347E+03	3.3762E+03	3.4034E+03
	Std	8.4208E+01	7.8698E+02	1.2221E+02	9.5715E+01	8.5920E+01	4.0946E+02	2.8236E+02	3.6152E+01	3.6732E+01
	Best	3.3671E+03	3.3115E+03	3.4225E+03	3.4486E+03	3.4690E+03	4.7115E+03	4.6989E+03	3.3049E+03	3.3175E+03
	Rank	3	7	6	4	5	9	8	1	2
f_{28}	Mean	3.3714E+03	4.3380E+03	3.5539E+03	3.3881E+03	3.3297E+03	1.1742E+04	6.9509E+03	3.3063E+03	3.3363E+03
	Std	2.8370E+01	3.0160E+03	7.9831E+01	4.5053E+01	1.8239E+01	6.1617E+02	5.3019E+02	1.9577E+01	2.1871E+01
	Best	3.3051E+03	3.2859E+03	3.4227E+03	3.2743E+03	3.3086E+03	9.9325E+03	6.1638E+03	3.2590E+03	3.3079E+03
	Rank	4	7	6	5	2	9	8	1	3
f_{29}	Mean	4.2785E+03	5.3666E+03	4.9793E+03	4.3811E+03	4.3433E+03	1.7498E+04	8.0596E+03	3.8978E+03	3.4900E+03
	Std	2.5635E+02	2.9291E+03	3.4847E+02	3.0967E+02	2.4368E+02	5.9182E+03	5.5613E+02	2.7497E+02	1.5421E+02
	Best	3.7064E+03	3.3612E+03	4.3510E+03	3.8095E+03	3.9506E+03	9.3514E+03	6.8548E+03	3.3479E+03	3.2554E+03
	Rank	3	7	6	5	4	9	8	2	**1**
f_{30}	Mean	1.5121E+06	2.4336E+06	7.2343E+06	9.3961E+05	1.5133E+06	3.6086E+09	6.6135E+08	1.1778E+06	1.0216E+06
	Std	2.1001E+05	1.7890E+06	4.2263E+06	1.1369E+05	3.5614E+05	9.5451E+08	1.4114E+08	1.8966E+05	1.2623E+05
	Best	1.0812E+06	8.7322E+05	1.8348E+06	6.8884E+05	9.4273E+05	2.1293E+09	3.7387E+08	8.3445E+05	8.0812E+05
	Rank	4	6	7	1	5	9	8	3	2
Total Rank		146	148	171	125	95	261	232	77	50
Final Rank		5	6	7	4	3	9	8	2	**1**

TABLE V
RESULTS OF COMPARING ALGORITHMS ON THE CEC2017 BENCHMARK FUNCTION (D=100)

f(x)	Index	TS22	HBA	FDA	AHA	SO	BWO	HGSO	QHGSO	E_HGSO
f_1	Mean	1.7726E+09	3.5173E+09	2.2732E+10	2.4507E+08	1.2680E+07	2.6722E+11	1.6462E+11	5.3570E+04	3.9760E+04
	Std	2.1151E+08	1.7941E+09	3.8859E+09	1.0235E+09	7.4599E+06	6.3959E+09	1.7326E+10	1.2800E+05	1.1985E+04
	Best	1.3219E+09	1.1495E+09	1.6677E+10	2.1467E+07	3.6474E+06	2.4629E+11	1.2459E+11	2.9385E+03	1.9781E+04
	Rank	5	6	7	4	3	9	8	2	**1**
f_3	Mean	3.0394E+05	4.0491E+05	2.5760E+05	1.9383E+05	3.1806E+05	4.5509E+05	3.2674E+05	3.3068E+05	3.3940E+05
	Std	1.3798E+04	5.6939E+04	2.8372E+04	1.5861E+04	2.1759E+04	1.4224E+05	1.3648E+04	3.7037E+04	4.3430E+04
	Best	2.7413E+05	3.2543E+05	2.0468E+05	1.5905E+05	2.2964E+05	3.4811E+05	2.9964E+05	2.6598E+05	2.2158E+05
	Rank	3	8	2	1	4	9	5	6	7
f_4	Mean	1.0839E+03	1.2651E+03	2.2510E+03	1.0215E+03	7.5159E+02	1.0895E+05	3.2301E+04	7.3367E+02	6.9432E+02
	Std	4.6361E+01	1.1842E+02	4.1669E+02	9.2692E+01	5.4313E+01	7.6207E+03	6.4003E+03	4.3995E+01	5.1198E+01
	Best	9.7748E+02	1.0201E+03	1.4717E+03	8.7738E+02	6.8254E+02	8.9847E+04	1.9248E+04	6.4728E+02	5.5317E+02
	Rank	5	6	7	4	3	9	8	2	**1**
f_5	Mean	1.4484E+03	1.2212E+03	1.3723E+03	1.3094E+03	8.3998E+02	2.1507E+03	1.8997E+03	9.7710E+02	6.7716E+02
	Std	4.7078E+01	5.8395E+01	1.0722E+02	6.2002E+01	3.4476E+01	2.9606E+01	4.3962E+01	1.2702E+02	3.0084E+01
	Best	1.3085E+03	1.1248E+03	1.1912E+03	1.1551E+03	7.5311E+02	2.0824E+03	1.8102E+03	8.0335E+02	6.0649E+02
	Rank	7	4	6	5	2	9	8	3	**1**
f_6	Mean	6.5688E+02	6.3517E+02	6.7298E+02	6.2679E+02	6.1938E+02	7.1548E+02	6.9998E+02	6.0755E+02	6.0021E+02
	Std	4.4501E+00	7.4057E+00	4.5269E+00	9.6939E+00	3.6834E+00	2.5224E+00	3.5140E+00	1.7970E+00	2.4678E-01
	Best	6.4716E+02	6.2015E+02	6.6298E+02	6.1012E+02	6.1099E+02	7.1099E+02	6.9066E+02	6.0353E+02	6.0006E+02
	Rank	6	5	7	4	3	9	8	2	**1**
f_7	Mean	2.6060E+03	2.2256E+03	2.5479E+03	2.6046E+03	1.2477E+03	3.9613E+03	3.4275E+03	1.4678E+03	9.8848E+02
	Std	1.3071E+02	2.1272E+02	1.8786E+02	3.3353E+02	6.3119E+01	6.4586E+01	1.7519E+02	1.1702E+02	3.1311E+01
	Best	2.4371E+03	1.8074E+03	2.1711E+03	2.0511E+03	1.1732E+03	3.8064E+03	3.0191E+03	1.1790E+03	9.1796E+02
	Rank	7	4	5	6	2	9	8	3	**1**
f_8	Mean	1.8343E+03	1.5253E+03	1.7015E+03	1.6966E+03	1.1408E+03	2.6506E+03	2.2970E+03	1.2690E+03	9.7842E+02
	Std	5.4774E+01	7.5541E+01	1.0187E+02	1.0583E+02	3.7740E+01	4.3042E+01	5.3760E+01	1.2094E+02	3.5206E+01
	Best	1.6494E+03	1.4216E+03	1.4869E+03	1.4809E+03	1.0692E+03	2.5130E+03	2.1761E+03	1.1044E+03	9.2141E+02
	Rank	7	4	6	5	2	9	8	3	**1**
f_9	Mean	4.8013E+04	5.2471E+04	4.1744E+04	2.4169E+04	7.0573E+03	8.4510E+04	7.1145E+04	1.8151E+04	1.0400E+03
	Std	2.2842E+03	7.0270E+03	6.3714E+03	9.9671E+02	1.9728E+03	3.5847E+03	4.2620E+03	1.2408E+04	1.1433E+02
	Best	4.4275E+04	3.8119E+04	2.5576E+04	2.1215E+04	4.7179E+03	7.7406E+04	6.2679E+04	5.1367E+03	9.2627E+02
	Rank	6	7	5	4	2	9	8	3	**1**
f_{10}	Mean	1.8675E+04	2.4699E+04	2.1400E+04	1.4678E+04	2.6690E+04	3.3222E+04	2.8846E+04	3.1308E+04	1.3945E+04
	Std	9.6106E+02	4.5190E+03	1.2156E+03	1.5798E+03	2.7706E+03	6.7858E+02	1.0260E+03	5.6450E+02	1.6036E+03
	Best	1.7126E+04	1.6571E+04	1.8908E+04	1.0915E+04	2.2147E+04	3.1759E+04	2.7002E+04	3.0306E+04	1.0113E+04
	Rank	3	5	4	2	6	9	7	8	**1**
f_{11}	Mean	2.5026E+04	4.6736E+04	3.2442E+04	2.8238E+04	3.0672E+04	4.6012E+05	1.3934E+05	7.8204E+03	3.0208E+03
	Std	6.5551E+03	9.5768E+03	7.2024E+03	1.1408E+04	7.8230E+03	1.4909E+05	1.2970E+04	4.5285E+03	5.3217E+02
	Best	1.2777E+04	3.5583E+04	2.1162E+04	1.1397E+04	1.5406E+04	3.1723E+05	1.0192E+05	3.1920E+03	2.2823E+03
	Rank	3	7	6	4	5	9	8	2	**1**
f_{12}	Mean	3.2149E+08	2.2163E+08	1.7624E+09	6.3639E+07	9.6005E+07	2.1173E+11	6.7385E+10	1.4993E+07	2.2154E+07
	Std	7.5053E+07	5.2044E+07	4.7361E+08	3.3909E+07	4.2567E+07	1.1253E+10	1.4989E+10	5.4846E+06	7.9968E+06
	Best	2.1656E+08	1.3043E+08	1.0374E+09	2.0765E+07	2.2066E+07	1.9005E+11	3.3262E+10	4.1786E+06	1.0128E+07
	Rank	6	5	7	3	4	9	8	1	2
f_{13}	Mean	8.5973E+05	1.0463E+05	1.1096E+07	3.7974E+04	6.6382E+04	4.8257E+10	9.9903E+09	5.4165E+03	1.2775E+04
	Std	2.5913E+05	3.3862E+05	8.7010E+06	2.4206E+04	6.5433E+04	4.3388E+09	2.8021E+09	3.7229E+03	2.9283E+03
	Best	4.8839E+05	6.6599E+03	1.3589E+06	9.0915E+03	2.0203E+04	3.4646E+10	5.2612E+09	1.7174E+03	8.4317E+03
	Rank	6	5	7	3	4	9	8	1	2
f_{14}	Mean	3.6613E+06	2.0243E+06	1.6768E+06	1.4545E+06	1.5570E+06	1.0032E+08	2.1829E+07	1.1659E+06	9.6420E+05
	Std	8.2599E+05	7.3598E+05	9.4981E+05	6.3300E+05	7.1287E+05	2.9933E+07	4.4292E+06	4.9469E+05	2.5516E+05
	Best	1.4239E+06	8.6491E+05	7.1055E+05	5.7804E+05	4.8115E+05	3.6724E+07	1.3871E+07	5.3021E+05	4.6235E+05
	Rank	7	6	5	3	4	9	8	2	**1**
f_{15}	Mean	5.1989E+04	1.2810E+04	1.0310E+05	7.5219E+03	1.9898E+04	2.5784E+10	2.7497E+09	3.7161E+03	4.7152E+03
	Std	1.3390E+04	1.1944E+04	5.5101E+04	5.9482E+03	2.0820E+04	3.2482E+09	8.0343E+08	2.9103E+03	1.2172E+03
	Best	2.0434E+04	2.7659E+03	1.9195E+04	2.1367E+03	6.4179E+03	1.5241E+10	1.0984E+09	1.7239E+03	3.1211E+03
	Rank	6	4	7	3	5	9	8	1	2
f_{16}	Mean	5.7828E+03	5.5894E+03	6.8122E+03	5.6335E+03	5.7068E+03	2.4343E+04	1.3527E+04	7.3350E+03	4.8470E+03
	Std	3.9898E+02	7.8154E+02	4.3684E+02	6.2622E+02	1.5711E+03	2.3693E+03	9.2290E+02	2.0435E+03	6.0997E+02
	Best	4.9386E+03	3.9835E+03	5.5973E+03	4.3265E+03	3.9297E+03	1.8888E+04	1.2046E+04	4.3045E+03	3.4979E+03

CONTINUED TABLE V

f(x)	Index	TS22	HBA	FDA	AHA	SO	BWO	HGSO	QHGSO	E_HGSO
	Rank	5	2	6	3	4	9	8	7	**1**
f_{17}	Mean	5.1413E+03	5.3263E+03	5.6631E+03	5.1050E+03	4.4345E+03	1.1590E+07	2.1449E+04	5.6181E+03	4.1470E+03
	Std	4.1836E+02	4.7711E+02	7.2421E+02	5.1426E+02	4.4748E+02	5.7951E+06	6.8468E+03	1.1997E+03	4.2745E+02
	Best	4.3593E+03	4.5342E+03	3.9741E+03	4.3004E+03	3.7082E+03	3.5977E+06	1.0315E+04	3.6412E+03	3.3026E+03
	Rank	4	5	7	3	2	9	8	6	**1**
f_{18}	Mean	3.3650E+06	5.1964E+06	2.5429E+06	2.3641E+06	3.8067E+06	3.0654E+08	3.1638E+07	5.0683E+06	1.5032E+06
	Std	9.3784E+05	2.4934E+06	1.2266E+06	1.1118E+06	1.6241E+06	1.1982E+08	6.7661E+06	2.7180E+06	6.5392E+05
	Best	2.0190E+06	2.6301E+06	5.6205E+05	9.5083E+05	1.4691E+06	1.1891E+08	1.6461E+07	8.5633E+05	6.4898E+05
	Rank	4	7	3	2	5	9	8	6	**1**
f_{19}	Mean	6.7948E+04	7.5914E+03	5.6672E+05	6.4181E+04	2.9208E+04	2.5525E+10	2.7179E+09	4.5156E+03	3.4420E+03
	Std	2.6280E+04	5.8269E+03	3.0382E+05	2.9799E+05	3.3815E+04	2.8809E+09	8.4608E+08	2.7326E+03	9.5555E+02
	Best	1.9517E+04	2.8294E+03	1.6119E+05	2.6697E+03	2.6824E+03	1.9871E+10	1.1205E+09	1.9995E+03	2.4817E+03
	Rank	6	3	7	5	4	9	8	2	**1**
f_{20}	Mean	5.0845E+03	5.1537E+03	5.8457E+03	5.3403E+03	5.8916E+03	8.1957E+03	7.2190E+03	7.0476E+03	4.3539E+03
	Std	3.5569E+02	7.2627E+02	5.5335E+02	5.8969E+02	1.0532E+03	3.1273E+02	3.5415E+02	8.4330E+02	4.5727E+02
	Best	4.2581E+03	4.2233E+03	5.0135E+03	4.0012E+03	3.9208E+03	7.3360E+03	6.4405E+03	3.5074E+03	3.5120E+03
	Rank	2	3	5	4	6	9	8	7	**1**
f_{21}	Mean	3.4825E+03	3.0111E+03	3.2226E+03	3.0095E+03	2.7252E+03	4.9805E+03	4.1528E+03	2.7529E+03	2.4914E+03
	Std	1.1089E+02	6.4004E+01	1.1747E+02	1.0591E+02	3.9652E+01	9.2243E+01	1.2150E+02	9.4975E+01	2.3382E+01
	Best	3.1481E+03	2.9124E+03	2.9688E+03	2.8333E+03	2.5964E+03	4.8466E+03	3.9085E+03	2.5447E+03	2.4459E+03
	Rank	7	5	6	4	2	9	8	3	**1**
f_{22}	Mean	2.2198E+04	2.6029E+04	2.4321E+04	1.8954E+04	2.6926E+04	3.5759E+04	3.2324E+04	3.3746E+04	1.6135E+04
	Std	1.0250E+03	3.9443E+03	1.4889E+03	1.8456E+03	5.1805E+03	5.0026E+02	7.5458E+02	7.8877E+02	1.6242E+03
	Best	2.0381E+04	1.8461E+04	2.1234E+04	1.4953E+04	1.3887E+04	3.4771E+04	3.0450E+04	3.1804E+04	1.2001E+04
	Rank	3	5	4	2	6	9	7	8	**1**
f_{23}	Mean	3.8890E+03	3.5714E+03	3.7993E+03	3.3552E+03	3.3612E+03	6.3120E+03	5.8992E+03	3.2211E+03	3.0301E+03
	Std	1.0514E+02	4.6237E+02	1.3075E+02	7.7857E+01	7.6539E+01	2.2757E+02	1.9533E+02	8.0582E+01	4.6852E+01
	Best	3.7367E+03	3.3091E+03	3.6207E+03	3.1957E+03	3.2346E+03	5.7097E+03	5.5277E+03	3.1011E+03	2.9434E+03
	Rank	7	5	6	3	4	9	8	2	**1**
f_{24}	Mean	4.6043E+03	5.7436E+03	4.5666E+03	4.2651E+03	4.0243E+03	1.0048E+04	8.5517E+03	3.6864E+03	3.4739E+03
	Std	1.3620E+02	2.6233E+03	1.9387E+02	1.2696E+02	8.3001E+01	6.9047E+02	5.7942E+02	9.7917E+01	6.8061E+01
	Best	4.3165E+03	3.8663E+03	4.1474E+03	4.0076E+03	3.8627E+03	8.4398E+03	7.5678E+03	3.5423E+03	3.3600E+03
	Rank	6	7	5	4	3	9	8	2	**1**
f_{25}	Mean	3.7796E+03	3.9589E+03	4.8135E+03	3.6619E+03	3.4806E+03	2.9258E+04	1.4307E+04	3.3751E+03	3.3422E+03
	Std	6.5351E+01	1.6883E+02	3.8537E+02	9.1837E+01	5.9667E+01	1.4089E+03	1.1628E+03	6.0355E+01	5.0269E+01
	Best	3.6672E+03	3.7204E+03	4.1093E+03	3.5024E+03	3.3632E+03	2.5937E+04	1.1919E+04	3.2731E+03	3.2453E+03
	Rank	5	6	7	4	3	9	8	2	**1**
f_{26}	Mean	6.6174E+03	1.5536E+04	1.9250E+04	1.8556E+04	1.2427E+04	5.2073E+04	3.6299E+04	1.0164E+04	8.3507E+03
	Std	3.5888E+03	9.7665E+03	1.7691E+03	6.9410E+03	7.9055E+02	1.1669E+03	2.5906E+03	8.4197E+02	9.6413E+02
	Best	5.0745E+03	1.1930E+04	1.5960E+04	4.5742E+03	1.0609E+04	5.0036E+04	3.1551E+04	8.0833E+03	6.9578E+03
	Rank	1	5	7	6	4	9	8	3	2
f_{27}	Mean	3.7339E+03	4.6181E+03	4.1877E+03	3.8974E+03	3.7520E+03	1.3273E+04	8.9646E+03	3.5951E+03	3.5317E+03
	Std	8.2458E+01	1.7103E+03	2.1511E+02	1.1643E+02	8.5377E+01	9.4258E+02	8.6002E+02	6.9604E+01	7.8054E+01
	Best	3.5744E+03	3.5136E+03	3.7237E+03	3.6722E+03	3.6040E+03	1.0945E+04	7.1254E+03	3.4406E+03	3.4310E+03
	Rank	3	7	6	5	4	9	8	2	**1**
f_{28}	Mean	3.8725E+03	4.3229E+03	6.3023E+03	3.9012E+03	3.6506E+03	2.8698E+04	2.0340E+04	3.5100E+03	3.5193E+03
	Std	6.5058E+01	2.9966E+02	7.7313E+02	2.2291E+02	4.3363E+01	1.0727E+03	2.1194E+03	3.4814E+01	3.0408E+01
	Best	3.7453E+03	3.8311E+03	5.0351E+03	3.6507E+03	3.5358E+03	2.6477E+04	1.4376E+04	3.4498E+03	3.4406E+03
	Rank	4	6	7	5	3	9	8	1	2
f_{29}	Mean	7.5663E+03	6.8469E+03	9.2789E+03	7.0654E+03	6.9978E+03	8.0198E+05	2.0022E+04	6.2503E+03	5.0244E+03
	Std	3.3960E+02	3.9495E+02	8.3390E+02	7.3917E+02	5.5059E+02	3.2353E+05	3.2784E+03	6.7028E+02	6.4296E+02
	Best	6.8225E+03	6.1911E+03	7.7768E+03	5.6073E+03	5.6149E+03	2.1847E+05	1.4217E+04	5.0007E+03	3.9643E+03
	Rank	6	3	7	5	4	9	8	2	**1**
f_{30}	Mean	3.0198E+06	3.6023E+05	2.3943E+07	3.3790E+05	4.8637E+05	4.3833E+10	8.7692E+09	2.7390E+04	9.9690E+04
	Std	5.6612E+05	2.0417E+05	1.1861E+07	2.1903E+05	2.6730E+05	4.1187E+09	2.9267E+09	1.9346E+04	4.0134E+04
	Best	1.8714E+06	1.2271E+05	8.0270E+06	6.3376E+04	1.5702E+05	3.0248E+10	5.2394E+09	8.3717E+03	4.7054E+04
	Rank	6	4	7	3	5	9	8	1	2
Total Rank		146	149	171	109	108	261	227	93	41
Final Rank		5	6	7	4	3	9	8	2	**1**

TABLE VIII
MEAN VALUES OF WILCOXON SIGNED RANK TEST ON CEC2017 BENCHMARK FUNCTIONS.

E_HGSO vs	Dimension							
	50				100			
	p-Value	*R*+	*R*-	+/=/-	*p*-Value	*R*+	*R*-	+/=/-
TS22	6.82E-03	54.72	410.28	26/2/1	9.20E-06	26.66	438.34	28/0/1
HBA	4.57E-02	122.00	343.00	24/3/2	9.33E-04	42.31	422.69	27/0/2
FDA	6.82E-03	54.72	410.28	26/2/1	9.20E-06	26.66	438.34	28/0/1
AHA	2.21E-02	137.86	327.14	19/5/5	3.38E-03	74.90	390.10	26/1/2
SO	1.00E-01	144.38	320.62	19/7/3	2.59E-03	59.17	405.83	26/1/2
BWO	1.73E-06	0.00	465.00	29/0/0	2.49E-06	4.83	460.17	29/0/0
HGSO	1.75E-06	1.03	463.97	29/0/0	3.23E-03	14.14	450.86	28/1/0
QHGSO	1.51E-01	221.14	243.86	12/8/9	3.94E-02	169.62	295.38	19/5/5
Mean Value	4.16E-02	91.98	373.02	23/3.4/2.6	6.19E-03	52.29	412.71	26.4/1/1.6

According to the data presented in Tables IV and V, the E_HGSO demonstrated superior performance. For the dimension of 50, the E_HGSO achieved the best results in 16 out of the tested functions, and the second-best result in 8 other functions. For the dimension of 100, the E_HGSO was the best performer in 22 functions and the second-best in 6 functions. This indicates an even greater overall advantage of the E_HGSO at the higher dimension. Importantly, the E_HGSO exhibited the best overall results across both problem dimensions.

A visual comparison of the above algorithms in 50 and 100 dimensions is given in Table VI. E_HGSO always significantly outperforms the other algorithms and there is no worst result, indicating that the improved strategy has significantly improved the performance of HGSO. The results of comparative tests show that the proposed strategy in the E_HGSO algorithm is more helpful in improving the performance of HGSO in high dimensional space.

Table VII shows the results of the Friedman test for all the algorithms. The p-values for 50 and 100 dimensions are 7.4399e-34 and 1.0728e-32, respectively. Since the p-values are much less than 0.05, we consider the statistical results to be statistically significant. The average rank of the Friedman test for the E_HGSO is 1.7241 and 1.4138, respectively, which is better than the other compared algorithms. In contrast, the mean ranks of the HGSO algorithm are 8.0000 and 7.8276, respectively, which are significantly worse than those of the E_HGSO algorithm. This suggests that the proposed enhancement strategy greatly improves the performance of the original HGSO algorithm.

In Table VIII, all p-values are found to be less than 0.05, suggesting that the statistical results are significant. The results of the Wilcoxon signed rank test demonstrate a considerable advantage of E_HGSO algorithm in both 50- and 100-dimensional test functions. Thus, it can be concluded that E_HGSO significantly outperforms the comparison algorithms, with statistical significance.

V. APPLICATION OF E_HGSO FOR FEATURE SELECTION

Feature selection is also called Feature Subset Selection (FS) [42]. It refers to the process of selecting N features from the existing M features to optimise the specific index of the system, selecting some most effective features from the original features to reduce the dimensionality of the dataset, which is an important means to improve the performance of learning algorithms, and is also a critical data preprocessing step in pattern recognition [43].

Feature selection is an important pre-processing step in classification, regression, and other data mining applications, as it helps to avoid the adverse effects of noisy, misleading, and inconsistent features on model performance. As a global combinatorial optimization problem, researchers have employed metaheuristic algorithms to select the most relevant features, with the aim of simplifying and improving the quality of high-dimensional datasets. However, when employing a wide range of datasets with large feature sizes, these methods tend to suffer from local optimization problems due to the considerable solution space. In this study, we propose a new dimensionality reduction approach to improve classification accuracy by selecting significant features using the HGSO.

The choice of feature selection method depends on the specific problem, the characteristics of the dataset, and the computational resources available. Filter methods are generally faster and simpler, but may not capture complex feature interactions. Wrapper methods can effectively explore the feature space and identify the most relevant features, but can be computationally intensive. Embedded methods offer a balance between the two, integrating feature selection within the model training process.

When solving classification problems, not all features in a data set are relevant, and often only a small number of features are relevant and can help determine the classification goal. In the era of big data, these worthless irrelevant features present in huge datasets usually take up a larger content. Selecting a subset of features is the best solution to the above problems. Feature selection is a process that aims to find a subset of relevant features from the original set. It can be seen that the subset of relevant features contains all the selected features and the remaining features are unselected.

Therefore, for each feature, there are two possibilities, "1" for selected feature and "0" for unselected feature. The number of feature subsets is 2N - 1 when the feature space is N. This problem has long been shown to be NP-hard and it appears to be difficult to find an optimal solution from a set of 2N - 1.

There are two important metrics in solving the feature selection problem, one is the size of the feature subset, i.e., the number of selected features, and a smaller number of selected features indicates a better feature selection. The other is the accuracy of the classification target. When the classification accuracy rate is higher, it also indicates that the feature selection effect is better. Therefore, the feature selection problem can be regarded as a multi-objective problem.

Objective 1: Feature subset size. Based on the number of "1" in the statistics set, we can get the number of currently selected features, so the first measure is shown in Equation 18:

$$f_1(X)=\sum_{1}^{D}x_i \tag{18}$$

$$f_2(X)=\frac{1}{n}\sum_{1}^{n}\frac{Nerr}{Nall} \tag{19}$$

Where *Nerr* denotes the number of classification errors; *Nall* denotes the number of all classified samples. *n* denotes the cross-validation Parameters.

The use of simple and easy-to-implement classification algorithms in wrapping methods can result in a good subset of features that are also applicable to complex classification algorithms. Therefore, this paper introduces the *K-NN* method as a classifier [44].

A. Model building

From the perspective of intelligent optimization, the feature selection problem is to obtain a solution that minimises the subset of features and maximises the classification accuracy through the process of population evolution for a family of solution vectors whose dimensions are the number of features of the problem, represented by 0 and 1. When solving the feature selection problem with E_HGSO, a feature selection solution is equivalent to an individual of the E_HGSO algorithm. If the component of the solution is "0", the feature is not selected; if the component of the solution is "1", the feature is selected. The feature selection problem involves designing a fitness function that considers two competing objectives: minimizing the number of selected features and maximizing the classification accuracy. The fitness function can be defined as follows:

$$fitness=\xi\cdot\Delta_R(D)+\psi\cdot\frac{|Y|}{|T_Z|} \tag{20}$$

Where $\Delta_R(D)$ denotes the classification error rate using K Nearest Neighbours (KNN) classification error rate, $|Y|$ denotes the number of feature subsets selected by the E_HGSO algorithm, $|T_Z|$ denotes the total number of features contained in the current dataset, and ξ is a parameter related to the classification error rate weights, ξ, $\Psi\in[0, 1]$, and $\xi+\Psi=1$.

B. Data sets and performance metrics

The capacity of the E_HGSO algorithm to perform feature selection was evaluated by conducting experiments on 8 standard datasets obtained from the UCI Machine Learning Repository. The specific details of these datasets are presented in Table IX. For the evaluation, each dataset underwent max-min normalization, whereby the data were scaled to the range [0, 1]. Subsequently, each dataset was divided into training and test subsets. The feature subsets obtained for each individual were then classified using the K-Nearest Neighbors (KNN) classifier.

In this case, we will continue to use the same 8 algorithms as the comparative algorithms, with all their parameters set the same as in the previous experiments. The E_HGSO algorithm is initialized randomly, with a population size N of 10 and a maximum number of iterations set to 100. Each algorithm is run independently for 5 times, with the dimension of the search space set to the number of features in the respective data sets. The average accuracy, the average number of feature selections, the average fitness value and the standard deviation of the fitness value are used to evaluate the advantages and disadvantages of each algorithm in the feature selection problem, and to test the optimization performance of the E_HGSO algorithm in solving the feature selection problem.

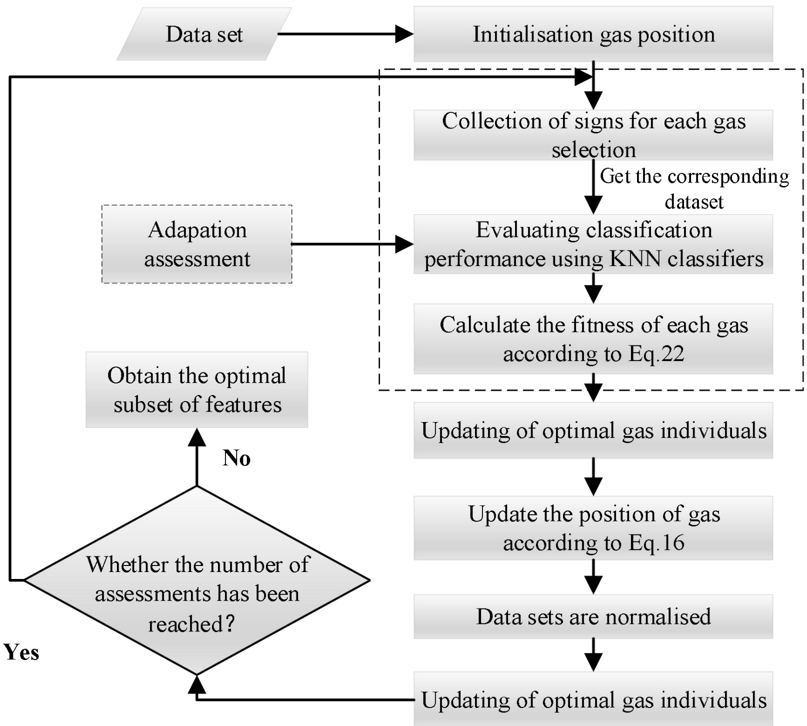

Fig.4. Flowchart for applying E_HGSO to feature selection

TABLE IX
CHARACTERISTICS OF THE DATA SET

No.	Data set	No. of samples	No. of features
1	arrhythmia	452	279
2	musk1	476	166
3	libras	309	90
4	connectionist	208	60
5	breast-cancer -wisc-diag	569	32
6	ionosphere	351	34
7	Sonar	208	60
8	cylinder-bands	512	39

Average accuracy: The average accuracy is the average of the classification accuracies of the optimization algorithm when performing feature selection and is defined as follows [41]:

$$AvgAcc = \frac{1}{Q}\sum_{i=1}^{Q} Acc_i \tag{21}$$

Where Q denotes the number of times the algorithm has been run, Acc_i denotes the optimal solution of classification accuracy obtained by the ith run of the algorithm, and the larger the value of $AvgAcc$, the better the classification. The value of k in the KNN classifier was set to 5 during the experimental analysis. So the adaptation is expressed as follows：

$$fiteness = \frac{0.95}{ACCcv5} + \frac{0.05 \times n_{sel}}{N_{tot}} \tag{22}$$

Where n_{sel} is the number of selected descriptors, N_{tot} is the total number of descriptors, and ACC_{cv5} is the precision of the five-fold cross-validation.

Average feature selection number: The average number of feature selections is the average of the number of features selected by the optimal solution obtained in Q runs of the algorithm over the total number of features in the dataset, which is expressed as follows Eq 23:

$$AvgSize = \frac{1}{Q}\sum_{i=1}^{Q} \frac{size_i}{D} \tag{23}$$

Where $size_i$ denotes the number of features selected by the optimal solution in the ith run, D is the total number of features in the original dataset, The smaller the value of $AvgSize$ (average number of selected features), the larger the value of $AvgAcc$ (average classification accuracy), which indicates that the algorithm performs better in the feature selection problem.

Average fitness: The average fitness is the average of the fitness of the optimal solution obtained from Q runs of the algorithm.

$$AvgAve = \frac{1}{Q}\sum_{i=1}^{Q} fitness_i \tag{24}$$

Where $fitness_i$ denotes the optimal solution fitness obtained in the ith run.

Standard deviation of fitness: The variance indicates the magnitude of volatility of the solution, the smaller its value, the more the algorithm can converge to the same value; the larger the value, the more the algorithm is volatile and unstable. The standard deviation is defined as:

$$AvgStd = \sqrt{\frac{1}{Q}\sum_{i=1}^{Q} (fitness_i - Ave)^2} \tag{25}$$

Where Ave denotes the average adaptation.

Table X presents the results of the E_HGSO algorithm and eight comparison algorithms for 8 datasets, including the mean and standard deviation of the fitness values and the optimal value. The best results are highlighted in bold. In terms of fitness values, a smaller value indicates better algorithm performance. Upon reviewing the results, it is evident that the E_HGSO algorithm consistently achieved the best results in all eight datasets. Additionally, the overall ranking of the E_HGSO algorithm is significantly ahead of the other algorithms. These findings underscore the superior performance and effectiveness of the E_HGSO algorithm in comparison to the alternative algorithms.

Fig.5 displays box plots representing the classification accuracies obtained by running the nine algorithms independently ten times on the eight datasets. These box plots provide a visual representation of the mean and dispersion of the data. From the figure, it is evident that the E_HGSO not only achieves high fitness on most datasets but also exhibits fewer outliers in the central distribution of results. This observation highlights the robustness of E_HGSO. The consistent high performance and reduced variability of the E_HGSO further validate its effectiveness.

In summary, our research results have conclusively demonstrated that the algorithm we proposed exhibits excellent performance in most cases. The algorithm not only outperforms the original algorithm, but also surpasses other comparative algorithms in optimizing feature subset selection and improving classification accuracy. These findings fully highlight the outstanding effectiveness and significant advantages of our algorithm in enhancing the efficiency and performance of feature selection and classification tasks.

Fig.6 illustrates the convergence curves of fitness for eight algorithms over 100 iterations. To facilitate a clearer comparison, we have excluded highly similar and redundant

TABLE X
MEAN FITNESS VALUES AND STANDARD DEVIATIONS OF DIFFERENT ALGORITHMS ON 8 DATA SETS

Data set		TS22	HBA	FDA	AHA	SO	BWO	HGSO	QHGSO	E_HGSO
arrhythmia	Mean	1.3094E+00	1.3954E+00	1.5258E+00	1.3344E+00	1.3388E+00	1.4117E+00	1.5751E+00	1.3408E+00	1.3135E+00
	Std	2.0163E-02	2.1968E-02	1.1775E-02	4.8746E-03	3.6802E-02	6.4036E-03	1.1940E-02	5.0713E-03	8.9191E-03
	Best	1.2841E+00	1.3651E+00	1.5080E+00	1.3310E+00	1.306242442	1.404880877	1.5530E+00	1.3312E+00	1.3086E+00
	Rank	1	6	8	3	4	7	9	5	2
musk1	Mean	1.0372E+00	1.0751E+00	1.0938E+00	1.0537E+00	1.0398E+00	1.0935E+00	1.1567E+00	1.0236E+00	1.0384E+00
	Std	5.4427E-03	2.1900E-02	1.5722E-02	1.3063E-02	3.5504E-03	4.0106E-03	9.0276E-03	7.6661E-03	5.4473E-03
	Best	1.0269E+00	1.0256E+00	1.0761E+00	1.0356E+00	1.033719302	1.082389558	1.1494E+00	1.0177E+00	1.0300E+00
	Rank	2	6	8	5	4	7	9	1	3
libras	Mean	1.1515E+00	1.1884E+00	1.2086E+00	1.1749E+00	1.1558E+00	1.1681E+00	1.2685E+00	1.1694E+00	1.1527E+00
	Std	7.4208E-03	1.3333E-02	1.3502E-02	1.0912E-02	4.6649E-03	1.3911E-02	1.4055E-03	5.0802E-03	3.0185E-03
	Best	1.1386E+00	1.1510E+00	1.1760E+00	1.1610E+00	1.145101513	1.161516517	1.2659E+00	1.1549E+00	1.1500E+00
	Rank	1	7	8	6	3	4	9	5	2
connectionist	Mean	1.0294E+00	1.0735E+00	1.0679E+00	1.0548E+00	1.0401E+00	1.0679E+00	1.1145E+00	1.0264E+00	1.0279E+00
	Std	7.8139E-03	2.5519E-02	1.4442E-02	8.3409E-03	1.0508E-02	1.8133E-02	2.2140E-03	2.6053E-03	4.2907E-03
	Best	1.0207E+00	1.0446E+00	1.0397E+00	1.0392E+00	1.027329932	1.042888307	1.1124E+00	1.0248E+00	1.0180E+00
	Rank	3	8	6	5	4	7	9	1	2
breast-cancer -wisc-diag	Mean	9.8498E-01	9.8916E-01	9.9478E-01	9.8585E-01	9.8298E-01	9.8794E-01	1.0079E+00	9.8903E-01	9.8242E-01
	Std	1.1621E-03	3.0614E-03	2.8661E-03	1.1678E-03	1.5567E-03	4.2784E-04	5.6406E-04	2.3617E-04	7.5014E-04
	Best	9.8388E-01	9.8582E-01	9.9063E-01	9.8373E-01	0.982060932	0.987701149	1.0077E+00	9.8888E-01	9.8206E-01
	Rank	3	7	8	4	2	5	9	6	**1**
ionosphere	Mean	1.0283E+00	1.0612E+00	1.0641E+00	1.0325E+00	1.0309E+00	1.0475E+00	1.1171E+00	1.0363E+00	1.0273E+00
	Std	1.9678E-03	8.2273E-03	1.1206E-02	3.9951E-03	3.2421E-03	5.5304E-03	1.0881E-02	9.5075E-05	2.3406E-16
	Best	1.0273E+00	1.0385E+00	1.0476E+00	1.0287E+00	1.0273E+00	1.038500183	1.1044E+00	1.0363E+00	1.0273E+00
	Rank	2	7	8	4	3	6	9	5	**1**
Sonar	Mean	1.0278E+00	1.0546E+00	1.0836E+00	1.0316E+00	1.0390E+00	1.0827E+00	1.1073E+00	1.0460E+00	1.0188E+00
	Std	1.0127E-02	7.5255E-03	2.7424E-03	1.5899E-03	8.4027E-03	4.2424E-04	3.1163E-02	7.7699E-03	4.1142E-03
	Best	1.0121E+00	1.0429E+00	1.0792E+00	1.0294E+00	1.0282E+00	1.0823E+00	1.0711E+00	1.0344E+00	1.0121E+00
	Rank	2	6	8	3	4	7	9	5	**1**
cylinder-bands	Mean	1.2181E+00	1.2767E+00	1.2726E+00	1.2537E+00	1.2302E+00	1.2628E+00	1.3500E+00	1.2264E+00	1.2046E+00
	Std	1.0030E-02	1.8628E-02	1.9601E-02	1.0850E-02	2.3308E-02	8.3031E-03	3.1029E-02	9.9729E-03	4.3574E-03
	Best	1.2064E+00	1.2321E+00	1.2483E+00	1.2305E+00	1.2079E+00	1.2477E+00	1.2862E+00	1.2167E+00	1.2021E+00
	Rank	2	8	7	5	4	6	9	3	**1**
Total Rank		16	55	61	35	28	49	72	31	13
Final Rank		2	7	8	5	3	6	9	4	**1**

curves, retaining four representative ones. The results clearly indicate that, across the majority of datasets, the three variants of the E_HGSO algorithm demonstrate notably faster convergence speeds compared to the other algorithms, ultimately reaching the lowest fitness values. Furthermore, the E_HGSO algorithm consistently exhibits the best performance by achieving the lowest fitness values. These findings highlight the superior convergence capabilities and overall effectiveness of the E_HGSO algorithm in optimizing fitness and finding optimal solutions.

VI. CONCLUSIONS AND PROSPECTS

This paper presents the utilization of the E_HGSO algorithm for solving the feature selection problem. The E_HGSO algorithm is an enhanced version of the HGSO algorithm, specifically designed to overcome the limitations of its predecessor. These limitations include insufficient population diversity, vulnerability to local optima, and slow convergence speed. The improvements made in this study focus on addressing these shortcomings. The new grouped search formula significantly expands the search range, resulting in improved efficiency and the ability to avoid local optima. The new position update formula introduces more diversity and randomness into the search process. The inclusion of the grouping mechanism based on the Lévy flight strategy further enhances the algorithm's ability to search effectively, which can better adapt to the search requirements of different distances. The E_HGSO algorithm has undergone a comprehensive comparison with eight other algorithms, E_HGSO consistently outperforms all other algorithms. Additional statistical evidence is presented

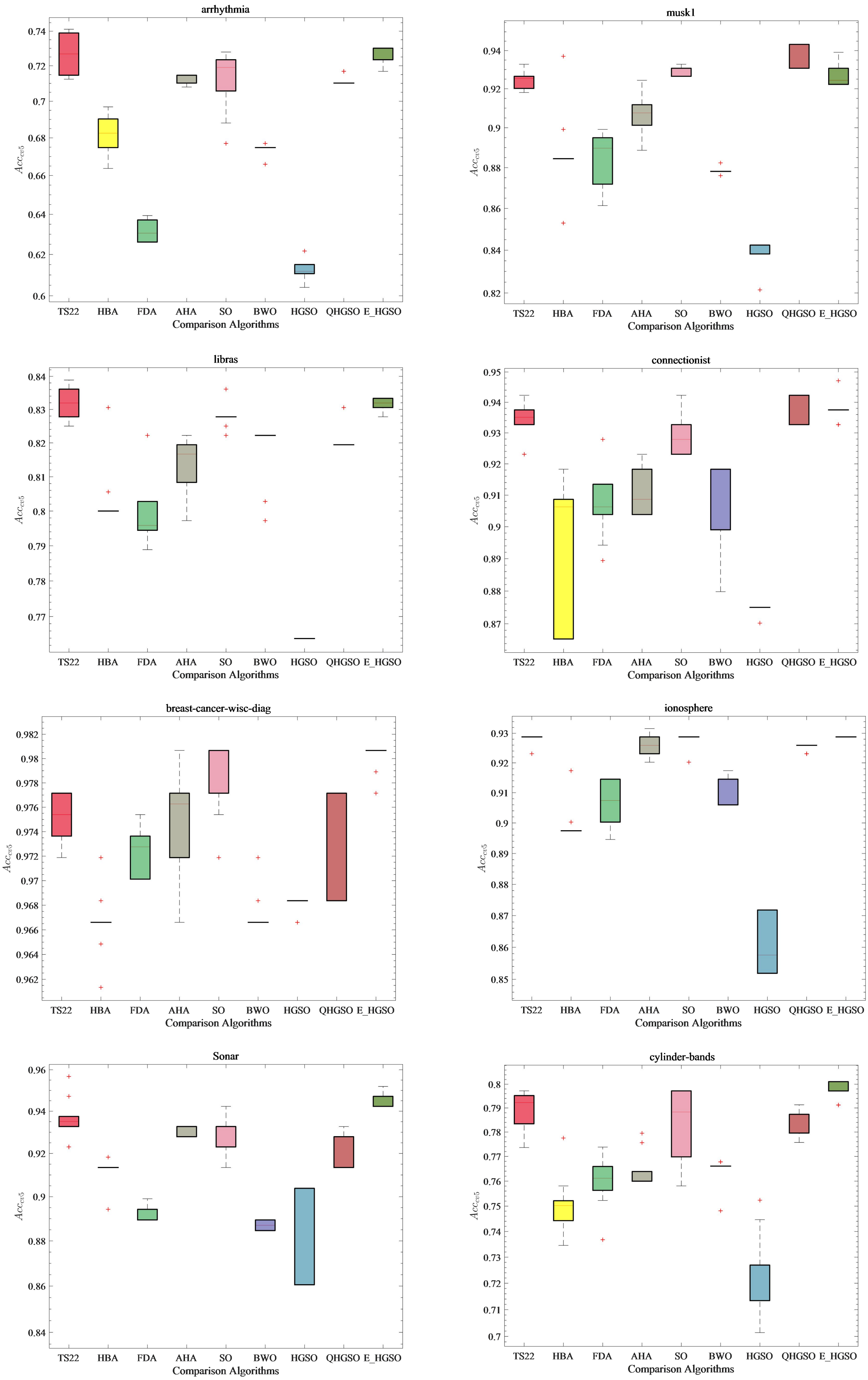

Fig.5. Accuracy boxplots of E_HGSO and other algorithms.

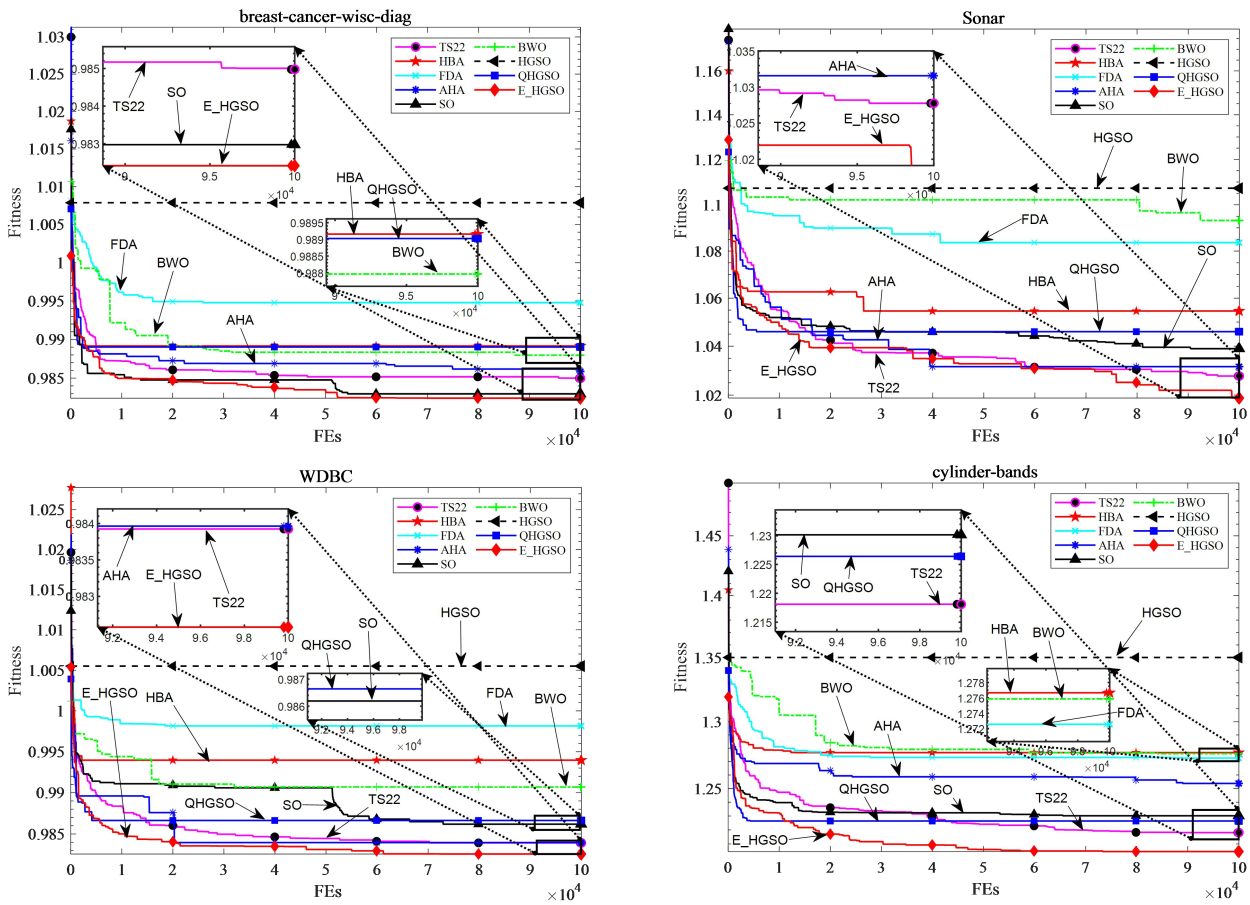

Fig.6. Convergence curve of E_HGSO and other algorithms.

through Friedman and Wilcoxon signed rank tests, which confirm the significant differences between the algorithms and reinforce the finding that E_HGSO consistently achieves superior results. Furthermore, the E_HGSO algorithm is applied to feature selection concurrently. To validate its performance in this context, eight datasets with varying dimensions and sizes are carefully selected from the UCI machine learning library. The experimental results demonstrate that E_HGSO exhibits superior efficiency, fitness, precision, and convergence speed when addressing feature selection problems.

In our ongoing research, we are dedicated to conducting a more comprehensive analysis and investigation of the HGSO algorithm. Our objective is to apply the algorithm to increasingly complex practical engineering problems. We strive to advance the HGSO algorithm, by combining comprehensive analysis, practical applications, and theoretical enhancements, we aim to provide valuable insights and solutions for addressing intricate optimization problems.

References

[1] O. H. Ibarra and C. E. Kim, "Heuristic algorithms for scheduling independent tasks on nonidentical processors," Journal of the ACM (JACM), vol. 24, no. 2, pp280-289, 1977.

[2] J. H. Holland, "Genetic algorithms," Scientific American, vol. 267, no. 1, pp66-73, 1992.

[3] S. Das and P. N. Suganthan, "Differential evolution: A survey of the state-of-the-art," IEEE transactions on evolutionary computation, vol. 15, no. 1, pp4-31, 2010.

[4] D. Wierstra, T. Schaul, T. Glasmachers, Y. Sun, J. Peters and J. Schmidhuber, "Natural evolution strategies," The Journal of Machine Learning Research, vol. 15, no. 1, pp949-980, 2014.

[5] X. Yao, Y. Liu, and G. Lin, "Evolutionary programming made faster," IEEE Trans. Evol. Computat, vol. 3, no. 2, pp82-102, 1999.

[6] J. Zhong, L. Feng and Y. S. Ong, "Gene expression programming: A survey," IEEE Computational Intelligence Magazine, vol. 12, no. 3, pp54-72, 2017.

[7] H. G. Beyer and B. Sendhoff, "Simplify your covariance matrix adaptation evolution strategy," IEEE Transactions on Evolutionary Computation, vol. 21, no. 5, pp746-759, 2017.

[8] D. Simon, "Biogeography-Based Optimization," IEEE Trans. Evol. Computat, vol. 12, no. 6, pp702-713, 2008.

[9] H. Sho, "Investigation of Particle Multi-Swarm Optimization with Diversive Curiosity," Engineering Letters, vol. 28, no. 3, pp960-969, 2020.

[10] A. Ghasemi-Marzbali, "A novel nature-inspired meta-heuristic algorithm for optimization: bear smell search algorithm," Soft computing, vol. 24, no. 17, pp13003-13035, 2020.

[11] C. Wang, P. Shang and L. Liu, "Improved artificial bee colony algorithm guided by experience," Engineering Letters, vol. 30, no. 1, pp261-265, 2022.

[12] M. A. Javarone, "An evolutionary strategy based on partial imitation for solving optimization problems," Physica A:Statistical Mechanics and its Applications, vol. 463, pp262-269, 2016.

[13] L. Sun, J. Li, X. Xiao, L. Zhang and J. Li, "Improved Differential Gray Wolf Algorithm Optimized Support Vector Regression Strip Thickness Prediction Method," Engineering Letters, vol. 29, no. 4, pp1462-1469, 2021.

[14] D. J. Kang, "A fast and stable snake algorithm for medical images," Pattern Recognition Letters, vol. 20, no. 5, pp507-512, 1999.

[15] J. Zhang, X. Xie and M. Wang, "A Hybrid Firefly Algorithm with Butterfly Optimization Algorithm and its Application," Engineering Letters, vol. 30, no. 2, pp453-462, 2022.

[16] F. A. Hashim, E. H. Houssein, K. Hussain, M. S. Mabrouk and W. Al-Atabany, "Honey Badger Algorithm: New metaheuristic algorithm for solving optimization problems," Mathematics and Computers in Simulation, vol. 192, pp84-110, 2022.

[17] R. V. Rao, V. J. Savsani and D. P. Vakharia, "Teaching-learning-based optimization: A novel method for constrained mechanical design optimization problems," Computer-Aided Design, vol. 43, no. 3, pp303-315, 2011.

[18] R. A. Gallego, R. Romero and A. J. Monticelli, "Tabu search algorithm for network synthesis," IEEE Transactions on Power Systems, vol. 15, no. 2, pp490-495, 2000.

[19] A. H. Kashan, "League Championship Algorithm (LCA): An algorithm for global optimization inspired by sport championships," Applied Soft Computing, vol. 16, pp171-200, 2014.
[20] C. Dai, W. Chen, Y. Song and Y. Zhu, "Seeker optimization algorithm: a novel stochastic search algorithm for global numerical optimization," Journal of Systems Engineering and Electronics, vol. 21, no. 2, pp300-311, 2010.
[21] N. Ghorbani and E. Babaei, "The exchange market algorithm with smart searching for solving economic dispatch problems," International Journal of Management Science and Engineering Management, vol. 13, no. 3, pp175-187, 2018.
[22] M. A. Eita and M. M. Fahmy, "Group counseling optimization," Applied Soft Computing, vol. 22, pp585-604, 2014.
[23] R. Cheng and Y. Jin, "A social learning particle swarm optimization algorithm for scalable optimization," Information Sciences, vol. 291, pp43-60, 2015.
[24] H. C. Kuo and C. H. Lin, "Cultural evolution algorithm for global optimizations and its applications," Journal of applied research and technology, vol. 11, no. 4, pp510-522, 2013.
[25] K. Salimifard, J. Li, D. Mohammadi and R. Moghdani, "A multi objective volleyball premier league algorithm for green scheduling identical parallel machines with splitting jobs," Applied Intelligence, vol. 51, no. 7, pp4143-4161, 2021.
[26] J. Derrac, S. Garc´ıa, D. Molina, and F. Herrera, "A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms," Swarm and Evolutionary Computation, vol. 1, no. 1, pp3-18, 2011.
[27] E. Rashedi, E. Rashedi and H. Nezamabadi-Pour, "A comprehensive survey on gravitational search algorithm," Swarm and evolutionary computation, vol. 41, pp141-158, 2018.
[28] O. K. Erol and I. Eksin, "A new optimization method: Big Bang-Big Crunch," Advances in Engineering Software, vol. 37, no. 2, pp106-111, 2006.
[29] E. Rashedi, H. Nezamabadi-pour, and S. Saryazdi, "GSA: A Gravitational Search Algorithm," Information Sciences, vol. 179, no. 13, pp2232-2248, 2009.
[30] R. A. Formato, "Central force optimization: A new deterministic gradient-like optimization metaheuristic," Opsearch, vol. 46, no. 1, pp25-51, 2009.
[31] A. Recioui, "Application of a galaxy-based search algorithm to MIMO system capacity optimization," Arabian Journal for Science and Engineering, vol. 41, no. 9, pp3407-3414, 2016.
[32] L. Abualigah, K. H. Almotairi, M. Abd Elaziz, M. Shehab and M. Altalhi, "Enhanced flow direction arithmetic optimization algorithm for mathematical optimization problems with applications of data clustering," Engineering Analysis with Boundary Elements, vol. 138, pp13-29, 2022.
[33] B. S. Yıldız, A. R. Yıldız, N. Pholdee, S. Bureerat, S. M. Sait and V. Patel, "The Henry gas solubility optimization algorithm for optimum structural design of automobile brake components," Materials Testing, vol. 62, no. 3, pp261-264, 2020.
[34] B. S. Yıldız, N. Pholdee, N. Panagant, S. Bureerat, A. R. Yildiz, and S. M. Sait, "A novel chaotic Henry gas solubility optimization algorithm for solving real-world engineering problems," Engineering with Computers, pp1-13, 2021.
[35] F. A. Hashim, E. H. Houssein, K. Hussain, M. S. Mabrouk and W. Al-Atabany, "A modified Henry gas solubility optimization for solving motif discovery problem," Neural Computing and Applications, vol. 32, no. 14, pp10759-10771, 2020.
[36] S. Ekinci, B. Hekimoğlu, and D. Izci, "Opposition based Henry gas solubility optimization as a novel algorithm for PID control of DC motor," Engineering Science and Technology, an International Journal, vol. 24, no. 2, pp331-342, 2021.
[37] W. Cao, X. Liu, and J. Ni, "Parameter Optimization of Support Vector Regression Using Henry Gas Solubility Optimization Algorithm," IEEE Access, vol. 8, pp88633-88642, 2020.
[38] D. Mohammadi, M. Abd Elaziz, R. Moghdani, E. Demir and S. Mirjalili, "Quantum Henry gas solubility optimization algorithm for global optimization," Engineering with Computers, pp1-20, 2021.
[39] J. Staudinger and P. V. Roberts, "A critical review of Henry's law constants for environmental applications," Critical Reviews in Environmental Science and Technology, vol. 26, no. 3, pp205-297, 1996.
[40] A. V. Chechkin, R. Metzler, J. Klafter and V. Y. Gonchar, "Introduction to the theory of Lévy flights," Anomalous transport: Foundations and applications, pp129-162, 2008.
[41] A. W. Mohamed, A. A. Hadi, A. M. Fattouh and K. M. Jambi, "LSHADE with semi-parameter adaptation hybrid with CMA-ES for solving CEC 2017 benchmark problems," IEEE, pp145-152, 2017.
[42] J. Li, K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang and H. Liu, "Feature selection: A data perspective," ACM computing surveys (CSUR), vol. 50, no. 6, pp1-45, 2017.
[43] H. Dai, W. Wu, J. Li, and Y. Yuan, "Incorporating Feature Selection in the Improved Stacking Algorithm for Online Learning Analysis and Prediction," Engineering Letters, vol. 28, no. 4, pp1011-1022, 2020.
[44] P. Cunningham and S. J. Delany, "k-Nearest neighbour classifiers-A Tutorial," ACM computing surveys (CSUR), vol. 54, no. 6, pp1-25, 2021.