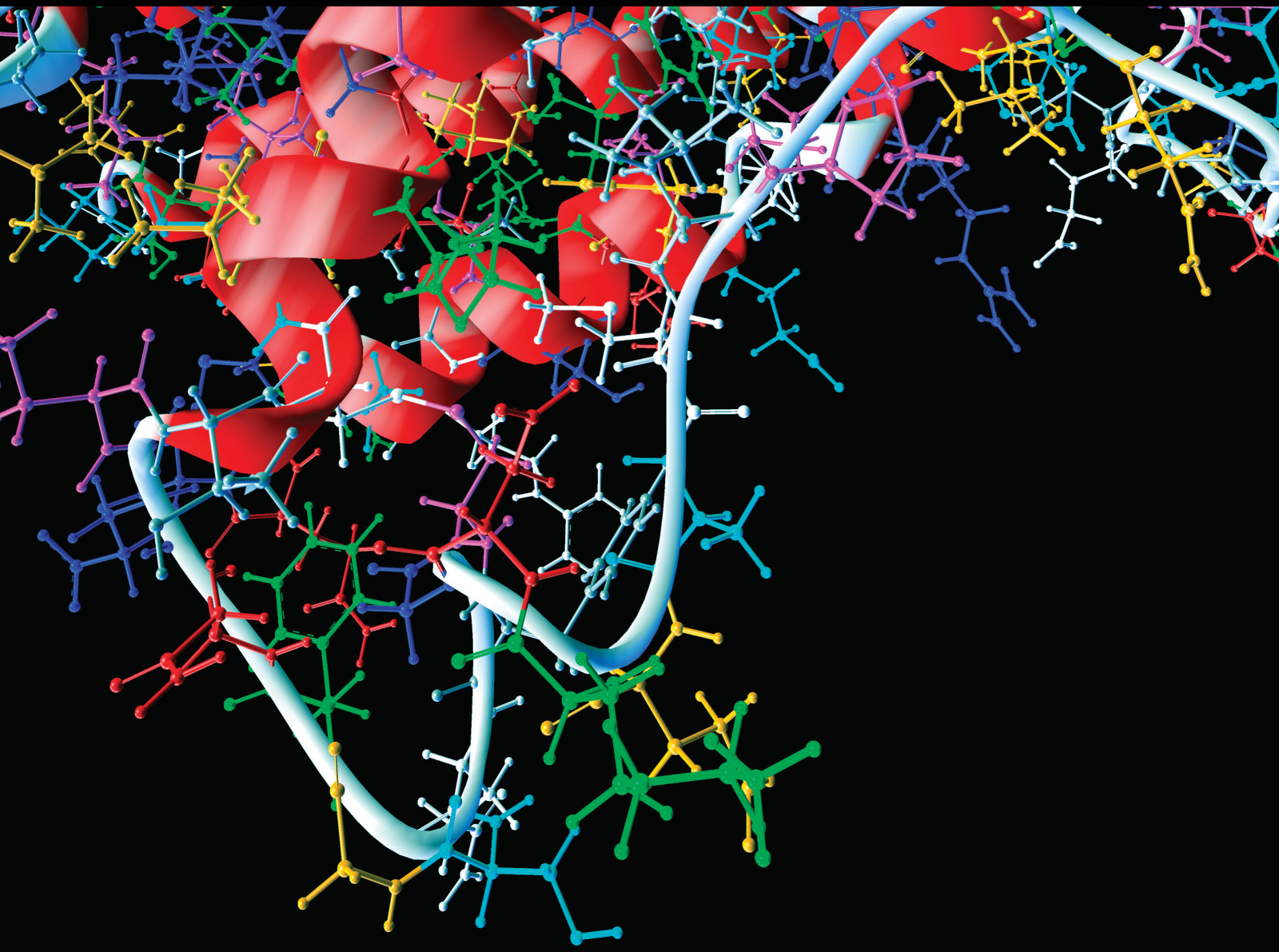
Computational and Mathematical Methods in Medicine

# Deep Learning and Machine Learning for Improving E-health and M-health

Lead Guest Editor: Kehua Guo
Guest Editors: Entao Luo, Kevin Wang, and Lu Liu



# Deep Learning and Machine Learning for Improving E-health and M-health

Computational and Mathematical Methods in Medicine

# Deep Learning and Machine Learning for Improving E-health and M-health

Lead Guest Editor: Kehua Guo
Guest Editors: Entao Luo, Kevin Wang, and Lu Liu

# Contents

**Boosting Slime Mould Algorithm for High-Dimensional Gene Data Mining: Diversity Analysis and Feature Selection**
Feng Qiu, Ran Guo, Huiling Chen, and Guoxi Liang
Research Article (27 pages), Article ID 8011003, Volume 2022 (2022)

**Primary Investigation of Deep Learning Models for Japanese "Group Classification" of Whole-Slide Images of Gastric Endoscopic Biopsy**
Xiaoya Fan, Lihui Yu, Xin Qi, Xue Lin, Junjun Zhao, Dong Wang, and Jing Zhang
Research Article (9 pages), Article ID 6899448, Volume 2022 (2022)

**Optimal Modeling of Anti-Breast Cancer Candidate Drugs Based on Graph Model Feature Selection**
Rongyuan Chen, Zhixiong He, Shaonian Huang, Lizhi Shen, and Xiancheng Zhou
Research Article (12 pages), Article ID 8418048, Volume 2022 (2022)

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 8011003, 27 pages
https://doi.org/10.1155/2022/8011003

*Research Article*

# Boosting Slime Mould Algorithm for High-Dimensional Gene Data Mining: Diversity Analysis and Feature Selection

**Feng Qiu**,[1] **Ran Guo**,[2] **Huiling Chen**,[1] **and Guoxi Liang**[3]

[1]*Department of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou 325035, China*
[2]*Cyberspace Institute Advanced Technology, Guangzhou University, Guangzhou 510006, China*
[3]*Department of Information Technology, Wenzhou Polytechnic, Wenzhou 325035, China*

Correspondence should be addressed to Ran Guo; guoran20@163.com, Huiling Chen; chenhuiling.jlu@gmail.com, and Guoxi Liang; guoxiliang2017@gmail.com

Received 28 July 2022; Revised 26 August 2022; Accepted 30 August 2022; Published 13 October 2022

Academic Editor: Kehua Guo



Slime mould algorithm (SMA) is a new metaheuristic algorithm, which simulates the behavior and morphology changes of slime mould during foraging. The slime mould algorithm has good performance; however, the basic version of SMA still has some problems. When faced with some complex problems, it may fall into local optimum and cannot find the optimal solution. Aiming at this problem, an improved SMA is proposed to alleviate the disadvantages of SMA. Based on the original SMA, Gaussian mutation and Levy flight are introduced to improve the global search performance of the SMA. Adding Gaussian mutation to SMA can improve the diversity of the population, and Levy flight can alleviate the local optimum of SMA, so that the algorithm can find the optimal solution as soon as possible. In order to verify the effectiveness of the proposed algorithm, a continuous version of the proposed algorithm, GLSMA, is tested on 33 classical continuous optimization problems. Then, on 14 high-dimensional gene datasets, the effectiveness of the proposed discrete version, namely, BGLSMA, is verified by comparing with other feature selection algorithm. Experimental results reveal that the performance of the continuous version of the algorithm is better than the original algorithm, and the defects of the original algorithm are alleviated. Besides, the discrete version of the algorithm has a higher classification accuracy when fewer features are selected. This proves that the improved algorithm has practical value in high-dimensional gene feature selection.

## 1. Introduction

With the development of modern social science and technology, a variety of problems have arisen in the society, requiring researchers to design more efficient and novel methods to put forward feasible solutions. In recent years, some metaheuristic algorithms have been developed to solve various optimization problems. Some studies also show that metaheuristic methods are more effective than traditional gradient-based methods [1]. Metaheuristic algorithms can be divided into several categories according to their causes: evolutionary algorithm (EAs), such as genetic algorithm (GA) [2] and differential evolution (DE) [3], and swarm intelligence algorithms (SI), such as particle swarm optimization (PSO) [4], Harris hawks algorithm (HHO) [5], RUNge Kutta optimizer (RUN) [6], hunger games search (HGS) [7], slime mould algorithm (SMA) [8], monarch butterfly optimization (MBO) [9], moth search algorithm (MSA) [10], colony predation algorithm (CPA) [11], and weighted mean of vectors (INFO) [12]. In addition, they have been widely used in various fields, such as solar cell parameter Identification [13], economic emission dispatch problem [14], image segmentation [15, 16], plant disease recognition [17], medical diagnosis [18, 19], scheduling problems [20–22], optimization of machine learning model [23], multiobjective problem [24, 25], fault diagnosis [26], object tracking [27, 28], expensive optimization problems [29, 30], medical diagnosis [31, 32], combination optimization problems [33], feature selection [34, 35], practical engineering problems [36, 37], and robust optimization [38, 39].

Among all the algorithms, SMA is a new one proposed in recent years. Because of its excellent performance in

dealing with complex problems and simple implementation, SMA has been widely applied in recent years. Because of its exploration and exploitation capabilities, it has been widely used in various fields to solve specific practical problems. For example, Kouadri et al. [40] proposed to use SMA in the actual power system to solve the optimal power problem and minimize the total cost of conventional and random power generation under the constraints of the power system. Khunkitti et al. [41] proposed the multiobjective optimal power flow (MOOPF) problem based on SMA, taking cost emission and transmission line loss as part of the objective function of the power system. Simulation results show that SMA has better solutions than other algorithms in the literature. Jafari-Asl et al. [42] proposed a method combining LS (line sampling) method with slime mould algorithm to solve the reliability problem under highly nonlinear and implicit limit states. Izci and Ekinci [43] evaluated the optimization ability of SMA by using a proportional integral derivative (PID) controller to adjust the speed of dc motor and maintain the terminal output of automatic voltage regulator (AVR) system and compared the performance of SMA with that of other controllers designed by competitive algorithms. Houssein et al. [44] proposed a multiobjective optimization algorithm based on SMA. The reliability of the proposed MOSMA was verified by the actual multiobjective optimization of automotive helical springs, and the effectiveness of MOSMA was evaluated by the Wilcoxon test and performance indicators. Houssein et al. [45] proposed a method combining SMA with adaptive guided differential evolution algorithm (AGDE) (SM-AGDE) to solve some of the defects of SMA. Gupta et al. [46] proposed a SMA to solve the estimation problem of proton exchange membrane fuel cell (PEMFC) model, which showed good performance in jumping out of local optimum, and the predicted results were basically consistent with the actual results. Therefore, SMA can be used for fuel cell problems. Elsayed et al. [47] used SMA to identify the parameters of transformer equivalent circuit and verified the ability and accuracy of SMA in parameter estimation of single-three-phase transformer, as well as its high performance and stability in determining the optimal parameters.

Hassan et al. [48] proposed an improved SMA (ISMA) to solve the problem of target and dual target economy and emission scheduling (EED) considering the valve point effect, in which the best solution was obtained by updating the position of the solution by using two equations in the sine and cosine algorithm. At the same time, on the basis of Pareto dominance concept and fuzzy decision, multiobjective ISMA is proposed, which has good performance and robustness. Jia et al. [49] optimized the SMA by introducing compound mutation strategy (CMS) and restart strategy (RS). CMS was used to increase population diversity, RS was used to avoid local optimization, and the effectiveness of the proposed CMSRSSMA was tested on the benchmark function. Meanwhile, the CMSRSSMA_SVM model was proposed and used for feature selection and parameter optimization. Altay [50] utilized 10 different chaotic mappings to generate chaotic rather than random values in SMA. By using chaotic mapping, the global convergence rate of SMA is improved and the local solution is avoided. Abdel-Basset et al. [51] integrated SMA and WOA algorithms to maximize Kapur's entropy and applied them in the field of image segmentation, achieving good results. Chauhan et al. [52] proposed a method combining arithmetic optimizer algorithm (AOA) and slime mould algorithm (SMA), namely, HAOASMA algorithm, which solved the problems of slime mould algorithm's insufficient memory and slow local convergence speed.

Since SMA was proposed, it has been applied in various fields and used to solve various problems, showing good performance. However, in the face of some complex optimization tasks, there are still problems of falling into local optimum and slow convergence. In order to cope with this situation and improve the performance of the algorithm, a combinatorial optimization method (GLSMA) based on Gaussian mutation and Levy flight is proposed in this paper. In GLSMA, the global exploration and local exploitation capabilities of the original algorithm are improved by introducing Gaussian mutation and Levy flight mechanism. In the optimization iteration process, the original position of slime moulds in the population was modified by Gaussian to enhance the diversity of the population and improve the global exploration ability of the algorithm, so that the algorithm could achieve a balance between global exploration and local exploitation. After that, Levy flight was used to improve the randomness of SMA and to jump out of local optimum. Benchmark function test results show that the improved version of GLSMA has better global search and local exploration ability compared with other advanced algorithms. The discrete version based on GLSMA also shows an ideal effect on feature selection.

The contributions and highlights of this paper are summarized below:

(1) An improved slime mould algorithm (GLSMA) based on Gaussian mutation and Levy flight is proposed to solve continuous optimization problems and high-dimensional gene feature selection problems

(2) The superiority of GLSMA is proved by comparing with several well-known algorithms on public datasets and achieved good results

(3) Proposed binary GLSMA to solve high-dimensional gene feature selection problems

(4) The developed GLSMA has faster exploration speed and convergence speed in the global optimization task

(5) Binary GLSMA has the highest classification accuracy and the least number of features in high-dimensional gene feature selection task

The remainder of this paper is organized as follows: the second part introduces the original SMA. The third part introduces Gaussian mutation mechanism and Levy flight in detail, as well as the improved SMA based on the two mechanisms. The fourth part introduces a series of

comparison experiments between GLSMA and other similar algorithms, including comparison experiments on continuous function and discrete function. The fifth part reviews and discusses of the proposed work. The sixth part summarizes the conclusions of this paper and gives several directions of future work.

## 2. Slime Mould Algorithm

Li et al. [8] established a mathematical model based on the oscillation behavior of slime moulds and thus proposed a metaheuristic slime mould algorithm (SMA).

The mathematical formula of slime moulds is shown in

$$X(t+1) = \begin{cases} \text{rand}\bullet(\text{UB} - \text{LB}) + \text{LB}, & \text{rand} < z, \\ X_b(t) + \text{vb}(t)\bullet(W\bullet X_A(t) - X_B(t)), & r < p, \\ \text{vc}(t)\bullet X(t), & r \geq p, \end{cases} \tag{1}$$

where vb represents a random value in the interval $[-a, a]$. Parameter vc ranges in the interval $[-b, b]$, which decreases with the number of iterations. Through the cooperative interaction between vb and vc, the selection behavior of slime moulds can be simulated and the optimal solution can be selected. $\text{Max}_t$ indicates the maximum number of iterations. UB and LB represent the upper and lower boundaries of the search space, respectively. $X_b$ represents the position vector of the current highest fitness (highest concentration) individual. $X_A(t)$ and $X_B(t)$ represent the position vectors of random individuals selected from the slime moulds in the $t$ iteration. rand and $r$ are the random values between 0 and 1. Parameter $z$ is set to 0.03. $X(t)$ and $X(t+1)$ represent position vectors of slime moulds at the $t$ and $(t+1)$ iterations, respectively.

$$a = \text{arctanh}\left(1 - \left(\frac{t}{\text{Max}_t}\right)\right), \tag{2}$$

$$b = 1 - \left(\frac{t}{\text{Max}_t}\right). \tag{3}$$

In addition, the decision parameter $p$ can be calculated as follows:

$$p = \tanh |S(i) - \text{DF}|, \tag{4}$$

where $S(i)$ indicates the fitness of the ith individual in the slime mould, $i \in 1, 2, \cdots, N$; $N$ denotes the size of population; and DF represents the best fitness, which is attained during all of the iterations.

$$W(\text{SmellIndex}(i)) = \begin{cases} 1 + r\bullet\log\left(\frac{\text{bF} - \text{SmellOrder}(i)}{\text{bF} - \text{wF}} + 1\right), & \text{condition}, \\ 1 - r\bullet\log\left(\frac{\text{bF} - \text{SmellOrder}(i)}{\text{bF} - \text{wF}} + 1\right), & \text{otherwise}, \end{cases} \tag{5}$$

$$[\text{SmellOrder}, \text{SmellIndex}] = \text{sort}(S), \tag{6}$$

where $W$ is the weight vector of slime moulds and $bF$ and $wF$ are the best and worst fitness obtained in the current iteration, respectively. SmellIndex and SmellOrder represent fitness ordering order (minimum problems are sorted in ascending order) and corresponding fitness values, respectively. condition represents the first half of SmellOrder.

## 3. Description of the GLSMA

*3.1. Gaussian Mutation.* Gaussian mutation (GM) operator is derived from Gaussian normal distribution, which is distinguished from Cauchy distribution. In the vertical direction, the Gaussian distribution is larger than the Cauchy distribution, and in the horizontal direction, the Gaussian distribution is smaller than the Cauchy distribution. Gauss mutations are more likely to produce new offspring in this part because of their narrow tail. In response, the search equation takes smaller steps to explore every corner of the search space in a better way. The Gaussian density function can be described as

$$f_{\text{Gaussian}(0,\sigma^2)}(\propto) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\sigma^2/2\sigma^2}, \tag{7}$$

where $\sigma^2$ is the variance of each member of the population. By setting the mean to 0 and the standard deviation to 1, this function is further simplified to generate an $N$-dimensional random variable. The generated random variables were applied to the exploration stage of slime moulds, as shown below:

$$X_i' = X_i \times (1 + G(\propto)), \tag{8}$$

where $G(\propto)$ is a uniformly distributed random number derived from Gaussian distribution, $X_i$ is a position in SMA during the current iteration, and $X_i'$ is the position corresponding to $X_i$ after Gaussian mutation. The introduction of Gaussian mutation mechanism enhances the diversity of population and improves the quality of SMA solution.

*3.2. Levy Flight.* Levy flight (LF) was first proposed by French mathematician Paul Levy in 1937, after which researchers used Levy statistics to describe various natural phenomena. Levy flight operator improves slime mould search capability by helping all search agents advance to better, more promising positions. A simple description of the Levy distribution is as follows:

$$\text{Levy}(\beta) \sim u = t^{-1-\beta}, 0 < \beta \leq 2, \tag{9}$$

where $\beta$ is an important index of regulatory stability. Levy random numbers can be described by the following formula:

$$\text{Levy}(\beta) \sim \frac{\varphi \times \mu}{|\nu|^{1/\beta}}, \tag{10}$$

where $\mu$ and $\nu$ are standard normal distributions, $\Gamma$ is a

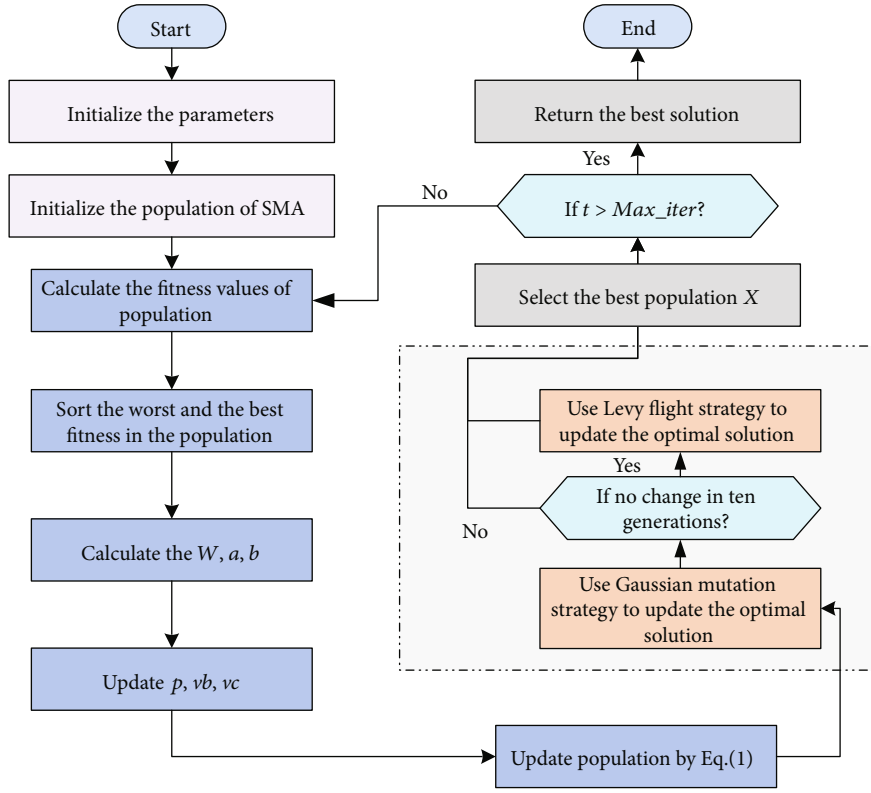



Figure 1: Flowchart of the proposed GLSMA.

standard gamma function, $\beta$ is set to 1.5, and $\varphi$ is defined as follows:

$$\varphi = \left[\frac{\Gamma(1+\beta)\times \sin(\pi\times\beta/2)}{\Gamma\left(((1+\beta)/2)\times\beta\times 2^{(\beta-1)/2}\right)}\right]^{1/\beta}. \quad (11)$$

In the exploration phase of slime mould algorithm, Levy strategy was used to update the location of search agents, so as to better balance exploration and search capabilities. The update formula is as follows:

$$X_i' = X_i \times (1 + \text{Levy}(\beta)). \quad (12)$$

In the formula, Levy($\beta$) is taken from the Levy distribution and is the number of random distribution. $X_i'$ is the new location of the $i$-th search agent $X_i$ after the update. The introduction of Levy flight can help all individuals to jump out of local optimum and improve the quality of the population.

*3.3. Framework of Proposed GLSMA.* In this section, we will describe GLSMA based on the Gaussian mutation mechanism and Levy flight strategy in detail. In the process of algorithm improvement, adding a mechanism can generally improve the algorithm in only one aspect but cannot improve the global exploration and local exploitation ability at the same time. By adding the Gaussian mutation mechanism, the corresponding value can be obtained from the current solution, but this can only improve the local exploitation ability and will fall into local optimal. The Levy flight mechanism can expand the search range of solutions, increase the possibility of obtaining the optimal solution, and avoid falling into local optimal. As a result, in the original SMA, two strategies (GM and LF) were introduced to facilitate the coordination of global exploration and local exploitation, forming a new SMA variant.

In the process of iterative optimization, Gaussian mutation was considered for individuals in the slime mould individuals after initial updating. The individuals obtained after mutation were compared with the individuals without that. If the fitness of the individuals in the mutation state was not improved, the original individuals were retained and the mutant individuals were discarded to ensure the quality of the population. Considering that the algorithm is easy to fall into local optimum, levy flight strategy is introduced to improve the randomness of SMA and the ability of jumping out of local optimum. The flowchart of GLSMA is shown in Figure 1. Experimental results show that compared with other swarm intelligence algorithms, GLSMA not only has stronger global exploration ability but also contributes to increase the quality of solutions and speed up convergence. The structure of the proposed GLSMA optimizer is shown in Algorithm 1.

*3.4. Computational Complexity Analysis.* According to the structure of GLSMA, it mainly includes initialization, fitness evaluation, fitness ranking, weight updating, position updating based on SMA strategy, position adjustment based on Gaussian mutation mechanism, and position updating based

**Begin**
  Initialize of the parameters: Max_*iter*, $N$
  Initialize of slime mould population $X$
  **While***t* ≤ Max_*iter*
    Calculate the fitness for each individual in slime mould
    Update $X_b$ and the best fitness
    Calculate the weight $W$, a, and b according to Equations (2), (3), (5)
    **For***i* = 1 : $N$
      Update $p$ using Equation (4)
      Update *vb* and *vc* based on *a* and *b*, respectively
      Update the positions by Equation (1)
    **End for**
    Utilize the best individual in the population to perform GM operations
    **If** (meet the condition)
    **Then** use Levy flight to avoid falling into local optimality
    Iteration = iteration + 1
  **End while**
  Return the best fitness and $X_b$ as the best solution
**End**

Algorithm 1: Pseudocode of GLSMA.

Table 1: Descriptions of unimodal benchmark functions.

| Function | Dim | Range | $f_{\min}$ |
|---|---|---|---|
| $F_1(x) = \sum_{i=1}^{n} x_i^2$ | 30 | [-100, 100] | 0 |
| $F_2(x) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} \lvert x_i \rvert$ | 30 | [-10, 10] | 0 |
| $F_3(x) = \sum_{i=1}^{n} \left( \sum_{j-1}^{i} x_j \right)^2$ | 30 | [-100, 100] | 0 |
| $F_4(x) = \max_i \{ \lvert x_i \rvert, 1 \le i \le n \}$ | 30 | [-100, 100] | 0 |
| $F_5(x) = \sum_{i=1}^{n-1} \left[ 100 \left( x_{i+1} - x_i^2 \right)^2 + (x_i - 1)^2 \right]$ | 30 | [-30, 30] | 0 |
| $F_6(x) = \sum_{i=1}^{n} (\lvert x_i + 0.5 \rvert)^2$ | 30 | [-100, 100] | 0 |
| $F_7(x) = \sum_{i=1}^{n} i x_i^4 + \text{random}[0, 1]$ | 30 | [-128, 128] | 0 |

on Levy flight strategy, where $N$ is the number of slime moulds, $D$ is the function's dimension, and $T$ is the maximum number of iterations. The calculation is as follows:

The time complexity for initialization is $O(D)$. In evaluating and ordering fitness, the computational complexity is $O(N + N \log N)$. The computational complexity of the update weight is $O(N \times D)$. The computational complexity of position updating process based on SMA is $O(N \times D)$. Similarly, the computational complexity of position updating process based on Gaussian mutation mechanism is $O(N \times D)$. The computational complexity of the position update process based on Levy flight is $O(N \log D)$. Therefore, the total computational complexity of GLSMA is $O(D + T \times N \times (1 + 3D + \log D + \log N))$.

## 4. Experiments and Results

In the experiment, to evaluate the continuous and discrete versions of GLSMA, the proposed SMA algorithm is compared with other optimizers on the continuous functions and feature problems, respectively. The effectiveness and competitiveness of the proposed algorithm are verified by two parts of experiments. In the first part, the strategies added on SMA were tested on 23 benchmark test functions (including 7 unimodal functions, 6 multimodal functions, and more than 10 fixed dimension multimodal functions) and 10 classic CEC2014 benchmark test functions (including 2 hybrid functions and 8 composition functions), to see whether the mechanism has a positive effect on the algorithm. Then, in the same test environment, GLSMA is compared with some original algorithms and advanced MA algorithms. In the second part, we compare the proposed binary GLSMA (BGLSMA) with other classifiers on feature selection problems.

All GLSMA experiments were written in the MATLAB R2014a compiler and run on Windows 10(64-bit) operating system. The computer hardware is Intel(R) Xeon(R) Silver 4110 CPU (2.40 GHz) 2.10 GHz (dual processors) and 32GB RAM.

In Section 4.1, we will test the influence of different mechanisms on the algorithm. In Section 4.2, GLSMA is

Table 2: Descriptions of multimodal benchmark functions.

| Function | Dim | Range | $f_{\min}$ |
|---|---|---|---|
| $F_8(x) = \sum_{i=1}^{n} -x_i \sin\left(\sqrt{\lvert x_i \rvert}\right)$ | 30 | [-500, 500] | $-418.9829 \times 30$ |
| $F_9(x) = \sum_{i=1}^{n} \left[x_i^2 - 10\cos(2\pi x_i) + 10\right]$ | 30 | [-5.12, 5.12] | 0 |
| $F_{10}(x) = -20\exp\left(-0.2\sqrt{1/n\sum_{i=1}^{n} x_i^2}\right) - \exp\left(1/n\sum_{i=1}^{n}\cos(2\pi x_i)\right) + 20 + e$ | 30 | [-32, 32] | 0 |
| $F_{11}(x) = 1/4000\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos\left(x_i/\sqrt{i}\right) + 1$ | 30 | [-600, 600] | 0 |
| $F_{12}(x) = \frac{\pi}{n}\left\{10\sin(\pi y_1) + \sum_{i=1}^{n-1}(y_i - 1)^2\left[1 + 10\sin^2(\pi y_{i+1})\right] + (y_n - 1)^2\right\} + \sum_{i=1}^{n} u(x_i, 10, 100, 4)$ <br> $y_i = 1 + \frac{x_i + 1}{4}$ <br> $u(x_i, a, k, m) = \begin{cases} k(x_i - a)^m & x_i > a \\ 0 & -a < x_i < a \\ k(-x_i - a)^m & x_i < -a \end{cases}$ | 30 | [-50, 50] | 0 |
| $F_{13}(x) = 0.1\left\{\sin^2(3\pi x_1) + \sum_{i=1}^{n}(x_i - 1)^2\left[1 + \sin^2(3\pi x_i + 1)\right] + (x_n - 1)^2\left[1 + \sin^2(2\pi x_n)\right]\right\} + \sum_{i=1}^{n} u(x_i, 5, 100, 4)$ | 30 | [-50, 50] | 0 |

Table 3: Descriptions of fixed dimension multimodal benchmark functions.

| Function | Dim | Range | $f_{\min}$ |
|---|---|---|---|
| $F_{14}(x) = \left(1/500 + \sum_{j=1}^{25}\left(1/j + \sum_{i=1}^{2}\left(x_i - a_{ij}\right)^6\right)\right)^{-1}$ | 2 | [-65, 65] | 1 |
| $F_{15}(x) = \sum_{i=1}^{11}\left[a_i - x_1\left(b_i^2 - b_i x_2\right)/b_i^2 + b_i x_3 + x_4\right]^2$ | 4 | [-5, 5] | 0.00030 |
| $F_{16}(x) = 4x_1^2 - 2.1x_1^2 + 1/3x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$ | 2 | [-5, 5] | -1.0316 |
| $F_{17}(x) = \left(x_2 - \left(5.1/4\pi^2\right)x_1^2 + (5/\pi)x_1 - 6\right)^2 + 10(1 - 1/8\pi)\cos x_1 + 10$ | 2 | [-5, 5] | 0.398 |
| $F_{18}(x) = \left[1 + (x_1 + x_2 + 1)^2\left(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2\right)\right] \times \left[30 + (2x_1 - 3x_2)^2 \times \left(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2\right)\right]$ | 2 | [-2, 2] | 3 |
| $F_{19}(x) = -\sum_{i=1}^{4} c_i \exp\left(-\sum_{j=1}^{3} a_{ij}\left(x_j - p_{ij}\right]\right)^2\right)$ | 3 | [1, 3] | -3.86 |
| $F_{20}(x) = -\sum_{i=1}^{4} c_i \exp\left(-\sum_{j=1}^{6} a_{ij}\left(x_j - p_{ij}\right]\right)^2\right)$ | 6 | [0, 1] | -3.32 |
| $F_{21}(x) = -\sum_{i=1}^{5}\left[(X - a_i)(X - a_i)^T + c_i\right]^{-1}$ | 4 | [0, 10] | -10.1532 |
| $F_{22}(x) = -\sum_{i=1}^{7}\left[(X - a_i)(X - a_i)^T + c_i\right]^{-1}$ | 4 | [0, 10] | -10.4028 |
| $F_{23}(x) = -\sum_{i=1}^{10}\left[(X - a_i)(X - a_i)^T + c_i\right]^{-1}$ | 4 | [0, 10] | -10.5363 |

compared with seven metaheuristic algorithms to prove its effectiveness. In Section 4.3, GLSMA is compared with eight advanced algorithms to verify its ability on exploration and exploitation. In Section 4.4, we use binary GLSMA (BGLSMA) to deal with feature selection in 14 UCI datasets.

*4.1. The Influence of Gaussian Mutation and Levy Flight.* As mentioned above, GLSMA consists of two main improved strategies: Gaussian mutation mechanism and Levy flight strategy. The purpose of this section is to validate the effectiveness of the combination of the two strategies. To this end, we compare GLSMA, SMA, and their variants GSMA and LSMA on 33 benchmark functions. GSMA only uses Gaussian mutation strategy, LSMA only uses Levy flight strategy, and SMA is the original algorithm.

All algorithm tests were carried out under the expected conditions to eliminate the influence of irrelevant factors on the experiment and ensure the fairness of the test. The population size was set to 30; the maximum evaluation test was uniformly set to 300,000. In order to weaken the influence of algorithm randomness on the experiment, we conducted 30 independence tests for each test case. In this paper, the average value of optimal function (Avg) and standard deviation (Std) of the selection algorithm results are compared. The global exploration ability and result quality of the algorithm were evaluated on the average (Avg), and Std of the optimal function was used to evaluate the robustness of the algorithm. In order to show the best results more clearly, all the best results are italicized.

In addition, nonparametric statistical verification Wilcoxon signed-rank test was used to measure the degree of improvement and whether it was statistically significant. The significance level was set at 0.05. The symbolic label "+/=/-" in the results states that the proposed method GLSMA is superior to, equal to, and inferior to other methods of competition, respectively. For a comprehensive statistical comparison, the Friedman [53] test was used to evaluate the average behavior of all different algorithms for further statistical comparison, and the average ranking was given in these comparison results, and the average rank value (ARV) of the Friedman test was used to evaluate the average performance of the compared methods.

Tables 1–4 contain 23 benchmark functions and 10 test functions in CEC2014. The selected 33 test functions include several different problems, covering unimodal function, multimodal function, fixed dimension multimodal function, hybrid function, and composition function. These test functions can be used to test the algorithm's global exploration capabilities and local exploitation capabilities and can be used to verify the balance between exploration and exploitation capabilities.

As can be seen from the results in Tables 5 and 6, GLSMA is significantly superior to other mechanism combinations and the original SMA. After careful analysis, Avg and Std in Table 5 represent the superiority of GLSMA over F1-F7, F9-F14, F17-F18, and F22-F33 functions. On the test functions F1-F4, F9-F11, F26-F28, and F30-F33, the GLSMA's Std value is 0, indicating that GLSMA has strong robustness. This is because the combination of Gaussian mutation and Levy flight mechanism improves the performance of the original SMA and can successfully find global optimal solutions for various complex problems. According to the statistical results of $p$ value in Table 6, many values of SMA column are less than 0.05, indicating that GLSMA has a certain improvement on the original SMA. As can be seen from the ARV tested by Friedman in Table 7, when comparing the four algorithms, GLSMA ranks first and is significantly superior to other algorithms. Moreover, it can be seen that the improvement effect of Gaussian mutation mechanism or Levy flight mechanism on the original SMA is not good, even can be said to be poor, but the combination of the two can achieve a good balance between exploration and exploitation, so as to achieve a good effect. In summary, the results show that the addition of Gaussian mutation mechanism and Levy flight mechanism is not only beneficial to the exploration and exploitation ability of GLSMA but also beneficial to the balance between the exploitation and exploration ability of GLSMA, which has a certain positive effect on the algorithm and improves the robustness of the original algorithm, which has improved significance.

Table 4: Descriptions of CEC2014 functions (search range: [−100, 100]D).

| Function | Class | Functions | Optimum |
|---|---|---|---|
| F24 | Hybrid | Hybrid 5: $N = 5$ | 2100 |
| F25 | | Hybrid 6: $N = 5$ | 2200 |
| F26 | Composition | Composition 1: $N = 5$ | 2300 |
| F27 | | Composition 2: $N = 3$ | 2400 |
| F28 | | Composition 3: $N = 3$ | 2500 |
| F29 | | Composition 4: $N = 5$ | 2600 |
| F30 | | Composition 5: $N = 5$ | 2700 |
| F31 | | Composition 6: $N = 5$ | 2800 |
| F32 | | Composition 7: $N = 3$ | 2900 |
| F33 | | Composition 8: $N = 3$ | 3000 |

Compared with the table, the image can more intuitively and clearly reflect the optimization results of GLSMA compared with other comparison objects. Figure 2 shows the convergence curves of the four comparison methods on nine functions. It is obvious that GLSMA using two mechanisms achieves better results than its variants. The combination of Gaussian mutation and Levy flight enables GLSMA to escape from local traps faster and obtain the global optimal solution. In the meantime, it can be seen that GLSMA has the fastest rate of convergence and can get the optimal value first. The results show that this combination of mechanisms can quicken the convergence of the algorithm while jumping out of local optimum. In general, the combination of GM and LF improves the overall performance of the original SMA.

*4.2. Comparison with Well-Known Algorithms.* In this experiment, 23 classical functions and 10 of the CEC2014 benchmark functions were selected to evaluate the performance of

Table 5: Experimental results of GLSMA, SMA, GSMA, and LSMA on 33 benchmark functions.

| | F1 | | F2 | | F3 | |
|---|---|---|---|---|---|---|
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *0.00E+00* | *0.00E+00* | *0.00E+00* | *0.00E+00* | *0.00E+00* | *0.00E+00* |
| SMA | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 |
| GSMA | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 |
| LSMA | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 |
| | F4 | | F5 | | F6 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *0.00E+00* | *0.00E+00* | *1.29E-03* | *1.16E-03* | *7.00E-06* | *4.61E-06* |
| SMA | 0.00E+00 | 0.00E+00 | 1.97E-03 | 1.01E-03 | 9.52E-06 | 4.12E-06 |
| GSMA | 0.00E+00 | 0.00E+00 | 1.96E-03 | 1.40E-03 | 1.32E-05 | 5.03E-06 |
| LSMA | 0.00E+00 | 0.00E+00 | 4.39E-03 | 2.62E-03 | 1.23E-05 | 5.56E-06 |
| | F7 | | F8 | | F9 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *5.98E-06* | *5.13E-06* | -1.26E+04 | 1.07E-02 | *0.00E+00* | *0.00E+00* |
| SMA | 9.56E-06 | 8.00E-06 | -1.26E+04 | 2.62E-04 | 0.00E+00 | 0.00E+00 |
| GSMA | 9.94E-06 | 9.90E-06 | *-1.26E+04* | *2.32E-04* | 0.00E+00 | 0.00E+00 |
| LSMA | 1.62E-05 | 1.41E-05 | -1.27E+04 | 4.23E-04 | 0.00E+00 | 0.00E+00 |
| | F10 | | F11 | | F12 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *8.88E-16* | *0.00E+00* | *0.00E+00* | *0.00E+00* | *2.35E-06* | *4.36E-06* |
| SMA | 8.88E-16 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 1.05E-05 | 1.00E-05 |
| GSMA | 8.88E-16 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 8.42E-06 | 1.16E-05 |
| LSMA | 8.88E-16 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 8.46E-06 | 8.61E-06 |
| | F13 | | F14 | | F15 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *3.40E-06* | *2.79E-06* | *9.98E-01* | *4.51E-16* | 3.12E-04 | 2.13E-05 |
| SMA | 6.59E-06 | 3.86E-06 | 9.98E-01 | 4.77E-16 | *3.08E-04* | *1.89E-06* |
| GSMA | 8.30E-06 | 5.62E-06 | 9.98E-01 | 5.49E-16 | 3.15E-04 | 3.67E-05 |
| LSMA | 8.12E-06 | 2.21E-06 | 9.98E-01 | 5.79E-16 | 3.21E-04 | 5.70E-05 |
| | F16 | | F17 | | F18 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | -1.03E+00 | 6.32E-14 | *3.98E-01* | *7.20E-12* | *3.00E+00* | *1.04E-14* |
| SMA | -1.03E+00 | 1.21E-14 | 3.98E-01 | 1.32E-11 | 3.00E+00 | 1.16E-14 |
| GSMA | *-1.03E+00* | *9.28E-15* | 3.98E-01 | 1.14E-11 | 3.00E+00 | 1.21E-14 |
| LSMA | -1.03E+00 | 1.31E-14 | 3.98E-01 | 2.58E-11 | 3.00E+00 | 1.03E-14 |
| | F19 | | F20 | | F21 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | -3.86E+00 | 6.18E-11 | -3.24E+00 | 5.70E-02 | -1.02E+01 | 1.27E-06 |
| SMA | *-3.86E+00* | *4.13E-11* | -3.23E+00 | 4.84E-02 | -1.02E+01 | 1.03E-06 |
| GSMA | -3.86E+00 | 1.24E-10 | -3.23E+00 | 4.84E-02 | *-1.02E+01* | *9.16E-07* |
| LSMA | -3.86E+00 | 6.17E-11 | *-3.21E+00* | *3.02E-02* | -1.02E+01 | 1.27E-06 |
| | F22 | | F23 | | F24 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *-1.04E+01* | *9.50E-07* | *-1.05E+01* | *8.73E-07* | *1.05E+05* | *8.26E+04* |
| SMA | -1.04E+01 | 1.15E-06 | -1.05E+01 | 1.06E-06 | 1.50E+05 | 4.96E+04 |
| GSMA | -1.04E+01 | 1.56E-06 | -1.05E+01 | 1.49E-06 | 1.16E+05 | 5.52E+04 |
| LSMA | -1.04E+01 | 1.18E-06 | -1.05E+01 | 1.44E-06 | 1.38E+05 | 6.74E+04 |

Table 5: Continued.

| | F25 | | F26 | | F27 | |
|---|---|---|---|---|---|---|
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *2.53E+03* | *1.78E+02* | *2.50E+03* | *0.00E+00* | *2.60E+03* | *0.00E+00* |
| SMA | 2.70E+03 | 2.13E+02 | 2.50E+03 | 0.00E+00 | 2.60E+03 | 0.00E+00 |
| GSMA | 2.71E+03 | 1.75E+02 | 2.50E+03 | 0.00E+00 | 2.60E+03 | 0.00E+00 |
| LSMA | 2.65E+03 | 2.04E+02 | 2.50E+03 | 0.00E+00 | 2.60E+03 | 0.00E+00 |
| | F28 | | F29 | | F30 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *2.70E+03* | *0.00E+00* | *2.70E+03* | *9.59E-02* | *2.90E+03* | *0.00E+00* |
| SMA | 2.70E+03 | 0.00E+00 | 2.70E+03 | 1.42E-01 | 2.90E+03 | 0.00E+00 |
| GSMA | 2.70E+03 | 0.00E+00 | 2.70E+03 | 1.46E-01 | 2.90E+03 | 0.00E+00 |
| LSMA | 2.70E+03 | 0.00E+00 | 2.70E+03 | 1.21E-01 | 2.90E+03 | 0.00E+00 |
| | F31 | | F32 | | F33 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *3.00E+03* | *0.00E+00* | *3.10E+03* | *0.00E+00* | *3.20E+03* | *0.00E+00* |
| SMA | 3.00E+03 | 0.00E+00 | 4.11E+03 | 1.10E+03 | 5.38E+03 | 1.64E+03 |
| GSMA | 3.00E+03 | 0.00E+00 | 4.20E+03 | 1.12E+03 | 5.14E+03 | 1.67E+03 |
| LSMA | 3.00E+03 | 0.00E+00 | 3.89E+03 | 1.10E+03 | 5.67E+03 | 2.19E+03 |

Table 6: Wilcoxon signed-rank test results between GLSMA, SMA, GSMA, and LSMA.

| Function | SMA | GMSA | LSMA | Function | SMA | GMSA | LSMA |
|---|---|---|---|---|---|---|---|
| F1 | 1.00E+00 | 1.00E+00 | 1.00E+00 | F18 | 2.40E-01 | 5.15E-01 | 5.92E-01 |
| F2 | 1.00E+00 | 1.0E+00 | 1.00E+00 | F19 | 4.72E-02 | 8.61E-01 | 5.04E-01 |
| F3 | 1.00E+00 | 1.00E+00 | 1.00E+00 | F20 | 3.71E-01 | 3.19E-01 | 1.38E-03 |
| F4 | 1.00E+00 | 1.00E+00 | 1.00E+00 | F21 | 4.65E-01 | 7.81E-01 | 2.13E-01 |
| F5 | 2.18E-02 | 1.75E-02 | 3.11E-05 | F22 | 8.94E-01 | 8.98E-02 | 1.41E-01 |
| F6 | 4.95E-02 | 4.53E-04 | 1.06E-04 | F23 | 7.04E-01 | 2.18E-02 | 3.82E-01 |
| F7 | 3.68E-02 | 1.41E-01 | 5.71E-04 | F24 | 9.27E-03 | 6.58E-01 | 7.87E-02 |
| F8 | 4.53E-04 | 1.38E-03 | 1.48E-03 | F25 | 1.40E-02 | 4.90E-04 | 1.96E-02 |
| F9 | 1.00E+00 | 1.00E+00 | 1.00E+00 | F26 | 1.00E+00 | 1.00E+00 | 1.00E+00 |
| F10 | 1.00E+00 | 1.00E+00 | 1.00E+00 | F27 | 1.00E+00 | 1.00E+00 | 1.00E+00 |
| F11 | 1.00E+00 | 1.00E+00 | 1.00E+00 | F28 | 1.00E+00 | 1.00E+00 | 1.00E+00 |
| F12 | 7.71E-04 | 8.73E-03 | 2.11E-03 | F29 | 8.22E-02 | 1.78E-01 | 2.90E-01 |
| F13 | 4.39E-03 | 7.71E-04 | 6.98E-06 | F30 | 1.00E+00 | 1.00E+00 | 1000E+00 |
| F14 | 5.08E-01 | 7.27E-01 | 5.081E-01 | F31 | 1.00E+00 | 1.00E+00 | 1.00E+00 |
| F15 | 1.11E-01 | 9.59E-01 | 7.66E-01 | F32 | 6.10E-05 | 4.38E-04 | 6.10E-05 |
| F16 | 3.37E-01 | 3.16E-01 | 1.32E-01 | F33 | 5.96E-05 | 8.86E-05 | 1.32E-04 |
| F17 | 3.49E-01 | 2.85E-02 | 2.99E-01 | +/=/- | 9/22/2 | 9/23/1 | 9/231 |

Table 7: Average ranking values using the Friedman test.

| Algorithm | GLSMA | SMA | GSMA | LSMA |
|---|---|---|---|---|
| AVR | *1.454545* | 1.878788 | 2.151515 | 2.30303 |
| Rank | *1* | 2 | 3 | 4 |

GLSMA. The 33 benchmark test functions used in all experiments of continuous optimization can be divided into four categories: unimodal function, multimodal function, hybrid function, and composition function. The unimodal function (F1-F7) has only one solution, which can be used to test the development ability of the algorithm. The multimodal function (F8-F23) has several local optimal solutions and is suitable for verifying the exploration ability of the algorithm. The hybrid function and composition function (F24-F33) selected from CEC2014 are used to verify the balance between algorithm exploration and exploitation. These functions are often used to assess the overall power of algorithms. In this experiment, the performance of the improved

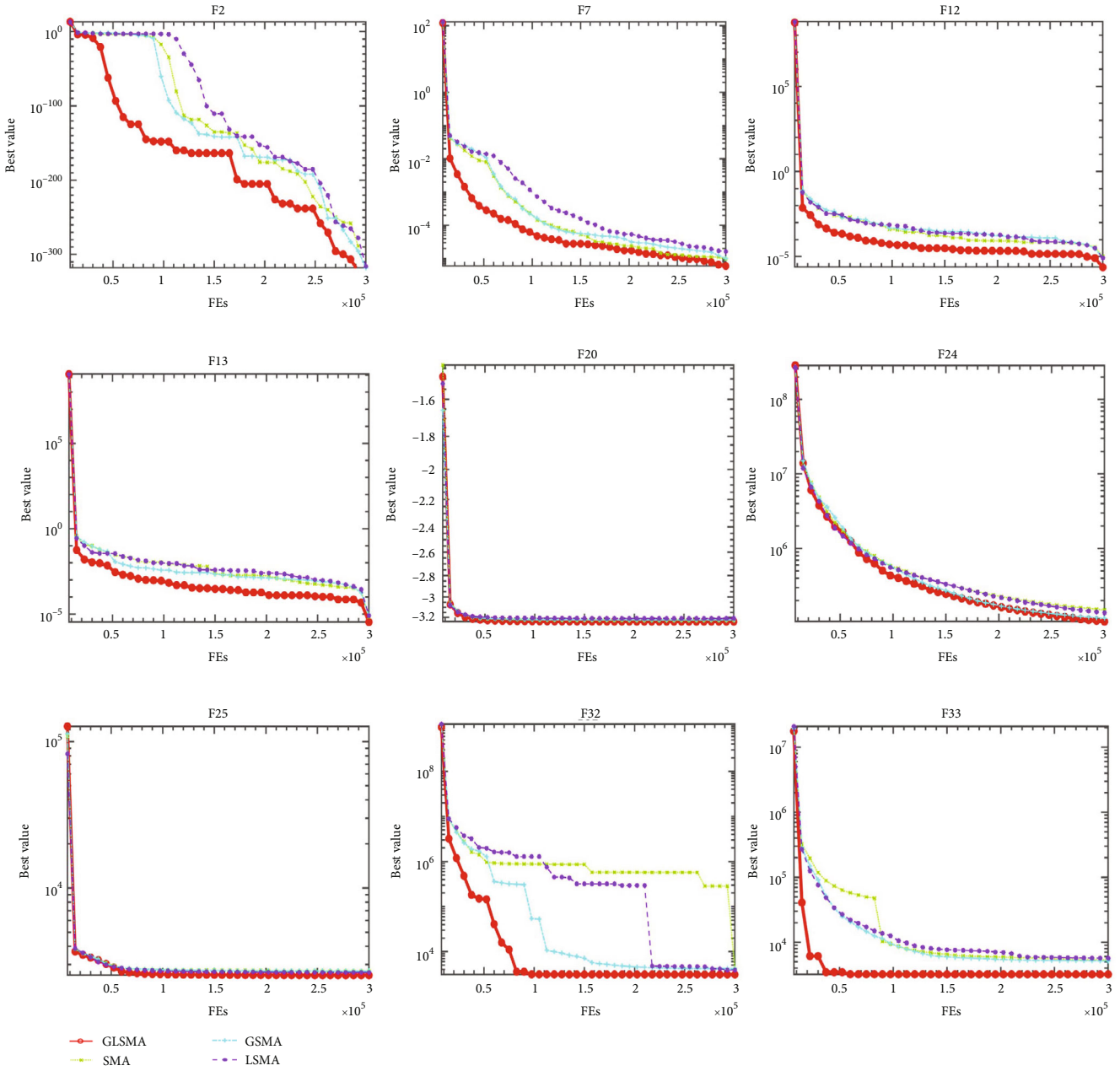



Figure 2: Convergence curves of GLSMA, SMA, GSMA, and LSMA on nine selected benchmark functions.

GLSMA was compared with PSO [4], WOA [54], GWO [55], SCA [56], FOA [57], DE [3], and SSA [58].

Tables 8–10 record the comparison results of GLSMA with seven well-known algorithms. The comparison results are shown in Table 10; among GLSMA and seven famous algorithms, the average Friedman test result of GLSMA is 2.328283, ranking first, and the average Friedman test result of DE is 2.730303, ranking second. It is obvious that the Friedman test of GLSMA and DE is obviously better than other algorithms. The average value (Avg) and standard deviation (Std) of optimal solution of GLSMA and other well-known algorithms are shown in Table 8. GLSMA has a significant advantage. Moreover, in all the comparison algorithms, GLSMA has Std 0 on more test functions, which proves that GLSMA algorithm has stronger stability. In addition, GLSMA shows obvious advantages and stability in almost all of the composition functions (F26-F28 and F30-F33). Table 9 shows the Wilcoxon symbol test results between GLSMA and other well-known algorithms. It can be seen that the $p$ value of GLSMA is less than 0.05 on almost all benchmark functions, which proves that GLSMA is significantly better than other algorithms, especially FOA, in all functions. Therefore, compared with these basic metaheuristic algorithms, GLSMA has statistical significance.

From the convergence curves of 8 algorithms on 9 functions shown in Figure 3, it can be seen that GLSMA

Table 8: Experimental results of GLSMA and seven original metaheuristic algorithms on 33 benchmark functions.

| | F1 | | F2 | | F3 | |
|---|---|---|---|---|---|---|
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *0.00E+00* | *0.00E+00* | *0.00E+00* | *0.00E+00* | *0.00E+00* | *0.00E+00* |
| PSO | 1.03E+02 | 1.01E+01 | 4.58E+01 | 3.92E+00 | 1.94E+02 | 2.47E+01 |
| WOA | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 4.18E+01 | 7.47E+01 |
| GWO | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 2.37E-183 | 0.00E+00 |
| SCA | 2.04E-56 | 6.25E-56 | 1.69E-58 | 5.76E-58 | 5.41E+00 | 1.51E+01 |
| FOA | 2.50E-09 | 9.63E-12 | 2.74E-04 | 4.17E-07 | 7.88E-07 | 1.85E-09 |
| DE | 3.70E-159 | 1.03E-158 | 2.85E-94 | 4.33E-94 | 1.27E+03 | 5.02E+02 |
| SSA | 3.79E-09 | 9.30E-10 | 2.96E-01 | 5.45E-01 | 5.90E-08 | 1.81E-08 |
| | F4 | | F5 | | F6 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *0.00E+00* | *0.00E+00* | 1.24E-03 | 1.37E-03 | 6.31E-06 | 3.72E-06 |
| PSO | 3.82E+00 | 1.26E-01 | *8.65E+04* | *2.0279E+04* | 1.02E+02 | 1.14E+01 |
| WOA | 4.37E+00 | 1.06E+01 | 2.44E+01 | 7.73E-01 | 5.80E-06 | 2.37E-06 |
| GWO | 1.81E-152 | 4.50E-152 | 2.65E+01 | 8.77E-01 | 4.89E-01 | 2.88E-01 |
| SCA | 1.52E-02 | 6.23E-02 | 2.75E+01 | 6.24E-01 | 3.53E+00 | 2.92E-01 |
| FOA | 9.14E-06 | 1.46E-08 | 2.87E+01 | 1.09E-04 | 7.50E+00 | 3.59E-07 |
| DE | 3.17E-15 | 5.95E-15 | 3.29E+01 | 2.01E+01 | *0.00E+00* | *0.00E+00* |
| SSA | 1.92E-01 | 3.42E-01 | 4.59E+01 | 3.57E+01 | 3.81E-09 | 7.66E-10 |
| | F7 | | F8 | | F9 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *7.45E-06* | *7.29E-06* | *-1.26E+04* | *1.27E-02* | *0.00E+00* | *0.00E+00* |
| PSO | 1.11E+02 | 2.41E+01 | -6.69E+03 | 8.24E+02 | 3.41E+02 | 1.52E+01 |
| WOA | 2.15E-04 | 2.53E-04 | -1.24E+04 | 5.62E+02 | 0.00E+00 | 0.00E+00 |
| GWO | 6.44E-05 | 4.41E-05 | -6.16E+03 | 6.56E+02 | 0.00E+00 | 0.00E+00 |
| SCA | 2.44E-03 | 2.59E-03 | -4.42E+03 | 2.83E+02 | 1.18E-12 | 6.47E-12 |
| FOA | 3.15E-05 | 1.44E-05 | -2.31E+02 | 1.13E+02 | 4.97E-07 | 1.73E-09 |
| DE | 2.39E-03 | 5.48E-04 | -1.24E+04 | 1.29E+02 | 1.33E-01 | 3.44E-01 |
| SSA | 7.64E-03 | 3.24E-03 | -7.76E+03 | 7.13E+02 | 6.91E+01 | 1.83E+01 |
| | F10 | | F11 | | F12 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *8.88E-16* | *0.00E+00* | *0.00E+00* | *0.00E+00* | *1.16E-06* | *1.62E-06* |
| PSO | 7.79E+00 | 3.02E-01 | 1.01E+00 | 9.86E-03 | 3.55E+00 | 3.93E-01 |
| WOA | 3.97E-15 | 2.59E-15 | 5.23E-04 | 2.86E-03 | *9.31E-07* | *4.33E-07* |
| GWO | 7.88E-15 | 6.49E-16 | 4.98E-04 | 1.90E-03 | 2.53E-02 | 1.77E-02 |
| SCA | 1.13E+01 | 9.36E+00 | 3.78E-09 | 2.07E-08 | 3.54E-01 | 1.30E-01 |
| FOA | 3.65E-05 | 4.17E-08 | 1.67E-10 | 6.21E-13 | 1.67E+00 | 6.05E-08 |
| DE | 7.53E-15 | 1.23E-15 | 0.00E+00 | 0.00E+00 | 1.57E-32 | 5.57E-48 |
| SSA | 1.78E+00 | 7.00E-01 | 1.28E-02 | 1.34E-02 | 9.54E-01 | 1.46E+00 |
| | F13 | | F14 | | F15 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | 4.57E-06 | 2.70E-06 | 9.98E-01 | 4.76E-16 | *3.08E-04* | *8.17E-07* |
| PSO | 1.59E+01 | 1.70E+00 | 2.78E+00 | 2.17E+00 | 9.76E-04 | 5.57E-05 |
| WOA | 7.74E-04 | 2.78E-03 | 9.98E-01 | 1.10E-14 | 4.70E-04 | 3.12E-04 |
| GWO | 4.08E-01 | 1.69E-01 | 3.15E+00 | 3.56E+00 | 7.53E-03 | 1.29E-02 |
| SCA | 2.03E+00 | 1.27E-01 | 9.98E-01 | 6.09E-07 | 5.15E-04 | 3.78E-04 |
| FOA | 5.79E-01 | 7.79E-02 | 1.27E+01 | 1.04E-15 | 8.36E-04 | 2.74E-04 |
| DE | *1.35E-32* | *5.57E-48* | *9.98E-01* | *0.00E+00* | 3.72E-04 | 1.83E-04 |
| SSA | 3.60E-03 | 6.49E-03 | 9.98E-01 | 1.13E-16 | 7.38E-04 | 3.71E-04 |

Table 8: Continued.

| | F16 | | F17 | | F18 | |
|---|---|---|---|---|---|---|
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | -1.03E+00 | 2.38E-14 | 3.98E-01 | 2.57E-12 | *3.00E+00* | *1.16E-14* |
| PSO | -1.03E+00 | 9.52E-05 | 3.98E-01 | 6.35E-05 | 3.01E+00 | 7.37E-03 |
| WOA | -1.03E+00 | 7.33E-15 | 3.98E-01 | 4.58E-10 | 3.00E+00 | 4.61E-08 |
| GWO | -1.03E+00 | 2.82E-11 | 3.98E-01 | 4.01E-09 | 3.00E+00 | 8.51E-08 |
| SCA | -1.03E+00 | 1.84E-06 | 3.98E-01 | 5.20E-05 | 3.00E+00 | 1.53E-07 |
| FOA | -1.55E-01 | 1.29E-01 | 1.34E+00 | 8.72E-01 | 6.00E+02 | 1.89E-05 |
| DE | -1.03E+00 | 6.78E-16 | *3.98E-01* | *0.00E+00* | 3.00E+00 | 2.03E-15 |
| **SSA** | *-1.03E+00* | *6.14E-16* | 3.98E-01 | 6.14E-16 | 3.00E+00 | 1.36E-14 |
| | F19 | | F20 | | F21 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | -3.86E+00 | 7.54E-11 | -3.23E+00 | 4.84E-02 | -1.02E+01 | 6.78E-07 |
| PSO | -3.86E+00 | 2.82E-03 | -2.92E+00 | 2.30E-01 | -7.10E+00 | 1.54E+00 |
| WOA | -3.86E+00 | 3.29E-03 | -3.25E+00 | 8.23E-02 | -1.02E+01 | 7.34E-07 |
| GWO | -3.86E+00 | 2.00 E-03 | -3.27E+00 | 7.29E-02 | -9.98E+00 | 9.31E-01 |
| SCA | -3.86E+00 | 3.107E-03 | -2.98E+00 | 2.31E-01 | -2.76E+00 | 2.52E+00 |
| FOA | -3.62E+00 | 2.47E-01 | -1.87E+00 | 4.40E-01 | -3.68E+00 | 7.81E-01 |
| DE | -3.86E+00 | 2.71E-15 | *-3.3E+00* | *2.83E-04* | -9.90E+00 | 1.36E+00 |
| SSA | *-3.86E+00* | *1.56E-15* | -3.23E+00 | 5.11E-02 | *-1.02E+01* | *2.65E-12* |
| | F22 | | F23 | | F24 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | -1.04E+01 | 9.95E-07 | -1.05E+01 | 7.81E-07 | 1.19E+05 | 6.92E+04 |
| PSO | -7.27E+00 | 1.19E+00 | -7.50E+00 | 1.12E+00 | 9.22E+04 | 4.53E+04 |
| WOA | -1.04E+01 | 2.57E-06 | -1.05E+01 | 5.21E-07 | 1.40E+06 | 1.10E+06 |
| GWO | -1.00E+01 | 1.35E+00 | -1.05E+01 | 1.01E-06 | 1.02E+06 | 2.20E+06 |
| SCA | -4.98E+00 | 3.34E+00 | -4.66E+00 | 2.44E+00 | 1.29E+06 | 9.63E+05 |
| FOA | -3.50E+00 | 8.81E-01 | -3.40E+00 | 7.13E-01 | 8.34E+08 | 3.13E+08 |
| DE | *-1.04E+01* | *1.81E-15* | *-1.05E+01* | *1.68E-15* | 3.04E+05 | 1.59E+05 |
| SSA | -9.70E+00 | 1.83E+00 | -1.04E+01 | 9.79E-01 | *5.96E+04* | *3.68E+04* |
| | F25 | | F26 | | F27 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | 2.64E+03 | 1.50E+02 | *2.50E+03* | *0.00E+00* | *2.60E+03* | *0.00E+00* |
| PSO | 2.90E+03 | 2.51E+02 | 2.62E+03 | 4.54E-01 | 2.63E+03 | 6.08E+00 |
| WOA | 2.97E+03 | 2.25E+02 | 2.63E+03 | 2.67E+01 | 2.61E+03 | 3.91E+00 |
| GWO | 2.59E+03 | 1.25E+02 | 2.64E+03 | 1.08E+01 | 2.60E+03 | 7.27E-04 |
| SCA | 2.99E+03 | 1.68E+02 | 2.66E+03 | 8.83E+00 | 2.60E+03 | 4.61E-02 |
| FOA | 1.25E+06 | 3.71E+05 | 2.50E+03 | 2.26E-06 | 2.60E+03 | 1.43E-05 |
| DE | *2.40E+03* | *7.71E+01* | 2.62E+03 | 1.39E-12 | 2.63E+03 | 3.42E+00 |
| SSA | 2.65E+03 | 2.00E+02 | 2.62E+03 | 1.77E-01 | 2.64E+03 | 8.96E+00 |
| | F28 | | F29 | | F30 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *2.70E+03* | *0.00E+00* | 2.70E+03 | 1.09E-01 | *2.90E+03* | *0.00E+00* |
| PSO | 2.71E+03 | 5.77E+00 | 2.78E+03 | 3.80E+01 | 3.45E+03 | 2.74E+02 |
| WOA | 2.72E+03 | 1.91E+01 | 2.70E+03 | 1.82E+01 | 3.72E+03 | 3.83E+02 |
| GWO | 2.71E+03 | 6.19E+00 | 2.76E+03 | 5.02E+01 | 3.37E+03 | 1.18E+02 |
| SCA | 2.73E+03 | 5.42E+00 | 2.70E+03 | 6.12E-01 | 3.44E+03 | 3.20E+02 |
| FOA | 2.70E+03 | 3.54E-08 | 2.80E+03 | 7.06E-11 | 2.90E+03 | 2.75E-07 |
| DE | 2.71E+03 | 1.04E+00 | 2.70E+03 | 4.42E-02 | 3.20E+03 | 8.08E+01 |

Table 8: Continued.

| | | | | | | |
|---|---|---|---|---|---|---|
| SSA | 2.71E+03 | 4.09E+00 | *2.70E+03* | *1.20E-01* | 3.45E+03 | 1.46E+02 |
| | F31 | | F32 | | F33 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *3.00E+03* | *0.00E+00* | *3.10E+03* | *0.00E+00* | *3.20E+03* | *0.00E+00* |
| PSO | 7.19E+03 | 7.21E+02 | 8.87E+04 | 1.18E+05 | 1.55E+04 | 6.61E+03 |
| WOA | 5.04E+03 | 6.16E+02 | 4.55E+06 | 4.93E+06 | 7.94E+04 | 4.77E+04 |
| GWO | 3.97E+03 | 2.24E+02 | 1.87E+06 | 4.50E+06 | 5.28E+04 | 3.81E+04 |
| SCA | 4.83E+03 | 3.05E+02 | 1.04E+07 | 6.28E+06 | 2.19E+05 | 8.49E+04 |
| FOA | 3.00E+03 | 6.77E-07 | 4.65E+03 | 2.03E+00 | 3.31E+03 | 1.804E-01 |
| DE | 3.64E+03 | 2.41E+01 | 2.77E+04 | 1.16E+05 | 5.82E+03 | 9.58E+02 |
| SSA | 3.79E+03 | 8.08E+01 | 2.20E+06 | 5.20E+06 | 1.20E+04 | 4.50E+03 |

Table 9: Wilcoxon signed-rank test results between GLSMA and seven original metaheuristic algorithms.

| Function | PSO | WOA | GWO | SCA | FOA | DE | SSA |
|---|---|---|---|---|---|---|---|
| F1 | 1.73E-06 | 1.00E+00 | 1.00E+00 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F2 | 1.73E-06 | 1.00E+00 | 1.00E+00 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F3 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F4 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F5 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F6 | 1.73E-06 | 7.50E-01 | 2.35E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F7 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 5.75E-06 | 1.73E-06 | 1.73E-06 |
| F8 | 1.73E-06 | 4.99E-03 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.36E-05 | 1.73E-06 |
| F9 | 1.73E-06 | 1.00E+00 | 1.00E+00 | 1.00E+00 | 1.73E-06 | 1.25E-01 | 1.73E-06 |
| F10 | 1.73E-06 | 4.15E-05 | 6.80E-08 | 1.73E-06 | 1.73E-06 | 1.96E-07 | 1.73E-06 |
| F11 | 1.73E-06 | 1.00E+00 | 5.00E-01 | 1.00E+00 | 1.73E-06 | 1.00E+00 | 1.73E-06 |
| F12 | 1.73E-06 | 5.72E-01 | 1.92E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.36E-05 |
| F13 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 6.73E-01 |
| F14 | 1.73E-06 | 6.13E-05 | 1.73E-06 | 1.73E-06 | 4.32E-08 | 4.32E-08 | 1.21E-07 |
| F15 | 1.73E-06 | 4.45E-05 | 9.75E-01 | 1.73E-06 | 1.73E-06 | 7.97E-01 | 2.60E-05 |
| F16 | 1.73E-06 | 1.55E-01 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.14E-05 | 1.34E-04 |
| F17 | 1.73E-06 | 3.16E-03 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F18 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.53E-06 | 6.87E-01 |
| F19 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F20 | 3.52E-06 | 8.13E-01 | 5.71E-02 | 1.73E-06 | 1.73E-06 | 1.92E-06 | 4.39E-03 |
| F21 | 1.73E-06 | 1.36E-01 | 7.73E-03 | 1.73E-06 | 1.73E-06 | 3.11E-05 | 1.73E-06 |
| F22 | 1.73E-06 | 4.11E-03 | 1.53E-01 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.48E-02 |
| F23 | 1.73E-06 | 4.99E-03 | 1.17E-02 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 3.11E-05 |
| F24 | 4.95E-02 | 2.60E-06 | 1.60E-04 | 1.73E-06 | 1.73E-06 | 2.88E-06 | 1.74E-04 |
| F25 | 1.48E-04 | 6.98E-06 | 2.18E-02 | 2.13E-06 | 1.73E-06 | 1.36E-05 | 8.77E-01 |
| F26 | 1.73E-06 | 2.56E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 4.32E-08 | 1.73E-06 |
| F27 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F28 | 1.73E-06 | 2.93E-04 | 1.22E-05 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F29 | 5.75E-06 | 9.84E-03 | 2.26E-03 | 1.73E-06 | 1.73E-06 | 3.18E-06 | 1.17E-02 |
| F30 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F31 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F32 | 1.73E-06 | 1.72E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F33 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| +/=/- | 32/0/1 | 22/91 | 25/7/1 | 31/2/0 | 33/0/0 | 17/3/13 | 21/3/9 |

Table 10: Average ranking values using the Friedman test.

| Algorithm | GLSMA | PSO | WOA | GWO | SCA | FOA | DE | SSA |
|---|---|---|---|---|---|---|---|---|
| AVR | *2.328283* | 6.591919 | 4.190404 | 4.187879 | 5.958081 | 5.80202 | 2.730303 | 4.211111 |
| Rank | *1* | 8 | 4 | 3 | 7 | 6 | 2 | 5 |

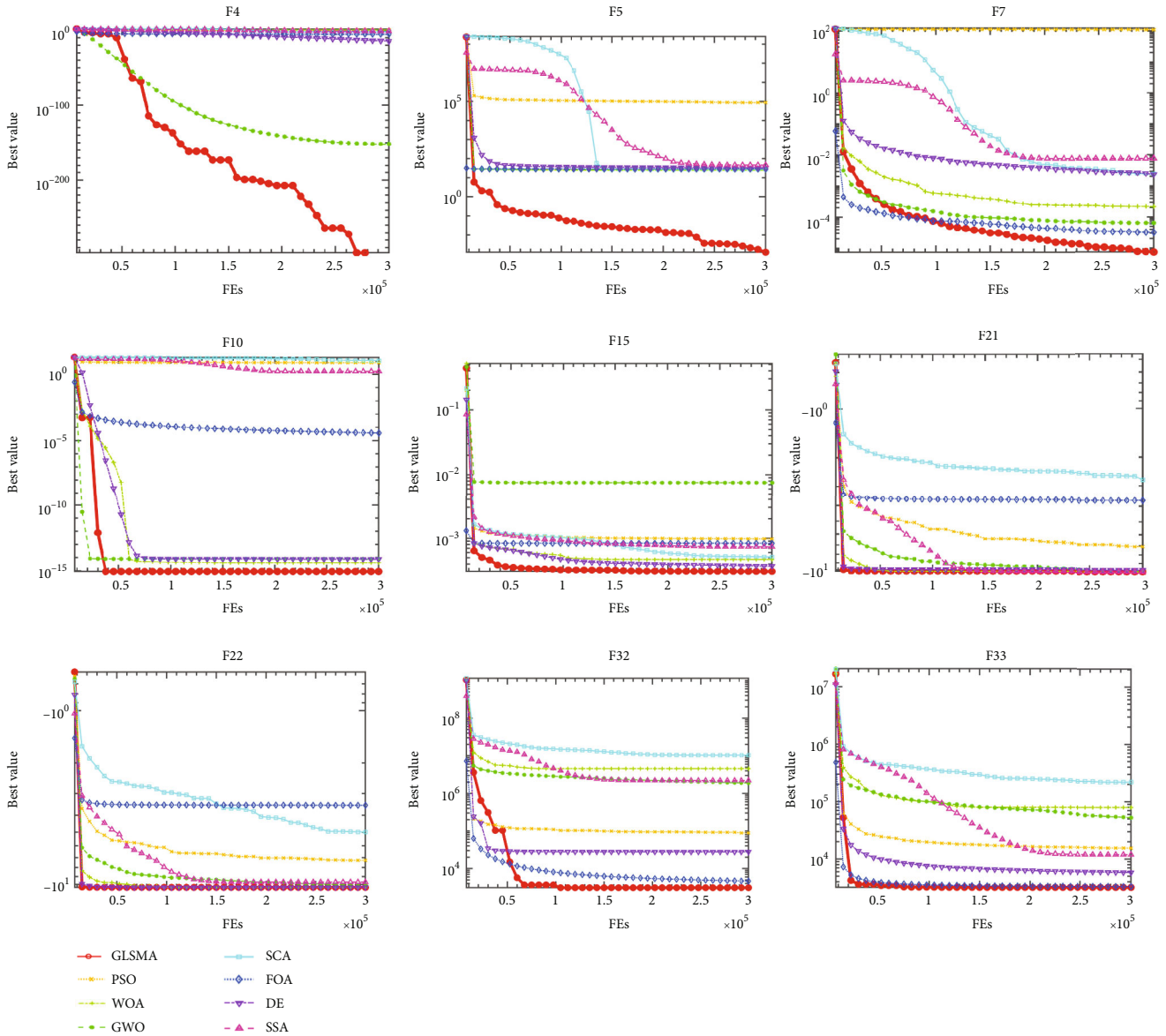



Figure 3: Convergence curves of GLSMA and seven original metaheuristic algorithms on nine selected benchmark functions.

improves the global search ability under the dual mechanism and can quickly escape from the local optimal trap and faster to find the global optimal.

In conclusion, compared with other well-known algorithms, GLSMA shows good overall superiority and stability. The strategy combination of Gaussian mutation and Levy flight enables the proposed GLSMA to obtain higher quality solutions in the optimization process, thus achieving a balance between exploration and exploitation.

*4.3. Comparison with Advanced Algorithms.* In this experiment, the proposed GLSMA algorithm is compared with 8 classical advanced algorithms in order to fully prove its global search and avoiding local optimality, including MPEDE [59], LSHADE [60], ALCPSO [61], CLPSO [62], CMAES [63], BMWOA [64], CESCA [65], and IGWO [66]. These include two classic DE variants, two superior PSO variants, and variations of WOA and GWO algorithms.

Table 11: Experimental results of GLSMA and eight advanced algorithms on 33 benchmark functions.

| | F1 | | F2 | | F3 | |
|---|---|---|---|---|---|---|
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *0.00E+00* | *0.00E+00* | *0.00E+00* | *0.00E+00* | *0.00E+00* | *0.00E+00* |
| MPEDE | 2.37E-225 | 0.00E+00 | 1.61E-102 | 8.82E-102 | 1.61E-102 | 8.82E-102 |
| LSHADE | 1.22E-202 | 0.00E+00 | 1.46E-86 | 4.69E-86 | 1.46E-86 | 4.69E-86 |
| ALCPSO | 6.12E-186 | 0.00E+00 | 6.26E-05 | 2.21E-04 | 6.26E-05 | 2.21E-04 |
| CLPSO | 3.07E-34 | 2.84E-34 | 6.48E-21 | 2.80E-21 | 6.48E-21 | 2.80E-21 |
| CMAES | 1.93E-29 | 1.55E-30 | 5.24E-02 | 2.49E-01 | 5.24E-02 | 2.49E-01 |
| BMWOA | 4.41E-04 | 7.36E-04 | 6.11E-03 | 4.84E-03 | 6.11E-03 | 4.84E-03 |
| CESCA | 1.13E+03 | 7.79E+02 | 8.02E+00 | 1.94E+00 | 8.02E+00 | 1.94E+00 |
| IGWO | 0.00E+00 | 0.00E+00 | 3.84E-260 | 0.00E+00 | 3.84E-260 | 0.00E+00 |
| | F4 | | F5 | | F6 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *0.00E+00* | *0.00E+00* | *1.69E-03* | *1.45E-03* | 5.79E-06 | 3.48E-06 |
| MPEDE | 7.92E-06 | 1.05E-05 | 1.46E+00 | 1.95E+00 | 3.08E-33 | 4.05E-33 |
| LSHADE | 1.34E-04 | 1.91E-04 | 1.33E+00 | 1.91E+00 | 2.36E-33 | 4.10E-33 |
| ALCPSO | 3.92E-05 | 4.38E-05 | 3.59E+01 | 3.40E+01 | 2.33E-31 | 4.95E-31 |
| CLPSO | 1.32E+00 | 2.40E-01 | 4.82E-01 | 4.50E-01 | *0.00E+00* | *0.00E+00* |
| CMAES | 2.11E-15 | 1.18E-16 | 3.99E-01 | 1.22E+00 | 2.00E-29 | 1.64E-30 |
| BMWOA | 4.50E-03 | 4.85E-03 | 2.26E-02 | 5.96E-02 | 1.54E-03 | 3.03E-03 |
| CESCA | 2.00E+01 | 7.77E+00 | 2.88E+05 | 3.21E+05 | 1.25E+03 | 9.97E+02 |
| IGWO | 1.33E-38 | 7.27E-38 | 2.32E+01 | 2.03E-01 | 1.19E-05 | 4.26E-06 |
| | F7 | | F8 | | F9 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *7.92E-06* | *5.81E-06* | *-1.26E+04* | *9.65E-03* | *0.00E+00* | *0.00E+00* |
| MPEDE | 2.98E-03 | 1.02E-03 | -1.19E+04 | 3.27E+02 | 7.67E+00 | 4.62E+00 |
| LSHADE | 6.16E-03 | 3.19E-03 | -1.89E+03 | 2.65E+01 | 2.06E+00 | 3.60E+00 |
| ALCPSO | 8.49E-02 | 2.98E-02 | -1.14E+04 | 2.88E+02 | 2.07E+01 | 7.79E+00 |
| CLPSO | 2.74E-03 | 7.42E-04 | -1.26E+04 | 2.16E+01 | 0.00E+00 | 0.00E+00 |
| CMAES | 5.52E-02 | 1.56E-02 | -7.13E+03 | 8.61E+02 | 2.30E+02 | 5.09E+01 |
| BMWOA | 1.17E-03 | 9.76E-04 | -1.26E+04 | 1.03E-02 | 3.69E-04 | 6.17E-04 |
| CESCA | 4.93E-01 | 3.46E-01 | -3.97E+03 | 2.51E+02 | 5.46E+01 | 1.26E+01 |
| IGWO | 3.76E-04 | 4.52E-04 | -7.61E+03 | 6.63E+02 | 0.00E+00 | 0.00E+00 |
| | F10 | | F11 | | F12 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *8.88E-16* | *0.00E+00* | *0.00E+00* | *0.00E+00* | 1.68E-06 | 2.71E-06 |
| MPEDE | 1.60E+00 | 8.45E-01 | 2.02E-02 | 2.11E-02 | 1.77E-01 | 3.04E-01 |
| LSHADE | 3.38E-14 | 3.72E-15 | 1.43E-02 | 1.73E-02 | 7.08E-01 | 8.66E-01 |
| ALCPSO | 6.93E-01 | 8.23E-01 | 1.47E-02 | 1.73E-02 | 1.83E-02 | 3.50E-02 |
| CLPSO | 1.24E-14 | 2.59E-15 | 0.00E+00 | 0.00E+00 | *1.57E-32* | *5.57E-48* |
| CMAES | 1.94E+01 | 1.33E-01 | 1.23E-03 | 2.80E-03 | 1.07E-30 | 1.63E-31 |
| BMWOA | 2.11E-03 | 2.22E-03 | 9.96E-04 | 3.66E-03 | 7.11E-06 | 2.34E-05 |
| CESCA | 6.61E+00 | 1.41E+00 | 9.54E+00 | 5.47E+00 | 9.55E+04 | 1.70E+05 |
| IGWO | 4.91E-15 | 1.23E-15 | 0.00E+00 | 0.00E+00 | 1.09E-06 | 4.02E-07 |
| | F13 | | F14 | | F15 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *3.08E-06* | *2.99E-06* | 9.98E-01 | 5.77E-16 | *3.09E-04* | *6.72E-06* |
| MPEDE | 3.97E-03 | 6.58E-03 | *9.98E-01* | *0.00E+00* | 2.37E-03 | 6.10E-03 |
| LSHADE | 4.20E-01 | 1.14E+00 | 9.98E-01 | 0.00E+00 | 9.80E-04 | 3.64E-05 |

Table 11: Continued.

| | | | | | | |
|---|---|---|---|---|---|---|
| ALCPSO | 7.73E-03 | 1.12E-02 | 9.98E-01 | 1.17E-16 | 3.69E-04 | 2.32E-04 |
| CLPSO | 1.35E-32 | 5.57E-48 | 9.98E-01 | 0.00E+00 | 5.50E-04 | 6.96E-05 |
| CMAES | 7.32E-04 | 2.79E-03 | 1.25E+01 | 6.58E+00 | 1.16E-02 | 2.06E-02 |
| BMWOA | 9.90E-05 | 1.76E-04 | 9.98E-01 | 3.13E-16 | 3.69E-04 | 2.32E-04 |
| CESCA | 7.14E+05 | 1.20E+06 | 2.87E+00 | 3.70E-01 | 4.38E-03 | 2.87E-03 |
| IGWO | 1.47E-02 | 3.56E-02 | 9.98E-01 | 3.01E-15 | 3.61E-04 | 2.05E-04 |
| | F16 | | F17 | | F18 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | -1.03E+00 | 2.81E-14 | 3.98E-01 | 4.35E-12 | 3.00E+00 | 1.26E-14 |
| MPEDE | -1.03E+00 | 6.78E-16 | *3.98E-01* | *0.00E+00* | 3.00E+00 | 1.87E-15 |
| LSHADE | -1.03E+00 | 6.78E-16 | 3.98E-01 | 0.00E+00 | 3.00E+00 | 1.20E-15 |
| ALCPSO | -1.03E+00 | 6.39E-16 | 3.98E-01 | 0.00E+00 | 3.00E+00 | 2.66E-15 |
| CLPSO | -1.03E+00 | 6.65E-16 | 3.98E-01 | 1.15E-15 | *3.00E+00* | *1.17E-15* |
| CMAES | -9.72E-01 | 2.07E-01 | 3.98E-01 | 0.00E+00 | 3.46E+01 | 1.53E+02 |
| BMWOA | *-1.0E+00* | *4.54E-16* | 3.98E-01 | 2.57E-13 | 3.00E+00 | 1.30E-14 |
| CESCA | -1.03E+00 | 5.79E-03 | 8.44E-01 | 3.91E-01 | 3.06E+00 | 8.45E-02 |
| IGWO | -1.03E+00 | 5.32E-13 | 3.98E-01 | 6.44E-11 | 3.00E+00 | 4.95E-14 |
| | F19 | | F20 | | F21 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | -3.86E+00 | 4.37E-11 | -3.23E+00 | 4.84E-02 | *-1.02E+01* | *1.22E-06* |
| MPEDE | -3.86E+00 | 2.71E-15 | -3.27E+00 | 5.99E-02 | -8.81E+00 | 2.53E+00 |
| LSHADE | -3.86E+00 | 7.59E-11 | -1.63E+00 | 1.65E+00 | -8.81E+00 | 2.53E+00 |
| ALCPSO | *-3.86E+00* | *2.63E-15* | -3.27E+00 | 5.92E-02 | -8.63E+00 | 2.36E+00 |
| CLPSO | -3.86E+00 | 2.71E-15 | *-3.32E+00* | *1.21E-12* | -1.02E+01 | 1.53E-06 |
| CMAES | -3.58E+00 | 8.73E-01 | -3.30E+00 | 4.84E-02 | -5.92E+00 | 3.77E+00 |
| BMWOA | -3.86E+00 | 2.26E-14 | -3.26E+00 | 6.05E-02 | -1.02E+01 | 3.44E-11 |
| CESCA | -3.53E+00 | 2.38E-01 | -2.05E+00 | 4.88E-01 | -8.77E-01 | 3.53E-01 |
| IGWO | -3.86E+00 | 2.66E-09 | -3.26E+00 | 6.05E-02 | -9.31E+00 | 1.93E+00 |
| | F22 | | F23 | | F24 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | -1.04E+01 | 1.03E-06 | -1.05E+01 | 1.23E-06 | 1.14E+05 | 8.22E+04 |
| MPEDE | -9.10E+00 | 2.70E+00 | -1.03E+01 | 1.48E+00 | 3.06E+03 | 4.43E+02 |
| LSHADE | -9.93E+00 | 1.82E+00 | -1.00E+01 | 1.89E+00 | *2.80E+03* | *2.66E+02* |
| ALCPSO | -9.52E+00 | 2.00E+00 | -10.00E+00 | 1.64E+00 | 1.32E+05 | 2.18E+05 |
| CLPSO | -1.04E+01 | 1.24E-06 | -1.05E+01 | 5.43E-10 | 8.77E+04 | 6.01E+04 |
| CMAES | -5.70E+00 | 3.22E+00 | -7.16E+00 | 3.75E+00 | 3.18E+03 | 3.38E+02 |
| BMWOA | *-1.04E+01* | *3.50E-11* | *-1.05E+01* | *2.87E-11* | 1.10E+06 | 1.12E+06 |
| CESCA | -1.18E+00 | 4.67E-01 | -1.31E+00 | 5.62E-01 | 3.40E+07 | 1.28E+07 |
| IGWO | -9.70E+00 | 1.83E+00 | -10.00E+00 | 1.64E+00 | 3.08E+05 | 2.48E+05 |
| | F25 | | F26 | | F27 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | 2.54E+03 | 1.70E+02 | *2.50E+03* | *0.00E+00* | *2.60E+03* | *0.00E+00* |
| MPEDE | 2.54E+03 | 1.54E+02 | 2.62E+03 | 1.68E-12 | 2.64E+03 | 5.74E+00 |
| LSHADE | 2.45E+03 | 1.17E+02 | 2.62E+03 | 2.13E-12 | 2.64E+03 | 5.18E+00 |
| ALCPSO | 2.67E+03 | 1.68E+02 | 2.62E+03 | 2.26E-02 | 2.64E+03 | 6.89E+00 |
| CLPSO | *2.42E+03* | *7.90E+01* | 2.62E+03 | 3.14E-06 | 2.62E+03 | 7.39E+00 |
| CMAES | 2.54E+03 | 2.59E+02 | 2.62E+03 | 1.39E-12 | 2.661E+03 | 9.29E+01 |
| BMWOA | 3.00E+03 | 2.03E+02 | 2.50E+03 | 4.14E-01 | 2.60E+03 | 1.99E-01 |
| CESCA | 5.95E+03 | 1.75E+03 | 3.03E+03 | 1.03E+02 | 2.66E+03 | 1.74E+01 |

Table 11: Continued.

| | | | | | | |
|---|---|---|---|---|---|---|
| IGWO | 2.55E+03 | 1.48E+02 | 2.62E+03 | 2.36E+00 | 2.60E+03 | 4.00E-03 |
| | F28 | | F29 | | F30 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *2.70E+03* | *0.00E+00* | 2.70E+03 | 1.27E-01 | *2.90E+03* | *0.00E+00* |
| MPEDE | 2.71E+03 | 5.13E+00 | 2.71E+03 | 3.04E+01 | 3.27E+03 | 1.51E+02 |
| LSHADE | 2.71E+03 | 3.82E+00 | 2.71E+03 | 3.45E+01 | 3.27E+03 | 9.40E+01 |
| ALCPSO | 2.71E+03 | 3.70E+00 | 2.76E+03 | 4.88E+01 | 3.47E+03 | 2.23E+02 |
| CLPSO | 2.71E+03 | 9.97E-01 | *2.70E+03* | *6.83E-02* | 3.13E+03 | 4.63E+01 |
| CMAES | 2.70E+03 | 2.03E+00 | 2.72E+03 | 5.81E+01 | 3.07E+03 | 3.98+01 |
| BMWOA | 2.70E+03 | 5.68E-03 | 2.71E+03 | 1.04E-01 | 2.90E+03 | 1.63E-01 |
| CESCA | 2.72E+03 | 8.91E+00 | 2.71E+03 | 1.44+00 | 3.98E+03 | 1.58E+02 |
| IGWO | 2.71E+03 | 2.75E+00 | 2.70E+03 | 1.71E-01 | 3.11E+03 | 4.03E+00 |
| | F31 | | F32 | | F33 | |
| | Avg | Std | Avg | Std | Avg | Std |
| GLSMA | *3.00E+03* | *0.00E+00* | *3.10E+03* | *0.00E+00* | *3.20E+03* | *0.00E+00* |
| MPEDE | 4.00E+03 | 3.22E+02 | 1.72E+06 | 3.94E+06 | 5.62E+03 | 1.14E+03 |
| LSHADE | 3.77E+03 | 1.41E+02 | 3.71E+03 | 1.61E+02 | 5.56E+03 | 9.79E+02 |
| ALCPSO | 4.31E+03 | 4.70E+02 | 1.45E+06 | 4.41E+06 | 1.82E+04 | 2.92E+04 |
| CLPSO | 3.70E+03 | 5.81E+01 | 3.86E+03 | 1.14E+02 | 6.15E+03 | 8.22E+02 |
| CMAES | 6.49E+03 | 2.96E+03 | 3.67E+03 | 1.31E+02 | 5.23E+03 | 6.37E+02 |
| BMWOA | 3.00E+03 | 5.17E-01 | 6.28E+05 | 6.18E+05 | 6.17E+04 | 5.38E+04 |
| CESCA | 5.43E+03 | 2.83E+02 | 1.78E+07 | 3.48E+06 | 1.42E+06 | 2.43E+05 |
| IGWO | 3.86E+03 | 1.94E+02 | 7.34E+05 | 2.75E+06 | 2.69E+04 | 1.15E+04 |

The comparison results of GLSMA with eight advanced algorithms are shown in Tables 11–13. Table 11 shows the average value and standard deviation of the optimal solution obtained by GLSMA and advanced algorithms. As can be seen, compared to other algorithms, GLSMA shows good superiority and stability in F1-F5, F7-F11, F13, F15, F21, F26-F28, and F30-33 functions. Table 12 shows the Wilcoxon rank test result's $p$ value among GLSMA and eight advanced algorithms. From the table values, it can be seen that GLSMA outperforms other comparison algorithms on most benchmark functions. GLSMA is superior to CESCA in all functions. Therefore, GLSMA is obviously competitive with other excellent algorithms. Table 13 shows the comparison results; among GLSMA and other 8 advanced algorithms, the average Friedman test result of GLSMA ranks the first, which is 3.629293.

The convergence curves of all nine algorithms on nine functions shown in Figure 4 show that GLSMA's convergence speed is faster than other advanced algorithms, and it can jump out of local optimum faster and avoid falling into local optimum better than other algorithms.

In summary, the introduction of Gaussian mutation mechanism and Levy flight mechanism gives GLSMA an advantage over competitive advanced algorithms, showing superior performance in different types of functions. GLSMA not only has stronger global search ability but also can avoid falling into local optimum.

*4.4. The Experiments for Feature Selection.* In this section, we transform the proposed algorithm GLSMA into a discrete version, namely, BGLSMA, which is applied to feature selection problems of high-dimensional gene data, thus making the proposed algorithm more realistic.

The purpose of feature selection problem is to remove some redundant and irrelevant features from the sample, so as to reduce the complexity of feature selection problem, reduce the subsequent calculation cost, and obtain higher classification accuracy. In the process of feature selection, it is necessary to determine which features should be selected. As a result, we transform continuous GLSMA into discrete GLSMA, namely, BGLSMA. The proposed GLSMA increases the population diversity, strengths the local exploitation ability, and helps us select favorable features in the search space, so as to obtain better feature subsets and improve classification accuracy.

*4.4.1. Binary GLSMA.* In feature selection algorithm based on GLSMA, $x = (x_{i,1}, x_{i,2}, \cdots, x_{i,n})$ represents a subset of features. In BGLSMA, if $x_{i,1} = 1$, this feature is selected, and conversely, if $x_{i,1} = 0$, this feature is discarded. In order to solve discretization problems, GLSMA needs to be discretized. The individual with binary position vector is initialized by random threshold, and then, the discretization of position $X_i$ can be expressed as

$$X_i^j(t+1) = \begin{cases} 1, \text{ if rand} \geq T\left(X_i^j(t)\right), \\ 0, \text{ if rand} < T\left(X_i^j(t)\right), \end{cases} \quad (13)$$

$$T(x) = \frac{1}{1 + e^{-2x}}$$

Table 12: Wilcoxon signed-rank test results between GLSMA and eight advanced algorithms.

| Function | MPEDE | LSHADE | ALCPSO | CLPSO | CMAES | BMWOA | CESCA | IGWO |
|---|---|---|---|---|---|---|---|---|
| F1 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.00E+00 |
| F2 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F3 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F4 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F5 | 3.82E-01 | 6.44E-01 | 1.73E-06 | 1.73E-06 | 2.77E-03 | 1.59E-01 | 1.73E-06 | 1.73E-06 |
| F6 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 5.75E-06 | 1.73E-06 | 5.22E-06 |
| F7 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F8 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 3.11E-05 | 1.73E-06 | 3.68E-02 | 1.73E-06 | 1.73E-06 |
| F9 | 1.73E-06 | 4.92E-06 | 1.73E-06 | 1.00E+00 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.00E+00 |
| F10 | 1.71E-06 | 9.53E-07 | 1.44E-06 | 1.12E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.96E-07 |
| F11 | 2.69E-05 | 8.79E-05 | 5.93E-05 | 1.00E+00 | 6.25E-02 | 1.73E-06 | 1.73E-06 | 1.00E+00 |
| F12 | 4.72E-02 | 5.31E-05 | 3.82E-01 | 1.73E-06 | 1.73E-06 | 1.92E-01 | 1.73E-06 | 5.72E-01 |
| F13 | 9.75E-01 | 3.71E-01 | 8.73E-03 | 1.73E-06 | 3.59E-04 | 9.71E-05 | 1.73E-06 | 1.73E-06 |
| F14 | 1.21E-07 | 1.21E-07 | 1.21E-07 | 1.21E-07 | 1.90E-06 | 2.75E-04 | 1.73E-06 | 6.33E-06 |
| F15 | 5.71E-02 | 1.73E-06 | 3.59E-04 | 1.73E-06 | 1.15-04 | 3.59E-04 | 1.73E-06 | 3.59E-04 |
| F16 | 3.20E-06 | 3.20E-06 | 4.85E-06 | 3.20E-06 | 6.95E-04 | 7.32E-06 | 1.73E-06 | 6.98E-06 |
| F17 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 2.84E-05 | 1.73E-06 | 2.16E-05 |
| F18 | 1.49E-06 | 1.54E-06 | 2.18E-06 | 1.54E-06 | 1.97E-02 | 1.12E-04 | 1.73E-06 | 6.62E-02 |
| F19 | 1.73E-06 | 3.11E-05 | 1.73E-06 | 1.73E-06 | 2.77E-03 | 1.73E-06 | 1.73E-06 | 1.92E-06 |
| F20 | 1.02E-05 | 1.38E-03 | 9.32E-06 | 1.73E-06 | 5.22E-06 | 1.80E-05 | 1.73E-06 | 5.30E-01 |
| F21 | 3.71E-01 | 3.71E-01 | 9.75E-01 | 1.60E-04 | 3.61E-03 | 1.73E-06 | 1.73E-06 | 8.47E-06 |
| F22 | 1.65E-01 | 3.59E-04 | 5.71E-02 | 3.11E-05 | 1.15E-04 | 1.73E-06 | 1.73E-06 | 6.34E-06 |
| F23 | 3.11E-05 | 3.59E-04 | 2.77E-03 | 1.73E-06 | 4.72E-02 | 1.73E-06 | 1.73E-06 | 2.16E-05 |
| F24 | 1.73E-06 | 1.73E-06 | 2.13E-01 | 1.36E-01 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.48E-03 |
| F25 | 9.92E-01 | 1.67E-02 | 3.33E-02 | 3.16E-03 | 6.58E-01 | 2.3534E-06 | 1.73E-06 | 8.77E-01 |
| F26 | 4.32E-08 | 4.32E-08 | 1.73E-06 | 1.73E-06 | 4.32E-08 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F27 | 1.73E-06 | 1.73-06 | 1.73E-06 | 2.13E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F28 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F29 | 9.78E-02 | 1.59E-01 | 8.92E-05 | 7.51E-05 | 8.97E-02 | 1.48E-02 | 1.73E-06 | 2.60E-05 |
| F30 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F31 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F32 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| F33 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 | 1.73E-06 |
| +/=/- | 18/7/8 | 21/4/8 | 22/4/7 | 16/3/14 | 25/3/5 | 21/2/10 | 33/0/0 | 26/7/0 |

Table 13: Average ranking values using the Friedman test.

| Algorithm | GLSMA | MPEDE | LSHADE | ALCPSO | CLPSO | CMAES | BMWOA | CESCA | IGWO |
|---|---|---|---|---|---|---|---|---|---|
| AVR | *3.62929* | 3.84293 | 4.20808 | 5.25202 | 3.93131 | 4.83384 | 5.16212 | 8.75455 | 5.38586 |
| Rank | *1* | 2 | 4 | 7 | 3 | 5 | 6 | 9 | 8 |

In the formula above, $X_i(t+1)$ indicates the value of the $i$-th dimension of the agent individual position $X$ searched in the discrete space, and rand means a random number within the range of [0, 1]. $T(x)$ means converting the value of the $i$-th dimension of $X$ in the continuous motion space to 0 or 1, thus realizing the discretization of the continuous space. The transformation of $T(x)$ does not change the structure of the algorithm.

As described above, feature selection is a process of using the least gene subset to obtain the optimal classification accuracy, that is, to improve the classification accuracy and reduce the number of features. This problem is described as a combinatorial optimization problem. In order to satisfy each objective, a linear combination of feature number and error rate is used to define fitness function, and the candidate solutions are evaluated

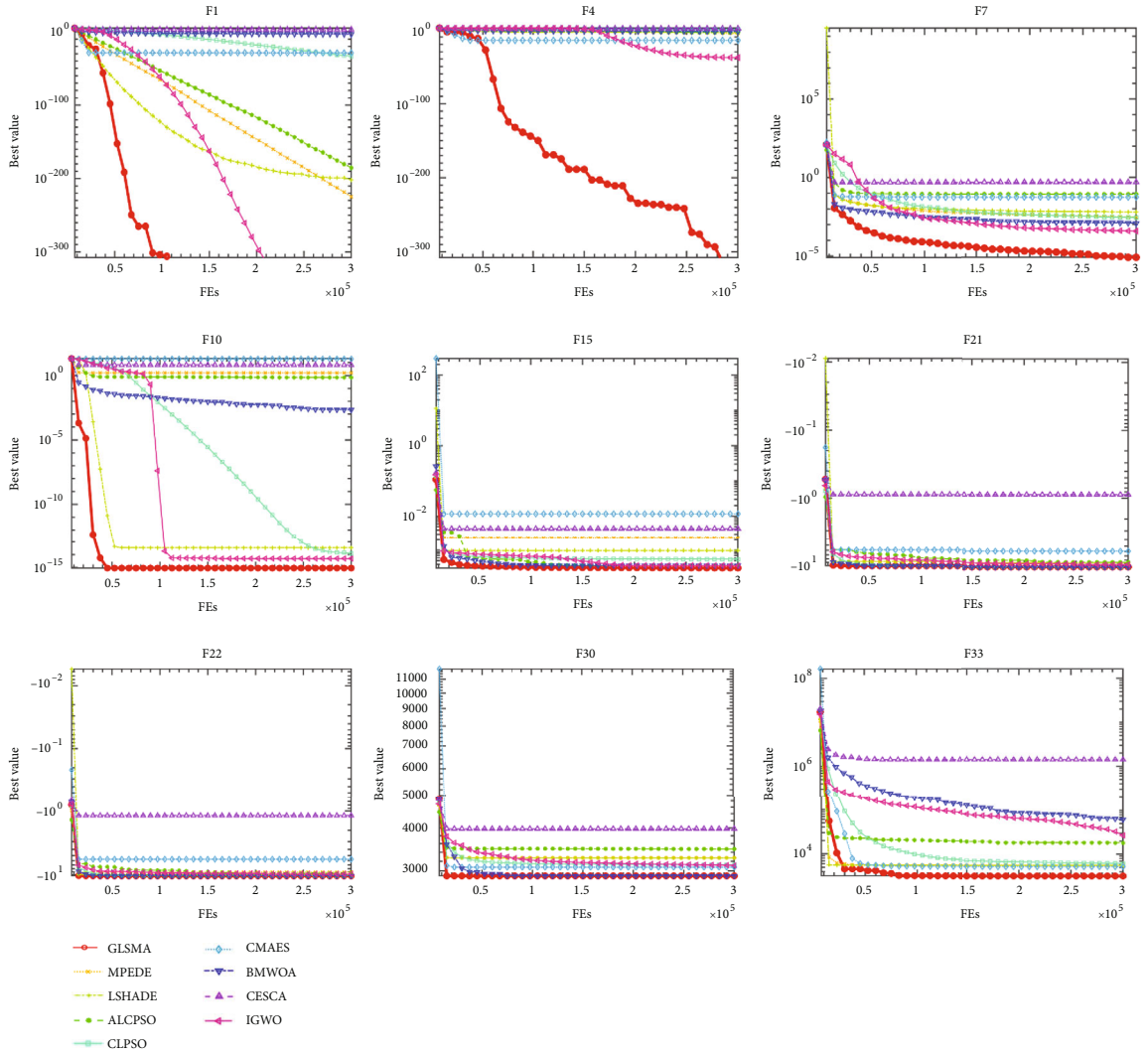



Figure 4: Convergence curves of GLSMA and eight advanced algorithms on nine selected benchmark functions.

comprehensively.

$$\text{fit} = \alpha \cdot E + \beta \cdot \frac{R}{D}. \tag{14}$$

In the above formula, $E$ is the classification accuracy rate of KNN classifier, the length of selected feature subset is represented by $R$, and the total number of features in the dataset is represented by $D$. $\alpha$ and $\beta$ are the weights of classification error rate and feature reduction, respectively. Compared with feature reduction, more attention is paid to accuracy; we set $\alpha$ to 0.95 and the latter to 0.05.

*4.4.2. Simulation Experiments.* In this experiment, the resulting BGLSMA is compared with other excellent metaheuristic optimizers on 14 UCI feature selection datasets. In Table 14, the details of these datasets are shown, including the number of samples, the number of features, and the number of categories. Table 14 shows that, in these datasets, the sample

Table 14: Characteristics of gene expression datasets.

| Datasets | Samples | Features | Categories |
|---|---|---|---|
| Colon | 62 | 2000 | 2 |
| SRBCT | 83 | 2309 | 4 |
| Leukemia | 72 | 7131 | 2 |
| Brain_Tumor1 | 90 | 5920 | 5 |
| Brain_Tumor2 | 50 | 10368 | 4 |
| CNS | 60 | 7130 | 2 |
| DLBCL | 77 | 5470 | 2 |
| Leukemia1 | 72 | 5328 | 3 |
| Leukemia2 | 72 | 11226 | 3 |
| Lung_Cancer | 203 | 12601 | 3 |
| Prostate_Tumor | 102 | 10509 | 2 |
| Tumors_9 | 60 | 5726 | 9 |
| Tumors_11 | 174 | 12534 | 11 |
| Tumors_14 | 308 | 15010 | 26 |

Table 15: Comparison between BGLSMA and other FS optimizers on average number of the selected features.

| Datasets | Metrics | bGWO | BBA | BGSA | BPSO | bALO | BSSA | BSMA | BGLSMA |
|---|---|---|---|---|---|---|---|---|---|
| Colon | Std | 20.7964 | 54.9142 | 26.5991 | 12.8966 | 18.8043 | 318.1085 | 17.0098 | *0.52705* |
| | Avg | 165.4 | 781.7 | 768.8 | 891.9 | 856.6 | 768.1 | 32 | *1.5* |
| SRBCT | Std | 12.3306 | 67.0439 | 18.7723 | 21.6972 | 21.4828 | 336.9647 | 24.8697 | *6.1968* |
| | Avg | 186.6 | 933 | 906.8 | 1029.1 | 998.2 | 978.9 | 34.5 | *11.2* |
| Leukemia | Std | 45.8786 | 186.7911 | 33.7377 | 22.1761 | 34.2379 | 1602.754 | 12.5526 | *0.5164* |
| | Avg | 793.2 | 2991.7 | 3119.7 | 3334 | 3269.3 | 2046.7 | 39.3 | *1.4* |
| Brain_Tumor1 | Std | 52.7674 | 163.1894 | 36.7792 | 51.0565 | 60.6781 | 1004.231 | 73.4078 | *7.5491* |
| | Avg | 623.2 | 2393.1 | 2564.6 | 2771.1 | 2732.5 | 2164.1 | 88.6 | *6.9* |
| Brain_Tumor2 | Std | 57.3682 | 129.1005 | 45.7117 | 57.4886 | 52.4006 | 1816.777 | 114.1627 | *5.1251* |
| | Avg | 1191 | 1191 | 1191 | 1191 | 1191 | 1191 | 1191 | *1191* |
| CNS | Std | 35.9365 | 367.5338 | 59.1534 | 53.6694 | 68.2541 | 1448.946 | 116.9699 | *1.4757* |
| | Avg | 863.9 | 2848.3 | 3202.7 | 3391.8 | 3340.2 | 2456.3 | 199.2 | *2.2* |
| DLBCL | Std | 20.1594 | 153.9154 | 40.0012 | 32.3934 | 21.6705 | 1104.873 | 13.0826 | *1.0593* |
| | Avg | 567.2 | 2231.2 | 2358.1 | 2545 | 2495.5 | 2102.2 | 28.6 | *1.7* |
| Leukemia1 | Std | 31.8462 | 133.4618 | 33.8856 | 36.5345 | 25.776 | 1035.249 | 62.1035 | *20.1* |
| | Avg | 564.8 | 2150.6 | 2283.7 | 2477.1 | 2433.8 | 1785.5 | 70.8 | *13.7* |
| Leukemia2 | Std | 85.2763 | 256.4667 | 67.8352 | 47.7568 | 53.6723 | 2515.541 | 61.4167 | *3.8355* |
| | Avg | 1253.6 | 4632.5 | 5042.5 | 5335.6 | 5279.6 | 2553 | 76.3 | *4.6* |
| Lung_Cancer | Std | 117.4218 | 256.5553 | 73.7741 | 51.4295 | 46.6363 | 2856.425 | 182.022 | *22.5982* |
| | Avg | 1548.9 | 5207.2 | 5766.8 | 6020.9 | 5957.5 | 4110.1 | 206.3 | *34.3* |
| Prostate_Tumor | Std | 56.0124 | 607.9035 | 107.2774 | 58.1687 | 63.221 | 2544.318 | 128.1768 | *2.5927* |
| | Avg | 1281.5 | 4093.7 | 4834 | 5058.4 | 4974.7 | 2748.4 | 171.8 | *4.5* |
| Tumors_9 | Std | 38.0935 | 139.8954 | 104.17 | 53.1402 | 48.8599 | 692.926 | 274.6595 | *50.4341* |
| | Avg | 661 | 2320.6 | 2572.5 | 2701.9 | 2661.8 | 2570.3 | 374.4 | *28.4* |
| Tumors_11 | Std | 97.3416 | 668.8177 | 117.447 | 77.4884 | *76.3216* | 2645.298 | 308.9831 | 1207.602 |
| | Avg | 1639.5 | 1639.5 | 1639.5 | 1639.5 | *1639.5* | 1639.5 | 1639.5 | 1639.5 |
| Tumors_14 | Std | 80.6821 | 499.8572 | 120.4111 | *57.0264* | 77.2704 | 1516.641 | 1914.11 | 2399.334 |
| | Avg | 2271.6 | 2271.6 | 2271.6 | *2271.6* | 2271.6 | 2271.6 | 2271.6 | 2271.6 |
| ARV | | 2.9286 | 4.9286 | 5.8571 | 8 | 6.9286 | 4.2857 | 2 | *1.0714* |
| Rank | | 3 | 5 | 6 | 8 | 7 | 4 | 2 | *1* |

number is 50-308, the feature number is 2000-15010, and the number of categories is 2-11. These datasets contain several different types of data. These high-dimensional gene datasets have such characteristics: the sample number is small, and the feature number is thousands, which has some impact on data dimension reduction.

In order to select fewer features while maintaining classification accuracy, K-nearest neighbor (KNN) [67] algorithm is used for data classification. K-nearest neighbor is a nonparametric regression statistical method with wide application in classification problems. The steps of KNN algorithm are as follows: firstly, the original data is preprocessed, and the processed data is divided into training set and test set; second, set the appropriate parameter $k$ to 1. Then, the initial group is selected in the training set, and the distance $D$ between the initial group and the test group is calculated. The distance calculation formula is shown in Equation (15). At the same time, calculate the distance $L$ between the training group and the test group, and compare whether $L$ is less than $D$. If $L$ is less than $D_{\max}$, repeat the above steps until the termination condition is reached.

$$D(x, y) = \left( \sum_{k}^{N} (x_k - y_k) \right)^{1/2}. \tag{15}$$

Metaheuristic classifiers used for comparison include bGWO [68], BBA [69], BGSA [70], BPSO [71], bALO [72], BSSA [73], and, the binary form of the original SMA, BSMA. The above classifiers are used for feature selection, and the relevant data of feature subsets found by various algorithms in the search process are learned in KNN classifier, and the corresponding result information is finally output for comparison. In order to reduce the influence of random factors, 10-fold cross validation was adopted, and the average value of multiple cross experimental results was taken as the result to evaluate the algorithm's accuracy.

Tables 15–18 list the statistical results of the average number of selected features, the average error rate, the average fitness, and the average calculation time. According to

Table 16: Comparison between BGLSMA and other FS optimizers on average error rate.

| Datasets | Metrics | bGWO | BBA | BGSA | BPSO | bALO | BSSA | BSMA | BGLSMA |
|---|---|---|---|---|---|---|---|---|---|
| Colon | Std | 6.549E-02 | 2.097E-01 | 1.277E-01 | 1.241E-01 | 1.365E-01 | 1.236E-01 | 0.000E+00 | *0.000E+00* |
| | Avg | 3.095E-02 | 2.619E-01 | 1.929E-01 | 1.571E-01 | 1.619E-01 | 1.810E-01 | 0.000E+00 | *0.000E+00* |
| SRBCT | Std | 0.000E+00 | 1.416E-01 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | *0.000E+00* |
| | Avg | 0.000E+00 | 1.237E-01 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | *0.000E+00* |
| Leukemia | Std | 0.000E+00 | 6.942E-02 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | *0.000E+00* |
| | Avg | 0.000E+00 | 5.357E-02 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | *0.000E+00* |
| Brain_Tumor1 | Std | 5.197E-02 | 9.735E-02 | 7.147E-02 | 5.520E-02 | 7.216E-02 | 5.636E-02 | 4.216E-02 | *3.162E-02* |
| | Avg | 3.222E-02 | 1.081E-01 | 6.222E-02 | 5.222E-02 | 4.222E-02 | 5.333E-02 | 2.000E-02 | *1.000E-02* |
| Brain_Tumor2 | Std | 7.770E-02 | 1.775E-01 | 1.370E-01 | 9.088E-02 | 9.875E-02 | 1.012E-01 | 0.000E+00 | *0.000E+00* |
| | Avg | 3.667E-02 | 2.850E-01 | 7.667E-02 | 7.000E-02 | 5.929E-02 | 6.167E-02 | 0.000E+00 | *0.000E+00* |
| CNS | Std | 5.271E-02 | 2.084E-01 | 8.784E-02 | 1.760E-01 | 8.635E-02 | 8.988E-02 | 4.518E-02 | *0.000E+00* |
| | Avg | 1.667E-02 | 4.171E-01 | 8.333E-02 | 1.143E-01 | 5.333E-02 | 8.429E-02 | 1.429E-02 | *0.000E+00* |
| DLBCL | Std | 3.953E-02 | 8.996E-02 | 4.518E-02 | 4.518E-02 | 3.953E-02 | 0.000E+00 | 0.000E+00 | *0.000E+00* |
| | Avg | 1.250E-02 | 7.857E-02 | 1.429E-02 | 1.429E-02 | 1.250E-02 | 0.000E+00 | 0.000E+00 | *0.000E+00* |
| Leukemia1 | Std | 0.000E+00 | 1.214E-01 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | *0.000E+00* |
| | Avg | 0.000E+00 | 7.143E-02 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | *0.000E+00* |
| Leukemia2 | Std | 0.000E+00 | 9.989E-02 | 0.000E+00 | 3.953E-02 | 0.000E+00 | 6.023E-02 | 0.000E+00 | *0.000E+00* |
| | Avg | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | 0.000E+00 | *0.000E+00* |
| Lung_Cancer | Std | 2.385E-02 | 4.691E-02 | 2.491E-02 | 3.310E-02 | 2.587E-02 | 3.376E-02 | 3.012E-02 | *1.664E-02* |
| | Avg | 1.479E-02 | 6.862E-02 | 1.929E-02 | 1.883E-02 | 2.452E-02 | 1.952E-02 | 9.524E-03 | *5.263E-03* |
| Prostate_Tumor | Std | 5.750E-02 | 1.470E-01 | 9.661E-02 | 5.182E-02 | 6.654E-02 | 6.542E-02 | 6.650E-02 | *0.000E+00* |
| | Avg | 1.818E-02 | 1.818E-02 | 1.818E-02 | 1.818E-02 | 1.818E-02 | 1.818E-02 | 1.818E-02 | *1.818E-02* |
| Tumors_9 | Std | *0.000E+00* | 2.171E-01 | *0.000E+00* | 1.125E-01 | 9.223E-02 | 1.265E-01 | 3.953E-02 | 7.027E-02 |
| | Avg | *0.000E+00* | 4.168E-01 | *0.000E+00* | 5.000E-02 | 5.417E-02 | 4.000E-02 | 1.250E-02 | 3.333E-02 |
| Tumors_11 | Std | *1.757E-02* | 8.280E-02 | 3.749E-02 | 4.035E-02 | 3.412E-02 | 6.182E-02 | 5.476E-02 | 3.857E-02 |
| | Avg | *5.556E-03* | 5.556E-03 | 5.556E-03 | 5.556E-03 | 5.556E-03 | 5.556E-03 | 5.556E-03 | 5.556E-03 |
| Tumors_14 | Std | *5.974E-02* | 8.847E-02 | 7.256E-02 | 5.155E-02 | 6.288E-02 | 7.715E-02 | 6.560E-02 | 3.950E-02 |
| | Avg | *1.908E-01* | 3.376E-01 | 2.339E-01 | 2.781E-01 | 2.669E-01 | 3.063E-01 | 2.736E-01 | 2.787E-01 |
| ARV | | 2 | 8 | 3.5714 | 4.4286 | 3.7857 | 4.7143 | 2 | 1.7143 |
| Rank | | 2 | 8 | 4 | 6 | 5 | 7 | 2 | 1 |

the average number of selected features in Table 15, the proposed BGLSMA has the least average number of selected features in all datasets except Tumors_11 and Tumors_14. On Colon and Leukemia datasets, BGLSMA obtained a small selected feature with a standard deviation less than 1. This proves that GLSMA can obtain fewer features and higher classification accuracy. As can be seen from the tables, BGLSMA has higher classification accuracy in processing some complex high-dimensional data and can find smaller number of features and reduce the data scale. In terms of ARV index, BGLSMA ranks first. This suggests that BGLSMA can obtain very competitive results in terms of the number of features selected.

Table 16 shows the comparison results of eight algorithms in terms of average error rate. It can be seen from the ranking that BGLSMA has the lowest average error rate, which proves that the proposed algorithm not only has better performance in global optimization problems but also has good classification ability in feature selection optimization. It can be seen that the average error rate of BGLSMA is significantly lower than that of BSMA. The Gaussian mutation mechanism enables the population to search a larger space. As the number of iterations increases, the most representative features in each dataset are gradually selected, and the classification accuracy is also improved.

It is clear from the key measurements listed in Table 17, namely, the weighted number of features and the weighted error rate, that BGLSMA outperformed other competitors on 78.6% of the dataset. In addition, both the detailed data and the final ARV value show that BGLSMA has greatly improved compared with BSMA, which is due to the introduction of Levy flight mechanism, which increases the diversity and randomness of the population and selects features from a wider range of features, thus achieving higher classification accuracy.

Table 18 shows the average calculation time results of algorithm comparison. The computation cost of BGLSMA optimizer proposed in this paper is higher than that of BBA, BGSA, and other optimizers, and the time complexity

Table 17: Comparison between BGLSMA and other FS optimizers on average fitness.

| Datasets | Metrics | bGWO | BBA | BGSA | BPSO | bALO | BSSA | BSMA | BGLSMA |
|---|---|---|---|---|---|---|---|---|---|
| Colon | Std | 6.219E-02 | 1.416E-01 | 1.214E-01 | 1.179E-01 | 1.298E-01 | 1.177E-01 | 4.253E-04 | *1.318E-05* |
| | Avg | 3.354E-02 | 2.032E-01 | 2.024E-01 | 1.716E-01 | 1.752E-01 | 1.911E-01 | 8.000E-04 | *3.750E-05* |
| SRBCT | Std | 2.671E-04 | 3.431E-02 | 4.067E-04 | 4.700E-04 | 4.654E-04 | 7.300E-03 | 5.388E-04 | *1.343E-04* |
| | Avg | 4.043E-03 | 2.862E-02 | 1.965E-02 | 2.229E-02 | 2.163E-02 | 2.121E-02 | 7.474E-04 | *2.426E-04* |
| Leukemia | Std | 3.217E-04 | 2.016E-03 | 2.366E-04 | 1.555E-04 | 2.401E-04 | 1.124E-02 | 8.803E-05 | *3.621E-06* |
| | Avg | 5.562E-03 | 1.924E-02 | 2.188E-02 | 2.338E-02 | 2.293E-02 | 1.435E-02 | 2.756E-04 | *9.818E-06* |
| Brain_Tumor1 | Std | 4.935E-02 | 6.543E-02 | 6.796E-02 | 5.232E-02 | 6.844E-02 | 4.909E-02 | 3.995E-02 | *3.005E-02* |
| | Avg | 3.588E-02 | 8.786E-02 | 8.077E-02 | 7.302E-02 | 6.319E-02 | 6.895E-02 | 1.975E-02 | *9.558E-03* |
| Brain_Tumor2 | Std | 7.366E-02 | 1.106E-01 | 1.302E-01 | 8.625E-02 | 9.379E-02 | 9.388E-02 | 5.506E-04 | *2.472E-05* |
| | Avg | 4.058E-02 | 1.352E-01 | 9.527E-02 | 9.028E-02 | 7.990E-02 | 7.645E-02 | 7.345E-04 | *2.122E-05* |
| CNS | Std | 5.002E-02 | 2.023E-01 | 8.344E-02 | 1.670E-01 | 8.186E-02 | 9.031E-02 | 4.304E-02 | *1.035E-05* |
| | Avg | 2.189E-02 | 2.113E-01 | 1.016E-01 | 1.324E-01 | 7.409E-02 | 9.730E-02 | 1.497E-02 | *1.543E-05* |
| DLBCL | Std | 3.749E-02 | 4.246E-02 | 4.290E-02 | 4.277E-02 | 3.763E-02 | 1.010E-02 | 1.196E-04 | *9.685E-06* |
| | Avg | 1.706E-02 | 3.127E-02 | 3.513E-02 | 3.684E-02 | 3.469E-02 | 1.922E-02 | 2.615E-04 | *1.554E-05* |
| Leukemia1 | Std | 2.989E-04 | 2.989E-04 | 2.989E-04 | 2.989E-04 | 2.989E-04 | 2.989E-04 | 2.989E-04 | *2.989E-04* |
| | Avg | 5.301E-03 | 4.443E-02 | 2.144E-02 | 2.325E-02 | 2.284E-02 | 1.676E-02 | 6.645E-04 | *1.286E-04* |
| Leukemia2 | Std | 3.799E-04 | 3.799E-04 | 3.799E-04 | 3.799E-04 | 3.799E-04 | 3.799E-04 | 3.799E-04 | *3.799E-04* |
| | Avg | 5.584E-03 | 4.486E-02 | 2.246E-02 | 3.564E-02 | 2.352E-02 | 3.852E-02 | 3.399E-04 | *2.049E-05* |
| Lung_Cancer | Std | 2.298E-02 | 4.072E-02 | 2.363E-02 | 3.148E-02 | 2.459E-02 | 3.729E-02 | 2.851E-02 | *1.579E-02* |
| | Avg | 2.019E-02 | 6.019E-02 | 4.121E-02 | 4.178E-02 | 4.694E-02 | 3.486E-02 | 9.866E-03 | *5.136E-03* |
| Prostate_Tumor | Std | 5.479E-02 | 7.082E-02 | 9.162E-02 | 4.910E-02 | 6.320E-02 | 6.686E-02 | 6.357E-02 | *1.234E-05* |
| | Avg | 2.337E-02 | 2.337E-02 | 2.337E-02 | 2.337E-02 | 2.337E-02 | 2.337E-02 | 2.337E-02 | *2.337E-02* |
| Tumors_9 | Std | 3.326E-04 | 1.828E-01 | 9.096E-04 | 1.071E-01 | 8.783E-02 | 1.214E-01 | *3.912E-02* | 6.714E-02 |
| | Avg | 5.772E-03 | 1.681E-01 | 2.246E-02 | 7.109E-02 | 7.470E-02 | 6.044E-02 | *1.514E-02* | 3.192E-02 |
| Tumors_11 | Std | *1.651E-02* | 6.199E-02 | 3.563E-02 | 3.821E-02 | 3.251E-02 | 5.882E-02 | 5.134E-02 | 3.794E-02 |
| | Avg | *1.182E-02* | 1.030E-01 | 6.002E-02 | 6.892E-02 | 8.389E-02 | 8.022E-02 | 4.772E-02 | 4.175E-02 |
| Tumors_14 | Std | 5.678E-02 | 8.676E-02 | 6.881E-02 | 4.890E-02 | 5.967E-02 | 7.255E-02 | 6.031E-02 | *3.776E-02* |
| | Avg | 1.888E-01 | 1.888E-01 | 1.888E-01 | 1.888E-01 | 1.888E-01 | 1.888E-01 | 1.888E-01 | *1.888E-01* |
| ARV | | 2.5 | 7.5714 | 5.0714 | 6.3571 | 5.7143 | 5.0714 | 2.2143 | *1.5* |
| Rank | | 3 | 8 | 4 | 7 | 6 | 4 | 2 | *1* |

of the BSMA and bGWO with better performance is also higher than that of other optimizers, as shown in Tables 15–18. The introduction of GM and LF strategies not only improves the performance of BGLSMA but also increases the cost of computing time. Meanwhile, the time cost of the original SMA is higher than that of other algorithms, which leads to the high time cost of BGLSMA to a certain extent.

To sum up, BGLSMA is found to be the best optimizer in the overall comparison with other optimizers. Although the time cost is relatively high, BGLSMA can select the optimal feature subset on the vast majority of high-dimensional gene datasets without losing meaningful features and achieve the best fitness and classification error rate at the same time. The experimental results show that the combined strategy of Gaussian mutation and Levy flight guarantees the good results of GLSMA in global exploration.

## 5. Discussions

In this part, the GLSMA algorithm proposed in this paper, its advantages, and the points that can be improved are discussed. In the original SMA, the slime mould is not able to find the optimal solution in the search space, and it will fall into the local optimum when encountering some problems, which limits the use of the algorithm. In this paper, Gaussian mutation and Levy flight are introduced to update the population, which can enhance the global exploration ability and avoid the algorithm falling into local optimum. Experimental results show that the optimization effect of the dual mechanism is better than that of the single mechanism, and the optimization effect of GLSMA is better than some advanced optimization algorithms.

We monitor the situation after the population updates its optimal fitness value to determine whether it falls into local optimum. If it falls into local optimum, then Levy flight

Table 18: Comparison between BGLSMA and other FS optimizers on average computational time.

| Datasets | Metrics | bGWO | BBA | BGSA | BPSO | bALO | BSSA | BSMA | BGLSMA |
|---|---|---|---|---|---|---|---|---|---|
| Colon | Std | 2.5514 | 2.5514 | 2.5514 | 2.5514 | *2.5514* | 2.5514 | 2.5514 | 2.5514 |
| | Avg | 31.9062 | 30.4906 | 13.6746 | 8.584 | *8.0335* | 37.1659 | 70.1156 | 114.1263 |
| SRBCT | Std | 2.9138 | 2.9361 | 1.8018 | 0.91374 | *0.8866* | 3.2522 | 9.1156 | 16.0452 |
| | Avg | 34.0148 | 34.1434 | 17.2141 | 10.1891 | *9.9331* | 45.3289 | 79.6426 | 139.0788 |
| Leukemia | Std | 9.0416 | 8.7894 | 3.5903 | 1.6512 | *1.437* | 13.8392 | 27.5824 | 46.7772 |
| | Avg | 91.0773 | 86.08 | 41.7215 | 19.6962 | *18.1027* | 122.9642 | 235.5459 | 383.8539 |
| Brain_Tumor1 | Std | 7.8949 | 7.2781 | 4.1147 | 1.5711 | *2.0128* | 11.921 | 15.8751 | 35.8179 |
| | Avg | 77.839 | 75.6561 | 37.1795 | 20.7058 | *19.0823* | 106.5918 | 204.8332 | 331.4628 |
| Brain_Tumor2 | Std | 13.1514 | 10.441 | 6.5964 | 2.3475 | *1.8691* | 17.5793 | 39.2047 | 53.3446 |
| | Avg | 129.8032 | 116.4358 | 53.2582 | 20.6694 | *18.4634* | 170.6433 | 339.6091 | 519.5528 |
| CNS | Std | 8.7883 | 7.4052 | 4.2408 | 1.4574 | *1.5028* | 13.3078 | 28.6106 | 44.1968 |
| | Avg | 89.6331 | 86.575 | 38.5343 | 17.6038 | *15.9767* | 120.3373 | 235.2898 | 363.2358 |
| DLBCL | Std | 7.2489 | 6.5617 | 3.362 | 1.5989 | *1.342* | 10.6978 | 20.8082 | 28.7903 |
| | Avg | 71.175 | 67.8196 | 33.2668 | 16.5984 | *16.1274* | 95.6511 | 181.2594 | 299.2078 |
| Leukemia1 | Std | 7.4559 | 7.4559 | 7.4559 | 7.4559 | *7.4559* | 7.4559 | 7.4559 | 7.4559 |
| | Avg | 70.1512 | 65.732 | 31.9728 | 15.5488 | *14.9732* | 92.9905 | 179.7238 | 284.7675 |
| Leukemia2 | Std | 14.1082 | 13.8476 | 8.4149 | 3.1481 | *3.1518* | 20.2226 | 43.5071 | 64.6559 |
| | Avg | 139.8351 | 130.1774 | 62.3828 | 28.8214 | *26.3602* | 192.8016 | 369.2325 | 591.9679 |
| Lung_Cancer | Std | 20.2852 | 25.0081 | 17.671 | 9.2944 | *3.67* | 35.113 | 52.5769 | 136.7772 |
| | Avg | 186.6111 | 219.1699 | 141.8596 | 115.52 | *114.9638* | 281.4307 | 436.7065 | 893.1362 |
| Prostate_Tumor | Std | 14.2167 | 14.4316 | 6.592 | 4.633 | *3.9045* | 22.1621 | 26.0276 | 61.9174 |
| | Avg | 135.8685 | 130.3268 | 71.3971 | 37.1729 | *36.0317* | 187.0535 | 376.2129 | 565.8891 |
| Tumors_9 | Std | 7.598 | 5.6978 | 3.1465 | 1.17 | *1.4113* | 10.018 | 18.957 | 26.6486 |
| | Avg | 72.981 | 73.0313 | 31.4313 | 15.3876 | *14.656* | 98.7644 | 228.8745 | 291.5208 |
| Tumors_11 | Std | 19.6112 | 18.8848 | 19.1443 | 5.764 | *9.1023* | 37.9402 | 45.4794 | 98.5864 |
| | Avg | 179.912 | 193.265 | 121.7997 | 94.3676 | *90.4458* | 265.7033 | 519.6271 | 661.6608 |
| Tumors_14 | Std | 32.6068 | 44.7828 | 15.9668 | 14.1586 | *17.9125* | 72.9596 | 45.6897 | 114.9044 |
| | Avg | 278.9632 | 368.5713 | 323.2499 | 287.037 | *289.2344* | 458.6548 | 714.792 | 835.4772 |
| ARV | | 4.4286 | 4.3571 | 3.0714 | 2 | *1.1429* | 6 | 7 | 8 |
| Rank | | 5 | 4 | 3 | 2 | *1* | 6 | 7 | 8 |

mechanism is invoked to help the algorithm increase the search space and jump out of local trap. The combination of the dual mechanism is significantly better than the single mechanism. However, it can be seen from Table 18 that the time cost of BGLSMA is relatively high, which is partly due to the high time cost of BSMA and partly due to the addition of mechanism, which leads to the increase of time cost. Correspondingly, the mechanism greatly improves the performance of the algorithm, allowing it to be applied to more domains, such as human activity recognition [74], microgrid planning [75], medical image augmentation [76], autism spectrum disorder classification [77], disease prediction [78, 79], named entity recognition [80], information retrieval services [81–83], and recommender systems [84–87].

## 6. Conclusions and Future Directions

In this paper, an improved SMA (GLSMA) algorithm based on Gaussian mutation and Levy flight is proposed. Experimental results show that the two mechanisms play an important role in further enhancing the global search of SMA and alleviating falling into local optimum. Firstly, the effectiveness of GLSMA method is verified by comparison with DE, PSO, GWO, and other well-known algorithms. Secondly, compared with other advanced swarm intelligence algorithms, such as MPEDE, LSHADE, ALCPSO, and CLPSO, GLSMA is able to find the optimal solution faster. Finally, in order to prove the performance of GLSMA in practical applications, BGLSMA is obtained by mapping GLSMA into binary space through transformation function, and it is applied to feature selection problems of 14 commonly used UCI high-dimensional gene datasets. Compared with excellent metaheuristic optimizer, general average characteristics selected number, average error rate, and average fitness and calculated the cost four aspects; it can be seen that GLSMA in the application of feature selection still has good global search ability and be able to select fewer features and higher classification accuracy. Therefore, the above conclusions indicate that GLSMA can be a promising method for not only function optimization problems but also practical feature selection problems.

There are still many aspects to explore in our research. We can consider applying GLSMA to other feature selection datasets and study the effectiveness of BGLSMA on other datasets. Further improvements to the SMA can be attempted to improve the balance between global exploration and local development. Finally, it is an interesting topic to apply SMA to more fields, such as photovoltaic parameter optimization and image segmentation (see Tables 5–18).

## Data Availability

The data involved in this study are all public data, which can be downloaded through public channels.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References


[1] X. Zhang, D. Wang, Z. Zhou, and Y. Ma, "Robust low-rank tensor recovery with rectification and alignment," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 238–255, 2021.

[2] W. Deng, H. Zhao, L. Zou, G. Li, X. Yang, and D. Wu, "A novel collaborative optimization algorithm in solving complex optimization problems," *Soft Computing*, vol. 21, no. 15, pp. 4387–4398, 2017.

[3] R. Storn, "Designing nonstandard filters with differential evolution," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 103–106, 2005.

[4] J. Kennedy and R. Mendes, "Neighborhood topologies in fully informed and best-of-neighborhood particle swarms," *Ieee Transactions on Systems Man and Cybernetics Part C-Applications and Reviews*, vol. 36, no. 4, pp. 515–519, 2006.

[5] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: algorithm and applications," *Future Generation Computer Systems-the International Journal of Escience*, vol. 97, pp. 849–872, 2019.

[6] I. Ahmadianfar, A. A. Heidari, A. H. Gandomi, X. Chu, and H. Chen, "RUN beyond the metaphor: an efficient optimization algorithm based on Runge Kutta method," *Expert Systems with Applications*, vol. 181, p. 115079, 2021.

[7] Y. Yang, H. Chen, A. A. Heidari, and A. H. Gandomi, "Hunger games search: visions, conception, implementation, deep analysis, perspectives, and towards performance shifts," *Expert Systems with Applications*, vol. 177, p. 114864, 2021.

[8] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, "Slime mould algorithm: a new method for stochastic optimization," *Future Generation Computer Systems-the International Journal of Escience*, vol. 111, pp. 300–323, 2020.

[9] G. G. Wang, S. Deb, and Z. H. Cui, "Monarch butterfly optimization," *Neural Computing & Applications*, vol. 31, no. 7, pp. 1995–2014, 2019.

[10] G. G. Wang, "Moth search algorithm: a bio-inspired metaheuristic algorithm for global optimization problems," *Memetic Computing*, vol. 10, no. 2, pp. 151–164, 2018.

[11] J. Tu, H. Chen, M. Wang, and A. H. Gandomi, "The colony predation algorithm," *Journal of Bionic Engineering*, vol. 18, no. 3, pp. 674–710, 2021.

[12] I. Ahmadianfar, A. A. Heidari, S. Noshadian, H. Chen, and A. H. Gandomi, "INFO: an efficient optimization algorithm based on weighted mean of vectors," *Expert Systems with Applications*, vol. 195, p. 116516, 2022.

[13] X. Ye, W. Liu, H. Li et al., "Modified whale optimization algorithm for solar cell and PV module parameter identification," *Complexity*, vol. 2021, Article ID 8878686, 23 pages, 2021.

[14] R. Dong, H. Chen, A. A. Heidari, H. Turabieh, M. Mafarja, and S. Wang, "Boosted kernel search: framework, analysis and case studies on the economic emission dispatch problem," *Knowledge-Based Systems*, vol. 233, p. 107529, 2021.

[15] A. G. Hussien, A. A. Heidari, X. Ye, G. Liang, H. Chen, and Z. Pan, "Boosting whale optimization with evolution strategy and Gaussian random walks: an image segmentation method," *Engineering with Computers*, 2022.

[16] H. Yu, J. Song, C. Chen et al., "Image segmentation of leaf spot diseases on maize using multi-stage Cauchy- enabled grey wolf algorithm," *Engineering Applications of Artificial Intelligence*, vol. 109, p. 104653, 2022.

[17] H. Yu, X. Cheng, C. Chen et al., "Apple leaf disease recognition method with improved residual network," *Multimedia Tools and Applications*, vol. 81, no. 6, pp. 7759–7782, 2022.

[18] J. Xia, Z. Wang, D. Yang et al., "Performance optimization of support vector machine with oppositional grasshopper optimization for acute appendicitis diagnosis," *Computers in Biology and Medicine*, vol. 143, p. 105206, 2022.

[19] J. Xia, D. Yang, H. Zhou et al., "Evolving kernel extreme learning machine for medical diagnosis via a disperse foraging sine cosine algorithm," *Computers in Biology and Medicine*, vol. 141, p. 105137, 2022.

[20] X. Han, Y. Han, Q. Chen et al., "Distributed Flow shop scheduling with sequence-dependent setup times using an improved iterated greedy algorithm," *Complex System Modeling and Simulation*, vol. 1, no. 3, pp. 198–217, 2021.

[21] D. Gao, G.-G. Wang, and W. Pedrycz, "Solving fuzzy job-shop scheduling problem using DE algorithm improved by a selection mechanism," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 12, pp. 3265–3275, 2020.

[22] G.-G. Wang, D. Gao, and W. Pedrycz, "Solving multi-objective fuzzy job-shop scheduling problem by a hybrid adaptive differential evolution algorithm," *IEEE Transactions on Industrial Informatics*, 2022.

[23] H. ling Chen, B. Yang, S. Jing Wang, G. Wang, H. Zhong Li, and W. Bin Liu, "Towards an optimal support vector machine classifier using a parallel particle swarm optimization strategy," *Applied Mathematics and Computation*, vol. 239, pp. 180–197, 2014.

[24] W. Deng, X. Zhang, Y. Zhou et al., "An enhanced fast non-dominated solution sorting genetic algorithm for multi- objective problems," *Information Sciences*, vol. 585, pp. 441–453, 2022.

[25] Y. Hua, Q. Liu, K. Hao, and Y. Jin, "A survey of evolutionary algorithms for multi-objective optimization problems with irregular Pareto fronts," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 2, pp. 303–318, 2021.

[26] H. Yu, K. Yuan, W. Li et al., "Improved butterfly optimizer-configured extreme learning machine for fault diagnosis," *Complexity*, vol. 2021, Article ID 6315010, 17 pages, 2021.

[27] K. Hu, J. Ye, E. Fan, S. Shen, L. Huang, and J. Pi, "A novel object tracking algorithm by fusing color and depth information based on single valued neutrosophic cross-entropy," *Journal of Intelligent Fuzzy Systems*, vol. 32, no. 3, pp. 1775–1786, 2017.

[28] K. Hu, W. He, J. Ye, L. Zhao, H. Peng, and J. Pi, "Online visual tracking of weighted multiple instance learning via neutrosophic similarity-based objectness estimation," *Symmetry*, vol. 11, no. 6, p. 832, 2019.

[29] S.-H. Wu, Z.-H. Zhan, and J. Zhang, "SAFE: scale-adaptive fitness evaluation method for expensive optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 3, pp. 478–491, 2021.

[30] J.-Y. Li, Z. H. Zhan, C. Wang, H. Jin, and J. Zhang, "Boosting data-driven evolutionary algorithm with localized data generation," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 5, pp. 923–937, 2020.

[31] Q. Li, H. Chen, H. Huang et al., "An enhanced grey wolf optimization based feature selection wrapped kernel extreme learning machine for medical diagnosis," *Computational and Mathematical Methods in Medicine*, vol. 2017, Article ID 9512741, 15 pages, 2017.

[32] Z. Cai, J. Gu, C. Wen et al., "An intelligent Parkinson's disease diagnostic system based on a chaotic bacterial foraging optimization enhanced fuzzy KNN approach," *Computational and Mathematical Methods in Medicine*, vol. 2018, Article ID 2396952, 24 pages, 2018.

[33] F. Zhao, S. di, J. Cao, J. Tang, and Jonrinaldi, "A novel cooperative multi-stage hyper-heuristic for combination optimization problems," *Complex System Modeling and Simulation*, vol. 1, no. 2, pp. 91–108, 2021.

[34] J. Hu, H. Chen, A. A. Heidari et al., "Orthogonal learning covariance matrix for defects of grey wolf optimizer: insights, balance, diversity, and feature selection," *Knowledge-Based Systems*, vol. 213, p. 106684, 2021.

[35] J. Hu, W. Gui, A. A. Heidari et al., "Dispersed foraging slime mould algorithm: continuous and binary variants for global optimization and wrapper-based feature selection," *Knowledge-Based Systems*, vol. 237, p. 107761, 2022.

[36] H. Chen, M. Wang, and X. Zhao, "A multi-strategy enhanced sine cosine algorithm for global optimization and constrained practical engineering problems," *Applied Mathematics and Computation*, vol. 369, p. 124872, 2020.

[37] H. Yu, S. Qiao, A. A. Heidari, C. Bi, and H. Chen, "Individual disturbance and attraction repulsion strategy enhanced seagull optimization for engineering design," *Mathematics*, vol. 10, no. 2, p. 276, 2022.

[38] Z. He, G. G. Yen, and J. Ding, "Knee-based decision making and visualization in many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 2, pp. 292–306, 2021.

[39] Z. He, G. G. Yen, and J. Lv, "Evolutionary multiobjective optimization with robustness enhancement," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 3, pp. 494–507, 2020.

[40] R. Kouadri, L. Slimani, and T. Bouktir, "Slime mould algorithm for practical optimal power flow solutions incorporating stochastic wind power and static var compensator device," *Electrical Engineering & Electromechanics*, vol. 6, pp. 45–54, 2020.

[41] S. Khunkitti, A. Siritaratiwat, and S. Premrudeepreechacharn, "Multi-objective optimal power Flow problems based on slime mould algorithm," *Sustainability*, vol. 13, no. 13, p. 7448, 2021.

[42] J. Jafari-Asl, S. Ohadi, M. E. Ben Seghier, and N. T. Trung, "Accurate structural reliability analysis using an improved line-sampling-method-based slime mold algorithm," *Asce-Asme Journal of Risk and Uncertainty in Engineering Systems Part a-Civil Engineering*, vol. 7, no. 2, 2021.

[43] D. Izci and S. Ekinci, "Comparative performance analysis of slime mould algorithm for efficient design of proportional-integral-derivative controller," *Electrica*, vol. 21, no. 1, pp. 151–159, 2021.

[44] E. H. Houssein, M. A. Mahdy, D. Shebl, A. Manzoor, R. Sarkar, and W. M. Mohamed, "An efficient slime mould algorithm for solving multi-objective optimization problems," *Expert Systems with Applications*, vol. 187, p. 115870, 2022.

[45] E. H. Houssein, M. A. Mahdy, M. J. Blondin, D. Shebl, and W. M. Mohamed, "Hybrid slime mould algorithm with adaptive guided differential evolution algorithm for combinatorial and global optimization problems," *Expert Systems with Applications*, vol. 174, p. 114689, 2021.

[46] J. Gupta, P. Nijhawan, and S. Ganguli, "Optimal parameter estimation of PEM fuel cell using slime mould algorithm," *International Journal of Energy Research*, vol. 45, no. 10, pp. 14732–14744, 2021.

[47] S. K. Elsayed, A. M. Agwa, M. A. El-Dabbah, and E. E. Elattar, "Slime mold optimizer for transformer parameters identification with experimental validation," *Intelligent Automation and Soft Computing*, vol. 28, no. 3, pp. 639–651, 2021.

[48] M. H. Hassan, S. Kamel, L. Abualigah, and A. Eid, "Development and application of slime mould algorithm for optimal economic emission dispatch," *Expert Systems with Applications*, vol. 182, p. 115205, 2021.

[49] H. Jia, W. Zhang, R. Zheng, S. Wang, X. Leng, and N. Cao, "Ensemble mutation slime mould algorithm with restart mechanism for feature selection," *International Journal of Intelligent Systems*, vol. 37, no. 3, pp. 2335–2370, 2022.

[50] O. Altay, "Chaotic slime mould optimization algorithm for global optimization," *Artificial Intelligence Review*, vol. 55, no. 5, pp. 3979–4040, 2022.

[51] M. Abdel-Basset, V. Chang, and R. Mohamed, "HSMA_WOA: a hybrid novel slime mould algorithm with whale optimization algorithm for tackling the image segmentation problem of chest X-ray images," *Applied Soft Computing*, vol. 95, p. 106642, 2020.

[52] S. Chauhan, G. Vashishtha, and A. Kumar, "A symbiosis of arithmetic optimizer with slime mould algorithm for improving global optimization and conventional design problem," *Journal of Supercomputing*, vol. 78, no. 5, pp. 6234–6274, 2022.

[53] J. Alcalá-Fdez, L. Sanchez, S. Garcia et al., "KEEL: a software tool to assess evolutionary algorithms for data mining problems," *Soft Computing*, vol. 13, no. 3, pp. 307–318, 2009.

[54] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.

[55] S. Mirjalili, S. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.

[56] S. Mirjalili, "SCA: a sine cosine algorithm for solving optimization problems," *Knowledge-Based Systems*, vol. 96, pp. 120–133, 2016.

[57] W.-T. Pan, "A new fruit fly optimization algorithm: taking the financial distress model as an example," *Knowledge-Based Systems*, vol. 26, pp. 69–74, 2012.

[58] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, and S. M. Mirjalili, "Salp swarm algorithm: a bio-inspired optimizer for engineering design problems," *Advances in Engineering Software*, vol. 114, pp. 163–191, 2017.

[59] G. Wu, R. Mallipeddi, P. N. Suganthan, R. Wang, and H. Chen, "Differential evolution with multi-population based ensemble of mutation strategies," *Information Sciences*, vol. 329, pp. 329–345, 2016.

[60] A. Piotrowski, "L-SHADE optimization algorithms with population-wide inertia," *Information Sciences*, vol. 468, pp. 117–141, 2018.

[61] W. Chen, J. Zhang, Y. Lin et al., "Particle swarm optimization with an aging leader and challengers," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 2, pp. 241–258, 2013.

[62] J. Liang, A. K. Qin, P. N. Suganthan, and S. Baskar, "Comprehensive learning particle swarm optimizer for global optimization of multimodal functions," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 3, pp. 281–295, 2006.

[63] N. Hansen, "The CMA Evolution Strategy: A Comparing Review," in *Towards a New Evolutionary Computation: Advances in the Estimation of Distribution Algorithms*, J. A. Lozan, P. Larrañaga, I. Inza, and E. Bengoetxea, Eds., pp. 75–102, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[64] A. Heidari, I. Aljarah, H. Faris, H. Chen, J. Luo, and S. Mirjalili, "An enhanced associative learning-based exploratory whale optimizer for global optimization," *Neural Computing & Applications*, vol. 32, no. 9, pp. 5185–5211, 2020.

[65] A. Lin, Q. Wu, A. A. Heidari et al., "Predicting intentions of students for master programs using a chaos-induced sine cosine-based fuzzy K-nearest neighbor classifier," *Ieee Access*, vol. 7, pp. 67235–67248, 2019.

[66] Z. Cai, J. Gu, J. Luo et al., "Evolving an optimal kernel extreme learning machine by using an enhanced grey wolf optimization strategy," *Expert Systems with Applications*, vol. 138, p. 112814, 2019.

[67] S. Jadhav, H. He, and K. Jenkins, "Information gain directed genetic algorithm wrapper feature selection for credit rating," *Applied Soft Computing*, vol. 69, pp. 541–553, 2018.

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

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

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

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

[72] E. Emary, H. Zawbaa, and A. Hassanien, "Binary ant lion approaches for feature selection," *Neurocomputing*, vol. 213, pp. 54–65, 2016.

[73] K. Reddy, L. K. Panwar, B. K. Panigrahi, and R. Kumar, "A new binary variant of sine-cosine algorithm: development and application to solve profit-based unit commitment problem," *Arabian Journal for Science and Engineering*, vol. 43, no. 8, pp. 4041–4056, 2018.

[74] S. Qiu, H. Zhao, N. Jiang et al., "Multi-sensor information fusion based on machine learning for real applications in human activity recognition: state-of-the-art and research challenges," *Information Fusion*, vol. 80, pp. 241–265, 2022.

[75] X. Cao, X. Sun, Z. Xu, B. Zeng, and X. Guan, "Hydrogen-based networked microgrids planning through two-stage stochastic programming with mixed-integer conic recourse," *IEEE Transactions on Automation Science and Engineering*, pp. 1–14, 2021.

[76] Q. Guan, Y. Chen, Z. Wei et al., "Medical image augmentation for lesion detection using a texture-constrained multichannel progressive GAN," *Computers in Biology and Medicine*, vol. 145, p. 105444, 2022.

[77] Z. Hu, J. Wang, C. Zhang et al., "Uncertainty modeling for multicenter autism spectrum disorder classification using Takagi–Sugeno–Kang fuzzy systems," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 2, pp. 730–739, 2022.

[78] Y. Su, S. Li, C. Zheng, and X. Zhang, "A heuristic algorithm for identifying molecular signatures in cancer," *IEEE Transactions on Nanobioscience*, vol. 19, no. 1, pp. 132–141, 2020.

[79] L. Li, Z. Gao, Y. T. Wang et al., "SCMFMDA: predicting microRNA-disease associations based on similarity constrained matrix factorization," *PLoS Computational Biology*, vol. 17, no. 7, article e1009165, 2021.

[80] Z. Yang, J. Ma, H. Chen, J. Zhang, and Y. Chang, "Context-aware attentive multi-level feature fusion for named entity recognition," *IEEE Transactions on Neural Networks and Learning Systems*, vol. PP, pp. 1–12, 2022.

[81] Z. Wu, R. Li, Z. Zhou, J. Guo, J. Jiang, and X. Su, "A user sensitive subject protection approach for book search service," *Journal of the Association for Information Science and Technology*, vol. 71, no. 2, pp. 183–195, 2020.

[82] Z. Wu, S. Shen, X. Lian, X. Su, and E. Chen, "A dummy-based user privacy protection approach for text information retrieval," *Knowledge-Based Systems*, vol. 195, p. 105679, 2020.

[83] Z. Wu, S. Shen, H. Zhou, H. Li, C. Lu, and D. Zou, "An effective approach for the protection of user commodity viewing privacy in e-commerce website," *Knowledge-Based Systems*, vol. 220, p. 106952, 2021.

[84] D. Wang, Y. Liang, D. Xu, X. Feng, and R. Guan, "A content-based recommender system for computer science publications," *Knowledge-Based Systems*, vol. 157, pp. 1–9, 2018.

[85] J. Li, C. Chen, H. Chen, and C. Tong, "Towards context-aware social recommendation via individual trust," *Knowledge-Based Systems*, vol. 127, pp. 58–66, 2017.

[86] J. Li and J. Lin, "A probability distribution detection based hybrid ensemble QoS prediction approach," *Information Sciences*, vol. 519, pp. 289–305, 2020.

[87] J. Li, X. L. Zheng, S. T. Chen, W. W. Song, and D. R. Chen, "An efficient and reliable approach for quality-of-service-aware service composition," *Information Sciences*, vol. 269, pp. 238–254, 2014.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 6899448, 9 pages
https://doi.org/10.1155/2022/6899448

*Research Article*

# Primary Investigation of Deep Learning Models for Japanese "Group Classification" of Whole-Slide Images of Gastric Endoscopic Biopsy

**Xiaoya Fan**,[1] **Lihui Yu**,[2] **Xin Qi**,[2] **Xue Lin**,[2] **Junjun Zhao**,[2] **Dong Wang**,[2] **and Jing Zhang**[2]

[1]*School of Software, Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, Dalian University of Technology, Dalian, Liaoning, China*
[2]*Dalian Municipal Central Hospital, Dalian, Liaoning, China*

Correspondence should be addressed to Dong Wang; wangdong0717@126.com and Jing Zhang; zj84402001@163.com

Received 28 June 2022; Accepted 8 August 2022; Published 26 September 2022

Academic Editor: Lu Liu



*Background*. Accurate pathological diagnosis of gastric endoscopic biopsy could greatly improve the opportunity of early diagnosis and treatment of gastric cancer. The Japanese "Group classification" of gastric biopsy corresponds well with the endoscopic diagnostic system and can guide clinical treatment. However, severe shortage of pathologists and their heavy workload limit the diagnostic accuracy. This study presents the first attempt to investigate the applicability and effectiveness of AI-aided system for automated Japanese "Group classification" of gastric endoscopic biopsy. *Methods*. In total, 260 whole-slide images of gastric endoscopic biopsy were collected from Dalian Municipal Central Hospital from January 2015 to January 2021. These images were annotated by experienced pathologists according to the Japanese "Group classification." Five popular convolutional neural networks, i.e., VGG16, VGG19, ResNet50, Xception, and InceptionV3 were trained and tested. The performance of the models was compared in terms of widely used metrics, namely, AUC (area under the receiver operating characteristic curve, i.e., ROC curve), accuracy, recall, precision, and F1 score. *Results*. Results showed that ResNet50 achieved the best performance with accuracy 93.16% and AUC 0.994. *Conclusion*. Our results demonstrated the applicability and effectiveness of DL-based system for automated Japanese "Group classification" of gastric endoscopic biopsy.

## 1. Background

Gastric cancer has long been acknowledged as a severe public health problem across the globe [1] with a 5-year survival rate of lower than 40% [2]. Despite the decrease of mortality over the past few decades in some countries, it remains the fourth leading cause of cancer death worldwide [3] and third in China [4]. In China, it is still highly prevalent and accounts for over 40% of new cases in the world [3, 4]. Most gastric cancer cases are diagnosed at an advanced stage due to its atypical symptoms in early stage and late aggressive behaviors [5]. The situation is more severe in China, where more than 60% of patients were diagnosed at an advanced stage [6]. However, the treatment options are often limited at this stage, resulting in unsatisfactory prognosis [7]. Early diagnosis of gastric cancer enables early clinical intervention, thus improving prognosis and survival rate. Currently, timely and correct diagnosis of gastric cancer relies heavily on pathological examination of gastric biopsy tissue, which is performed by highly trained pathologists with an optical microscopy. However, this process is tedious and time-consuming. Moreover, the accuracy of pathological diagnosis of gastric biopsy is quite limited [4] due to the shortage of pathologists worldwide. Such shortage leads to heavy workload for pathologists and possible errors in diagnosis.

Therefore, there is a great need for automated and accurate pathology diagnosis of gastric cancer. Several applications of deep learning (DL) models have emerged in digital pathology image analysis. The common task is either binary classification task as tumor detection or three-way

classification task as cancer classification. For instance, Wang et al. [8] proposed GastricNet, a DL-based framework for automatic gastric cancer identification. The model achieved 100% accuracy for slice-based classification, outperforming the state-of-the-art networks like DenseNet and ResNet. Leon et al. [9] proposed two independent approaches based on convolutional neural network (CNN) for gastric cancer detection using histopathological image samples. The first one analyzed the morphological features of the whole image. The second one analyzed the local characteristic properties. The histopathological image was classified as either benign or malignant. Qu et al. [10] performed the same task with a transfer learning strategy. The model was first pretrained with ImageNet and further fine-tuned with a well-annotated benign/malignant dataset. A detection accuracy of up to 89.72% was achieved, showing the promise of automated gastric cancer detection. On the other hand, Sharma et al. [11] explored the deep learning methods for classification in H&E-stained histopathological whole slide images (WSIs) of gastric carcinoma. The WSIs were classified as HER2+ tumor, HER2- tumor, or nontumor. The model achieved an overall accuracy of 0.6690. More recently, Iizuka et al. [12] investigated the feasibility of CNNs and recurrent neural networks (RNNs) for classifying WSI into adenocarcinoma, adenoma, and nonneoplastic, achieving area under the curves (AUCs) up to 0.97 and 0.99 for gastric adenocarcinoma and adenoma, respectively. The work described above suggests that DL models are promising for pathological image analysis for gastric cancer. However, clinical application remains challenging. The ultimate purpose of automated pathological image analysis is to better serve for decision making in treatment. However, this work does not immediately translate to clinic decisions.

In China, the majority of pathologists are trained with *WHO classification* [13] while most treatment options are adopted from Japan [4, 5] with reference to the *Japanese Gastric Cancer Treatment Guidelines* [14] (Japanese guidelines). However, the pathological classification adopted in Japanese guidelines is the "group classification" of gastric biopsy specimens (Groups 1~5), rather than WHO classification. Thereby, the pathologists and gastroenterologists speak different languages, and the pathological diagnosis does not correspond well with clinical treatment decisions.

In this paper, the feasibility of deep learning models for automated Japanese "Group classification" of WSIs of gastric endoscopic biopsy was investigated. Five popular DL models, VGG-16 [15], VGG-19 [15], ResNet-50 [16], Xception [17], and InceptionV3 [18], were trained and compared. Results showed that ResNet50 achieved the best performance with an accuracy of 93.16% and an AUC of 0.994. To the best of our knowledge, this is the first attempt to investigate the applicability and effectiveness of AI-aided system for pathological group classification of human gastric epithelial lesions.

## 2. Methods

*2.1. Whole-Slide Image Preparation.* In total, 260 cases of gastric endoscopic biopsy from 173 patients (128 males and 45 females, aged from 27 to 92 years old, mean ± std: 65.2 ± 11.4) with human gastric epithelial lesions were collected from Dalian Municipal Central Hospital from January 2015 to January 2021. The WSIs were stained with hematoxylin and eosin (H&E) and further produced at ×40 magnification (0.238 $\mu$m/pixel) by the digital scanner (KF-PRO-005). All procedures performed in studies involving human participants were approved by the Medical Ethics Committee of Dalian Municipal Central Hospital and in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed written consent was also obtained from individual participants included in the study. The workflow of the study is illustrated in Figure 1.

*2.2. Annotation Procedure.* All the WSIs were annotated by two experienced pathologists using an open-source annotation tool *labelme* (https://github.com/wkentaro/labelme). According to the Japanese "Group classification," gastric endoscopic biopsy can be classified as 5 groups (see Figure 2 for illustration). The description of the 5 groups is shown in Table 1. Briefly, it defines Group 1 as normal tissue or nonneoplastic lesion tissue, Group 2 as tissue that is difficult to make diagnosis between neoplastic and nonneoplastic lesions, Group 3 as adenoma, Group 4 as tissue with neoplastic lesion that is suspected to be carcinoma, and Group 5 as tissue with carcinoma [14].

The annotation tool, *labelme*, enables the pathologists to segment a WSI into various regions and label each region with the group they belong to. It has to be noted that different regions from a single WSI could be labeled as different groups, whereas a single region can only be labeled as a single group. The annotation process was accomplished by two pathologists. Pathologist 1 first drew the outline of each region and annotated with Groups 1~5. The initial annotation was then modified, confirmed, or verified by a senior pathologist. An example of WSI with final annotation is illustrated in Figure 3. The region outlined in red was annotated as Group 5, while the region outlined in green was annotated as Group 1. The remaining regions that were not outlined were the regions that are hard to classify.

*2.3. Preprocessing and Datasets.* A big challenge faced in computational pathology is the huge size of a WSI. A single image can contain hundreds of millions of pixels. To apply deep learning models, a WSI image was first segmented into small tiles with $400 \times 400$ pixel size. The tiles with a tissue area less than 50% were discarded. In clinical practice, in order to increase diagnostic accuracy, pathologists often observe the specimens under various magnifications of the view (40×, 100×, 200×, and 400×). Therefore, we varied the size of the tiles from $400 \times 400$ pixel, $600 \times 600$ pixel, $800 \times 800$ pixel to $1000 \times 1000$ pixel and established 5 datasets consisting of tiles with different sizes (see Table 2). All datasets underwent the tile selection process described above.

*2.4. Model Training, Testing, and Evaluation.* Each dataset was split into a training set (60%), validation set (20%),

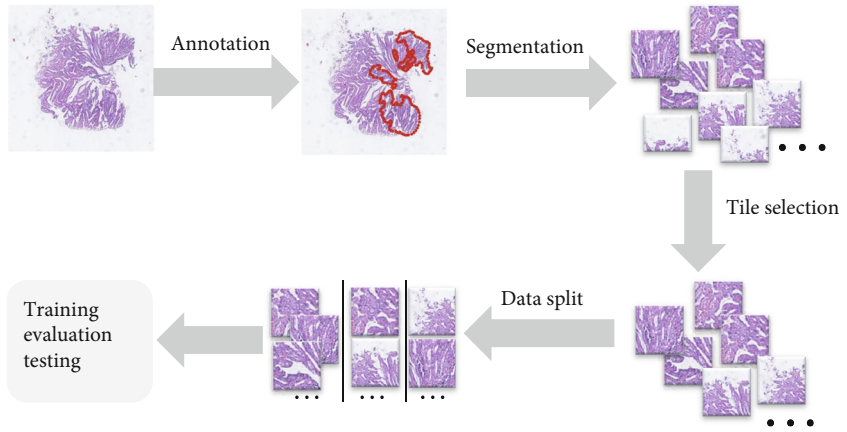



Figure 1: Workflow of the study.

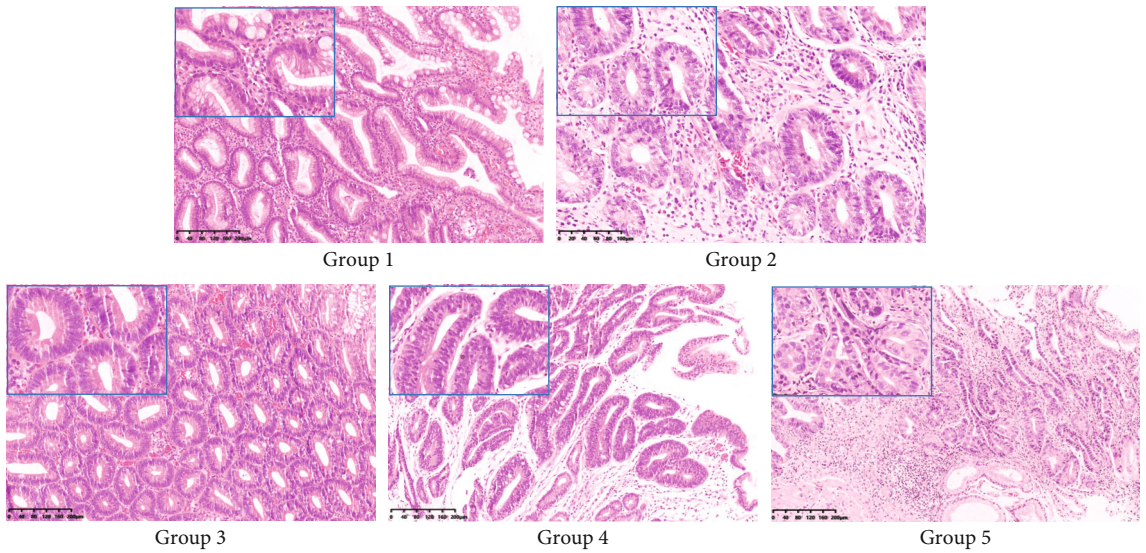



Figure 2: Typical examples of Groups 1~5 according to the Japanese "Group classification." Framed in the left corner of each subfigure is the typical appearance of the corresponding class.

Table 1: Description of Japanese "Group classification".

| Group | Description |
|---|---|
| Group 1 | Normal tissue or nonneoplastic lesion tissue |
| Group 2 | Difficult to make diagnosis between neoplastic and nonneoplastic lesions |
| Group 3 | Adenoma |
| Group 4 | Can be diagnosed as neoplastic lesion and is suspected to be carcinoma |
| Group 5 | Carcinoma |

and testing set (20%). To avoid data imbalance, the splitting process was done within each group. Five popular models were trained, i.e., VGG16, VGG19, ResNet50, Xception, and InceptionV3. Standard data augmentation techniques (such as reflection, rotation, and shift) and early stopping were employed to avoid overfitting. TensorFlow was used as the framework to build DL models. All models were trained/tested on one Nvidia GeForce RTX 2080Ti 8 GB GPU.

Commonly used metrics, namely, overall accuracy (Acc) and area under the receiver operator characteristic (ROC) curve (AUC), were used to evaluate the performance of the models, which were calculated from comparing model prediction with the annotation of pathologists. The accuracy, recall, precision, and F1 score were also calculated for each group to provide detailed information of model performance.

Acc represents the overall accuracy, which was defined as the percentage of correctly predicted tiles in all tiles.

The recall for Group $i$ was defined as

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{1}$$

where TP (true positive) is the number of tiles that were annotated as Group $i$ by pathologists and correctly predicted

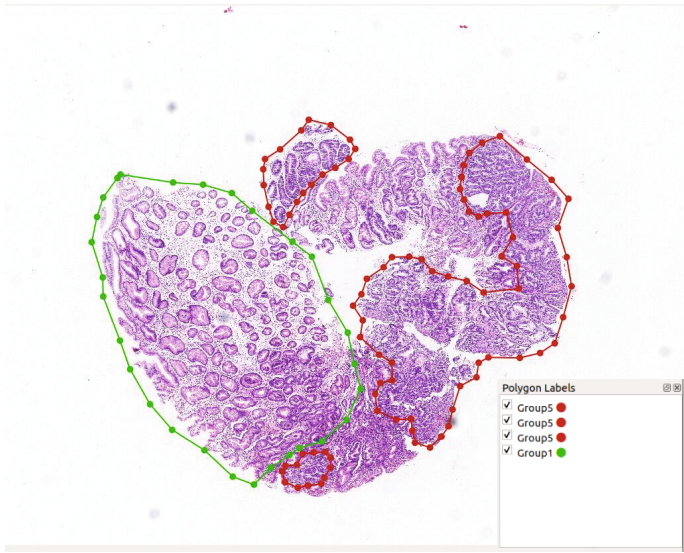



Figure 3: An example of annotated WSI, segmented with *labelme*. Outlined in red was annotated as Group 5 while in green was annotated as Group 1.

Table 2: Information of datasets.

| | Group 1 | Group 2 | Group 3 | Group 4 | Group 5 |
|---|---|---|---|---|---|
| Number of WSIs | 166 | 84 | 45 | 68 | 58 |
| Dataset 1 (400 × 400) | 45001 | 10274 | 17999 | 20845 | 41521 |
| Dataset 2 (600 × 600) | 21430 | 4307 | 7584 | 8967 | 17834 |
| Dataset 3 (800 × 800) | 10875 | 2291 | 4204 | 4931 | 9944 |
| Dataset 4 (1000 × 1000) | 7243 | 1387 | 2652 | 3044 | 6318 |
| Dataset 5 (mixed size*) | 39548 | 7985 | 14440 | 16942 | 34096 |

*Tiles with various sizes (600 × 600, 800 × 800, and 1000 × 1000) were mixed to construct Dataset 5.

as Group $i$ by the model; FN (false negative) is the number of tiles that were annotated as Group $i$ but predicted incorrectly as any other groups by the model. It represents the percentage of tiles that were correctly predicted by the model in all tiles annotated as Group $i$ by pathologists.

The precision for Group $i$ was defined as

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{2}$$

where FP (false positive) is the number of tiles that were predicted as Group $i$ but annotated as any other groups by pathologists. It stands for the percentage of tiles that were annotated as Group $i$ by pathologists in all tiles that were predicted as Group $i$ by the model.

The F1 score, a metric that combines both precision and recall, was defined as

$$\text{F1 score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{Precision} + \text{recall}}. \tag{3}$$

Table 3: Test accuracy of five different models*.

| Model | AUC | Acc (%) |
|---|---|---|
| ResNet50 | 0.988 | 89.5 |
| VGG16 | 0.970 | 83.1 |
| VGG19 | 0.949 | 76.6 |
| Xception | 0.894 | 66.2 |
| InceptionV3 | 0.881 | 63.0 |

*The models were trained on Dataset 1.

## 3. Results

*3.1. Among the Five Model Architectures, ResNet50 Performs the Best.* We first compared the performance of five different models trained with Dataset 1 (see Table 3). Among the five models, ResNet50 achieved the best performance with an AUC of 0.988 and an Acc of 89.5%, followed by VGG16 and VGG19 with an AUC of 0.970 and 0.949, respectively. We further analyzed the accuracy, recall, precision, and F1 score of the five models for each group individually (see Figure 4). Results showed that ResNet50 achieved 90.33%, 81.56%, 88.33%, 81.15%, and 95.18% recall for Groups

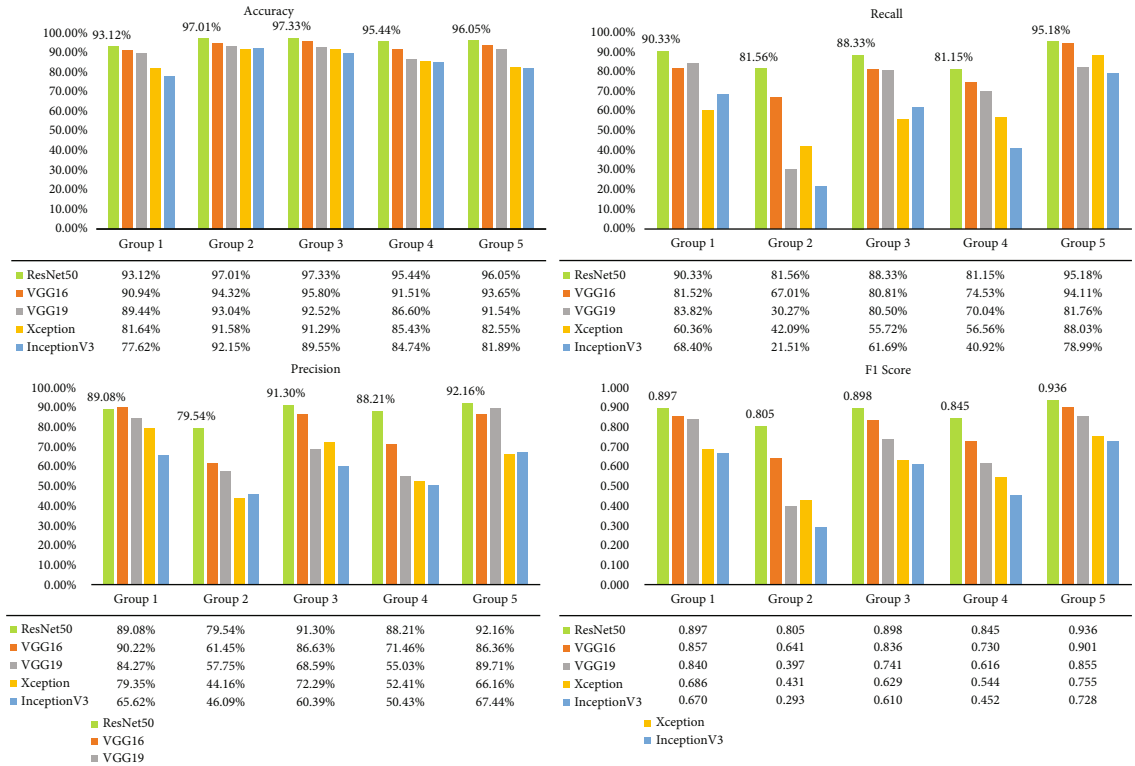



Figure 4: Performance (accuracy, recall, precision, and F1 score) of five different models for each group evaluated on the test set. All models were trained on Dataset 1. The values of the metrics for ResNet50 are highlighted above the corresponding green bars.

1~5, respectively, leading all the other models for all groups. As also can be seen, ResNet50 outperformed all other models in terms of all metrics (higher green bars in Figure 4), except precision for Group 1, which is slightly lower compared to VGG16 (89.08% compared to 90.22%).

*3.2. The Best Performance Was Achieved with Dataset 5 with Mixed Size of Tiles.* We further tested the performance of the ResNet50 models trained with different datasets. The AUC and Acc are listed in Table 4. As can be seen, ResNet50 performed the best when trained with Dataset 5 with an AUC of 0.994 and an Acc of 93.16%, which consists of tiles with mixed sizes. Apart from the overall AUC and Acc, we also calculated the accuracy, recall, precision, and F1 score of ResNet50 models for each group trained with different datasets (Figure 5). Results showed that ResNet50 achieved the highest accuracy, recall, precision, and F1 score for all groups when trained on Dataset 5 (higher green bars in Figure 5), except the highest recall for Group 5, which was achieved when the model was trained on Dataset 3. Figure 6 shows the normalized confusion matrix of ResNet50 trained on Dataset 5 to better illustrate the model performance.

Table 4: AUC and accuracy of ResNet-50 models trained on different datasets.

| Dataset | AUC | Acc (%) |
|---|---|---|
| Dataset 1 (400 × 400) | 0.988 | 89.48% |
| Dataset 2 (600 × 600) | 0.989 | 90.64% |
| Dataset 3 (800 × 800) | 0.983 | 88.09% |
| Dataset 4 (1000 × 1000) | 0.979 | 87.21% |
| Dataset 5 (mixed size) | 0.994 | 93.16% |

## 4. Discussion

The present study is the first attempt to investigate the feasibility of deep learning models for automated Japanese "Group classification" based on WSIs. Specifically, we trained 5 popular CNN models, namely, VGG16, VGG19, ResNet50, Xception, and InceptionV3. Results showed that ResNet50 achieved the leading performance in terms of AUC and Acc. This is not surprising since comprehensive empirical evidence has shown that residual network can gain accuracy from considerably increased depth. To apply DL models to huge WSIs, each WSI was first segmented into small tiles. We varied the size of tiles and built five datasets that constitute tiles of different sizes (see Table 2). In clinic, pathologists observe the biopsy under various magnifications of view to enhance diagnosis. Similarly, a dataset with mixed sized tiles was also constructed in this work. We further trained ResNet50 with different datasets. Results showed that when trained with mixed sized tiles (Dataset 5), the model achieved the best performance, with an AUC of 0.994 and an Acc of 93.16%. With tiles of different sizes, the model is able to "see" the samples at various spatial scales, as the pathologists observe the biopsy under various magnifications of view. Our results suggest that DL models

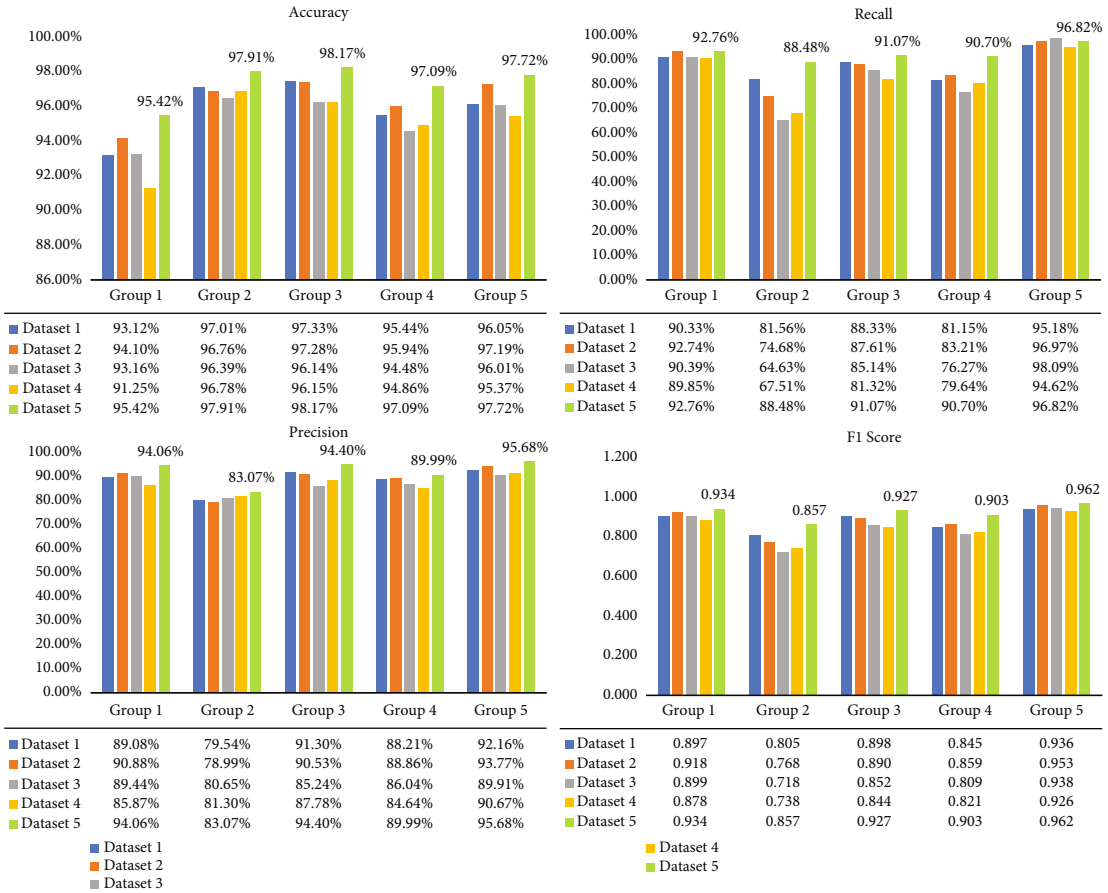



Figure 5: Performance (accuracy, recall, precision, and F1 score) of ResNet50 evaluated on the test set. The models were trained on five different datasets. The values of the metrics for ResNet50 trained on Dataset 5 are highlighted above the green bars.

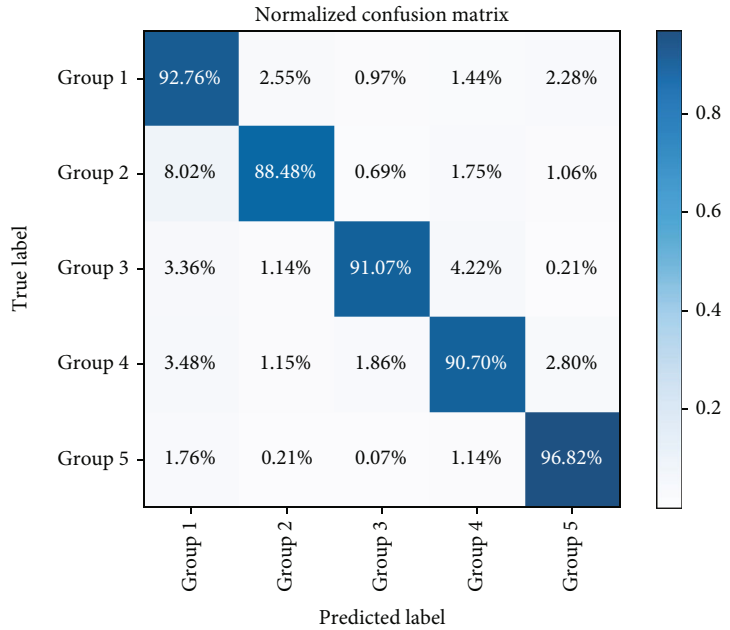



Figure 6: Normalized confusion matrix of ResNet50 on the test set. The model was trained on Dataset 5.

for automated Group classification of neoplastic lesion biopsy is promising and could help relief the workload of pathologists and increase diagnosis accuracy.

A closer investigation suggested that such improvement lies in a much higher recall for Group 2 (88.48% compared to 81.56%, 74.68%, 64.63%, and 67.51%) and Group 4 (90.70% compared to 81.15%, 83.21%, 76.27%, and 79.64%) and a fairly higher recall for Group 3 (91.07% compared to 88.33%, 87.61%, 85.14%, and 81.32%). The correct recognition of Groups 2~4 gastric biopsy is challenging in clinical practice. Meanwhile, it is critical for early detection of gastric cancer, enabling good prognosis with proper intervention. Therefore, higher recall for Groups 2~4 is of high clinical importance. Our results suggest that training the model with mixed sized tiles enables a higher recall for Groups 2~4.

In addition, our results suggest that the size of tiles has an effect on the model performance. An interesting trend revealed by our results is that the model degenerates when trained with bigger tiles. Two possible reasons may explain this. First, when segmented into big-sized tiles, the number of tiles is smaller. The samples in the training set could be insufficient for model training, leading to underfitting of the model. Second, the big-sized tiles are more likely to contain unwanted information, increasing the difficulty of model training.

There is currently no standardized treatment protocol that is globally accepted, and clinical practice alters across countries. This is mainly due to the fact that gastric cancer populations from different countries have distinct etiology, epidemiological characteristics, and clinicopathological features, especially between the East and the West [2, 4, 5]. There are several pathological diagnosis systems of gastric cancer around the world. Two main systems are the "WHO classification of tumors of digestive system" [13] and the Japanese "Group classification" [14]. Most pathologists in China are trained with the "WHO classification" while the gastroenterologists are trained with the Japanese guidelines. Such discrepancy causes a reduction in collaboration efficiency between the pathologists and gastroenterologists. Therefore, current pathological diagnosis system in China does not fully play its due role in guiding clinical treatment. The Japanese "Group classification," on the other hand, corresponds well with the endoscopic diagnostic system and could provide direct guidance for clinical treatment. Another problem that cannot be ignored in China is that the pathologists are extremely in short. Therefore, AI-aided automatic system for Japanese "Group classification" of gastric epithelial lesions is of clinical importance and the present study is the first attempt. If validated, such a system could be applied, not only to relieve the workforce of pathologists and improve their diagnosis accuracy but also to better translate pathological diagnosis to clinical practice.

To the best of our knowledge, the present study is the first to investigate the feasibility of automated Japanese "Group classification" of gastric biopsies based on pathological images. Automated pathological image analysis for human gastric epithelial lesions is not a novel topic. Most studies focus on early cancer diagnosis or cancer classification and treat the task as a binary [9, 10] or three-way classification problem [19]. For instance, Qu et al. [10] proposed a step-wise fine-tuning approach for gastric pathology image classification, where the model was first pretrained with ImageNet and further fine-tuned with a well-annotated benign/malignant dataset. Similarly, Leon et al. [9] proposed a CNN-based approach to classify gastric histopathological images as benign or malignant. On the other hand, Li et al. [19] proposed a DL-based approach for early diagnosis of gastric cancer, where non-precancerous lesion, precancerous lesion, and gastric cancer were automatedly differentiated. Sharma et al. [11] proposed a convolutional neural network for cancer classification based on immunohistochemical response and achieved 0.669 accuracy, in which the WSIs of surgical sections were classified as HER2+ tumor, HER2- tumor, or nontumor.

Table 5: The false-positive rate (FPR) and false-negative rate (FNR) of each group.

| | Group 1 | Group 2 | Group 3 | Group 4 | Group 5 |
|---|---|---|---|---|---|
| FPR | 5.94% | 16.93% | 5.60% | 10.01% | 4.32% |
| FNR | 7.24% | 11.52% | 8.93% | 9.30% | 3.18% |

A common limitation of DL models for medical applications is that their interpretability is very weak. Their decision should be suggestive or assisted, rather than deterministic. However, they are still quite helpful. One important application of the present study is to quickly screen out Group 1 and Group 5 biopsy, which is defined as normal tissue or nonneoplastic lesion tissue and carcinoma, respectively. Although the pathological diagnosis of Group 1 and Group 5 is relatively easy, it still takes time. Such automated screening could greatly reduce the workload of pathologists so that they have more time spending on other suspicious specimens, thus increasing the diagnosis accuracy, as well as efficiency. ResNet50 achieved 7.24% false-negative rate for Group 1 and 4.32% false-positive rate for Group 5, respectively, suggesting its potential to be applied as so. The false-positive rate and false-negative rate for each group are provided in Table 5. The false-negative rate for Group 5 indicates the rate of missed diagnosis of carcinoma, which can have a more negative effect on the patient. Therefore, a low value is highly expected. Our results showed 3.18% false-negative rate for Group 5, which is acceptable for an assistive screening system. Moreover, the inference time for one WSI is about 30 seconds, which is shorter than conventional diagnosis by pathologists with a microscope. The elapsed training time is about 7 hours.

One limitation of the present study is that all WSIs were collected from one center and produced by an identical digital scanner. WSIs from multiple centers with various digital scanners should be included for further validation of this approach.

## 5. Conclusions

This paper presents the first attempt to investigate the applicability of convolutional neural networks for automated

Japanese "Group classification" of WSIs of gastric endoscopic biopsy. Five popular CNNs were trained and tested. Results showed that ResNet50 achieved the best performance with an accuracy of 93.16% and an AUC of 0.994. Our results demonstrated the applicability and effectiveness of DL-based system for automated Japanese "Group classification" of gastric endoscopic biopsy.

## Abbreviations

AI: Artificial intelligence
AUC: Area under the ROC curve
CNN: Convolutional neural network
DL: Deep learning
FN: False negative
FP: False positive
HER2+: Human epidermal growth factor receptor-2 positive
HER2-: Human epidermal growth factor receptor-2 negative
H&E: Hematoxylin and eosin
RNN: Recurrent neural network
ROC: Receiver operating characteristic curve
WHO: World Health Organization
WSI: Whole slide image
TP: True positive.

## Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding authors on reasonable request.

## Ethical Approval

All procedures performed in studies involving human participants were approved by the Medical Ethics Committee of Dalian Municipal Central Hospital and in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

## Consent

Informed written consent was also obtained from individual participants included in the study.

## Conflicts of Interest

The authors declare that they have no conflict of interest.

## Authors' Contributions

Xiaoya Fan, Lihui Yu, Xin Qi, and Xue Lin contributed equally to this work.

## Acknowledgments

This study was supported by the Dalian Science and Technology Innovation Fund (grant number 2021JJ12WZ46) from the Dalian Science and Technology Bureau, China.

## References

[1] Y. He, Y. Wang, F. Luan et al., "Chinese and global burdens of gastric cancer from 1990 to 2019," *Cancer Medicine*, vol. 10, no. 10, pp. 3461–3473, 2021.

[2] A. Jemal, E. M. Ward, C. J. Johnson et al., "Annual report to the nation on the status of cancer, 1975–2014, featuring survival," *JNCI: Journal of the National Cancer Institute*, vol. 109, no. 9, 2017.

[3] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[4] Z. Bu and J. Ji, "A current view of gastric cancer in China," *Translational Gastrointestinal Cancer*, vol. 2, pp. 1–4, 2013.

[5] F.-H. Wang, X.-T. Zhang, Y.-F. Li et al., "The Chinese Society of Clinical Oncology (CSCO): clinical guidelines for the diagnosis and treatment of gastric cancer, 2021," *Cancer Communications*, vol. 41, no. 8, pp. 747–795, 2021.

[6] L. Zong, M. Abe, Y. Seto, and J. Ji, "The challenge of screening for early gastric cancer in China," *The Lancet*, vol. 388, no. 10060, p. 2606, 2016.

[7] K. Yang and J.-K. Hu, "Gastric cancer treatment: similarity and difference between China and Korea," *Translational Gastroenterology and Hepatology*, vol. 2, p. 36, 2017.

[8] Y. Li, X. Li, X. Xie, and L. Shen, "Deep learning based gastric cancer identification," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 182–185, Washington, DC, USA, 2018.

[9] F. Leon, M. Gelvez, Z. Jaimes, T. Gelvez, and H. Arguello, "Supervised classification of histopathological images using convolutional neuronal networks for gastric cancer detection," in *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, pp. 1–5, Bucaramanga, Colombia, 2019.

[10] J. Qu, N. Hiruta, K. Terai, H. Nosato, M. Murakawa, and H. Sakanashi, "Gastric pathology image classification using stepwise fine-tuning for deep neural networks," *Journal of Healthcare Engineering*, vol. 2018, Article ID 8961781, 13 pages, 2018.

[11] H. Sharma, N. Zerbe, I. Klempert, O. Hellwich, and P. Hufnagl, "Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology," *Computerized Medical Imaging and Graphics*, vol. 61, pp. 2–13, 2017.

[12] O. Iizuka, F. Kanavati, K. Kato, M. Rambeau, K. Arihiro, and M. Tsuneki, "Deep learning models for histopathological classification of gastric and colonic epithelial tumours," *Scientific Reports*, vol. 10, no. 1, p. 1504, 2020.

[13] I. D. Nagtegaal, R. D. Odze, D. Klimstra et al., "The 2019 WHO classification of tumours of the digestive system," *Histopathology*, vol. 76, no. 2, pp. 182–188, 2020.

[14] Japanese Gastric Cancer Association, "Japanese gastric cancer treatment guidelines 2014 (ver. 4)," *Gastric Cancer*, vol. 20, no. 1, pp. 1–19, 2017.

[15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015, https://arxiv.org/abs/1409.1556.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on*

*computer vision and pattern recognition*, pp. 770–778, Las Vegas, NV, USA, 2016.

[17] F. Chollet, "Xception: deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251–1258, Honolulu, HI, USA, 2017.

[18] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, Las Vegas, NV, USA, 2016.

[19] Y. Li, Y. Li, L. Deng et al., "Early diagnosis of gastric cancer based on deep learning combined with the spectral-spatial classification method," *Biomedical Optics Express*, vol. 10, no. 10, pp. 4999–5014, 2019.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 8418048, 12 pages
https://doi.org/10.1155/2022/8418048

*Research Article*

# Optimal Modeling of Anti-Breast Cancer Candidate Drugs Based on Graph Model Feature Selection

**Rongyuan Chen**,[1] **Zhixiong He**,[2] **Shaonian Huang**,[1] **Lizhi Shen**,[1] **and Xiancheng Zhou**[1]

[1]*Key Laboratory of Hunan Province for Statistical Learning and Intelligent Computation, Hunan University of Technology and Business, Hunan Changsha 410205, China*
[2]*Hunan University of Technology and Business, Hunan Changsha 410205, China*

Correspondence should be addressed to Xiancheng Zhou; zxc6501@126.com

Received 17 May 2022; Revised 27 June 2022; Accepted 26 July 2022; Published 30 August 2022

Academic Editor: Kehua Guo



Breast cancer is one of the most widespread and fatal cancers in women. At present, anticancer drug-inhibiting estrogen receptor $\alpha$ subtype (ER$\alpha$) can greatly improve the cure rate for breast cancer patients, so the research and development of this kind of drugs are very urgent. In this paper, the problem of how to screen excellent anticancer drugs is abstracted as an optimization problem. Firstly, the graph model is used to extract low-dimensional features with strong distinguishing and describing ability according to various attributes of candidate compounds, and then, kernel functions are used to map these features to high-dimensional space. Then, the quantitative analysis model of ER$\alpha$ biological activity and the classification model based on ADMET properties of the support vector machine are constructed. Finally, sequential least square programming (SLSQP) is utilized to solve the ER$\alpha$ biological activity model. The experimental results show that for anticancer data sets, compared with principal component analysis (PCA), the error rate of the graph model constructed in this paper is reduced by 6.4%, 15%, and 7.8% on mean absolute error (MAE), mean squared error (MSE), and root mean square error (RMSE), respectively. In terms of classification prediction, compared with principal component analysis (PCA), the recall and precision rates of this method are enhanced by 19.5% and 12.41%, respectively. Finally, the optimal biological activity value (IC50_nM) 34.6 and inhibitory biological activity value (pIC50) 7.46 were obtained.

## 1. Introduction

In countries all over the world, the proportion of cancer in many factors harmful to people's health is increasing year by year. Breast cancer has been the most common cancer, with a mortality rate of 11.5% to 28.4% [1]. At present, in the process of studying the treatment of breast cancer, some scholars have found that an estrogen receptor $\alpha$ subtype (estrogen receptor alpha, ER$\alpha$) can be used as a key target for effective treatment of breast cancer, and compounds that can antagonize the activity of ER$\alpha$ can be used as candidate drugs for the treatment of breast cancer.

There are many kinds of anticancer compounds, and the extraction of low-dimensional features with strong description ability from various properties of the compounds can greatly improve the efficiency of screening anticancer drug candidates. Many scholars' abstract anticancer drug selection as feature extraction, such as Tassenberg et al. [2], use automatic feature extraction algorithm DenMap to detect single crystal dendrite core quickly, accurately, and repeatably and realize average automated feature extraction. Xue et al. [3] adopted an analytic hierarchy process entropy (HDE) feature extraction method based on the analytic hierarchy process to effectively diagnose the fault of rolling bearings, which eliminates the redundant information between features and retains the fault-related information. Yang et al. [4] used nonlinear simulation feature extraction to complete the tasks of

speech detection and keyword location in inference sensor system with low power consumption. Xue et al. [5] proposed a feature extraction method based on asymmetric probability distribution function to reconstruct the distillation curve in industrial refining process, which is beneficial to the modeling and optimization of the oil refining process. Liu et al. [6] used feature extraction method to classify the biogenetic mechanism of circular RNA, which confirmed the view that multiple biogenetic mechanisms of different subsets of human CircRNA coexist. Zhu et al. [7] proposed a lightweight single image superresolution network with an expectation-maximization attention mechanism (EMASRN) to extract feature maps of different sizes. The experimental results demonstrate the superiority of EMASRN over state-of-the-art lightweight SISR methods in terms of both quantitative metrics and visual quality.

On the other hand, whether anticancer compounds can be selected as drug candidates, we also need to consider the following five characteristics: (1) intestinal epithelial cell permeability, (2) cytochrome P450 enzyme, (3) cardiac safety evaluation of compounds, (4) human oral bioavailability, and (5) micronucleus test. These five characteristics are often referred to as ADMET properties [8–10].

For the evaluation of the screening of anticancer compounds listed above, this problem is abstracted as the problem of classification and prediction of anticancer drugs based on ADMET properties. For example, Guo et al. [11] proposed a novel Relation Separation Network (RSNet) in this paper, aiming to boost few-shot learning by improving similar-class recognition performance. Compared to PT+MAP, RSNet improves the accuracy of classification on the CUB data set by approximately 5% and that of similar-class classification by more than 10%. Wang et al. [12] used a scalable window waveform sampling method (SWWS) based on the classification pattern to classify the workload requested by all users and then reasonably predict the usage of user cloud resources to minimize the cost of use. Wang et al. [13] combined missing value analysis with likelihood ratio test, introduced weighted decay random forest model, realized ICU readmission classification based on sparse data, and greatly reduced patients' expenses. Zhang et al. [14] adopted recursive partition classification method, established a classification prediction model of 58 between derivative inhibitors and du's amastigotes, and determined its molecular target and molecular mechanism. Steckenrider and Furukawa [15] adopted a highly random road crack perception network detection and classification method based on probability formula, which allows features to be extracted from crack images and retains the uncertainty in the detection. Chen et al. [16] adopted a classification diagnosis method combining FTIR near-infrared spectroscopy (NIRS) and support vector machine to differentiate malignant pleural effusion (MPE) from benign pleural effusion (BPE). Barth et al. [17] adopted the classification method of combining principal component analysis (PCA) with $K$-nearest neighbor (KNN) to address the problem of high correlation variables in wine classification. Lamge et al. [18] adopted a skin disease detection and classification method based on the combination of image processing technology and neural network to classify and evaluate patients' skin lesions images. Schultz et al. [19] used recurrent neural network and convolution neural network to classify airport performance on the basis of weather data. This method quantifies the correlation between airport performance decline and weather severity, and the prediction accuracy of aircraft take-off can reach more than 90%.

At present, the research and development of anticancer compounds in the medical field can be roughly classified into three steps. Firstly, the properties of the compounds were examined, and then, the activity model of the compound against cancer cells and the classification model of ADMET properties were constructed. Finally, the characteristic value of the antagonistic activity of the compound to cancer cells was obtained by solving the model. Therefore, in this paper, the research and development of anti-breast cancer drugs are abstracted as an optimization problem, and an optimization method based on graph model feature extraction is constructed. In this paper, the kernel function is used to map the features to a higher dimensional space to construct a nonlinear quantitative prediction function of biological activity, and then, a classification prediction model of ADMET properties of anticancer drugs based on support vector machine (SVM) is constructed as a constraint. Sequential least square programming (SLSQP) is used to efficiently and quickly solve the distinguishing variable value of the optimal biological activity value, the optimal biological activity value (IC50_nM), and the inhibitory biological activity value (pIC50). Specifically, we have studied the following four specific issues:

(i) Question 1: For a wide variety of compounds, can the graph model method designed in this paper extract low-dimensional features with stronger description ability from many attributes than the principal cost analysis method?

(ii) Question 2: Can the quantitative prediction model constructed by regression method accurately predict the biological activity of ER$\alpha$?

(iii) Question 3: Compared with principal component analysis, can support vector machine predict ADMET properties more accurately in a shorter time?

(iv) Question 4: Can the improved sequential least square programming (SLSQP) be solved faster and more accurately than the traditional intelligent optimization algorithm?

Through the research to solve the above problems, we can efficiently and intelligently predict whether the compound can become a candidate for breast cancer treatment and assist human doctors to accurately select effective anti-breast cancer drugs for breast cancer patients. Effectively

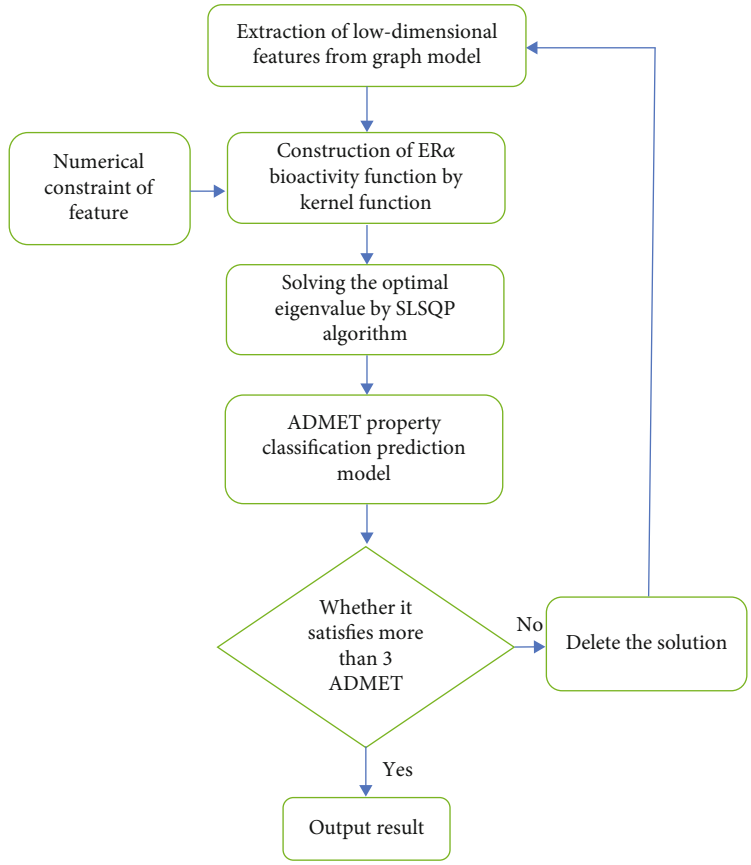



Figure 1: the overall framework.

improve the cure rate of breast cancer. The specific contributions of this article are as follows:

(1) A graph model method is proposed to extract low-dimensional features with strong descriptive ability and eliminate redundant features

(2) Through the graph model-SVM classification prediction method, five classification models for the properties of ADMET are constructed to test whether the candidate drugs are suitable for patients

(3) The bioactivity function of anticancer drug ER$\alpha$ was constructed, and the candidate anticancer drug ER$\alpha$ with optimal bioactivity was obtained by SLSQ algorithm as the best anticancer drug

## 2. Overview of Methods

The purpose of this paper is to select candidate drugs with superior efficacy from many anti-breast cancer compounds. First of all, the low-dimensional features with strong ability to describe effectively are selected from the many attributes of the compound. Then, the activity of the compound against cancer cells is measured to determine whether the compound can be used as a candidate drug. Below, we will describe the methodology framework and technical details in detail.

*2.1. General Framework.* The overall framework of this approach is shown in Figure 1. The steps for screening anticancer drugs are as follows:

(i) Step 1: use the graph model method to extract low-dimensional features with strong distinguishing ability from anti-breast cancer drug candidates

(ii) Step 2: the kernel function is used to map the extracted features to high-dimensional space to construct a quantitative analysis model of ER$\alpha$ biological activity

(iii) Step 3: at the same time, anti-breast cancer drugs also need to consider the properties of ADMET, using SVM to build a classification prediction model

(iv) Step 4: taking the ER$\alpha$ biological activity function in step 2 as the objective function and the classification prediction model in step 3 as the constraint, the improved least square method-sequential least square programming (SLSQP) is used to solve the optimal ER$\alpha$ biological activity value

*2.2. Graph Model-Minimum Spanning Tree (MST).* In view of the large number of attributes of anticancer drugs, we compare the attributes of compounds to nodes in graph theory and the correlation between attributes to the

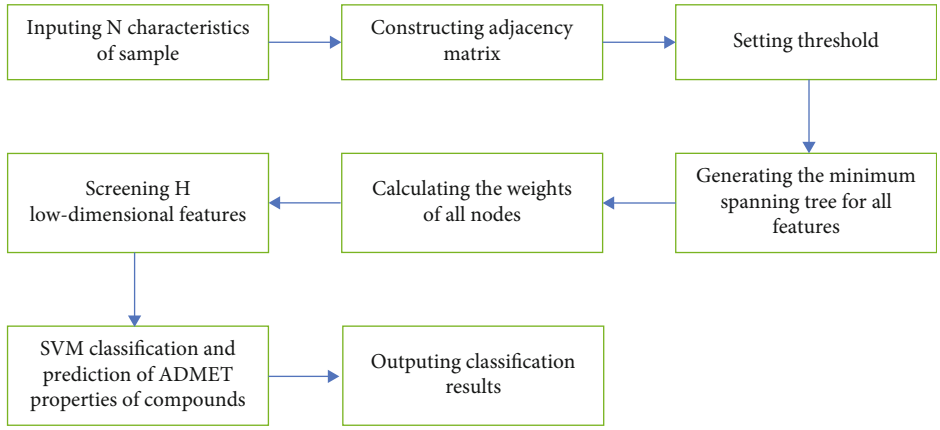



Figure 2: Graph model framework of ADMET property classification and prediction model for SVM compounds.

distance between nodes, from which the adjacency matrix between attributes is established. Finally, all nodes can generate a minimum spanning tree to extract low-dimensional features with strong description ability. When constructing the adjacency matrix, we need to choose an appropriate threshold to construct the adjacency matrix, so the crux of the problem is how to select the applicable threshold.

First of all, $N$ characteristic attributes of anticancer drugs are expressed as follows:

$$T^0 = \left\{ T^O_{m,n}, m = 1, 2, \cdots, M ; n = 1, 2 \cdots, N \right\}, \quad (1)$$

where $T^O_{m,n}$ is the data of the $n$(th) feature in the $m$(th) sample and $N = 729$, $M$ is the number of samples. The correlation coefficient matrix is

$$R_{n_1,n_2} = \frac{\sum_{m=1}^{M} \left[ T_{n_1,m} - \left( T_{n_1} \right) \right] \times \left[ T_{n_2,m} - \left( T_{n_2} \right) \right]}{\sqrt{\sum_{m=1}^{M} \left[ T_{n_1,m} - \left( T_{n_1} \right) \right]^2 \sum_{m=1}^{M} \left[ T_{n_2,m} - \left( T_{n_2} \right) \right]^2}}, \quad (2)$$

where $T_{n_1,}, T_{n_2,}$are the two features of$n_1, n_2$(th) and the similarity degree $D$ between the features is defined as follows:

$$D_{n_1,n_2} = 1 - \left| R_{n_1,n_2} \right|. \quad (3)$$

Then, the Kruskal algorithm is used to generate the minimum spanning tree according to the distance matrix and then according to the correlation coefficient between nodes. $H(H < N)$ important features (nodes) are selected. If these nodes are connected, the similarity between them is calculated, and the maximum distance value is selected as the threshold $DNI_{\min}$ of the adjacency matrix.

*2.3. Support Vector Machine (SVM).* ADMET properties of anti-breast cancer compounds determine whether they can be used as candidate drugs, and the quality of the index can be regarded as a multiattribute dichotomous problem. In order to meet the requirements of Question 3, this paper takes $N$ compound attributes of anticancer drugs as independent variables and ADMET properties as dependent variables and then constructs five classification prediction models, which are compared with the principal component analysis.

The attribute data dimension of anticancer compounds is quite high, and there is a lot of redundancy. If the classification algorithm is directly used for classification, it is difficult to get satisfactory results in a short time. Therefore, this paper comprehensively applies SVM algorithm and graph model to classify and predict. The model framework is shown in Figure 2. The ideas are as follows:

(i) For high-dimensional attribute data of compounds, the graph model is used to extract low-dimensional features with strong descriptive ability to reduce redundant information

(ii) SVM is a good binary classifier, and satisfactory results can be obtained with fewer samples

*2.4. Kernel Function and Sequential Least Square Programming (SLSQP).* In this paper, the low-dimensional features of compounds with strong descriptive ability are screened out by the graph model, and these features are mapped to high-dimensional space by kernel function to further strengthen the ability of differentiation. Then, the best nonlinear ER$\alpha$ bioactivity model was fitted by the least square method.

The relationship between features is nonlinear, and the least square method cannot effectively fit the nonlinear relationship, so we increase the least square algorithm. For solving nonlinear programming problems, it is of great importance whether the objective function and constraint conditions are continuous and smooth. If smooth, all decision variables are differentiable. The vector composed of partial derivatives of multivariate function can be used as the gradient direction indicating the fastest growth of empirical function. As the introduction of ADMET property increases the complexity of solving the problem, we add a Lagrange multiplier method (Lagrange multiplier) to the

Table 1: The degree change of the minimum spanning tree of some feature nodes.

| Features | Nodes Degree_1 | Nodes Degree_0 | Nodes Degree_diff | Features | Nodes Degree_1 | Nodes Degree_0 | Nodes Degree_diff |
|---|---|---|---|---|---|---|---|
| nHCsatu | 190 | 23 | 167 | MDEN-11 | 142 | 33 | 109 |
| nG12Ring | 156 | 0 | 156 | MLFER_BH | 178 | 69 | 109 |
| SHCsatu | 167 | 16 | 151 | maxdNH | 140 | 31 | 109 |
| nHBint3 | 166 | 29 | 137 | mindNH | 139 | 31 | 108 |
| ETA_Shape_Y | 70 | 207 | 137 | minHdNH | 138 | 31 | 107 |
| maxHBd | 38 | 167 | 129 | maxaasC | 73 | 179 | 106 |
| SsNH2 | 158 | 29 | 129 | nBondsD2 | 211 | 106 | 105 |
| nsNH2 | 158 | 30 | 128 | minsCH3 | 31 | 135 | 104 |
| nHsNH2 | 158 | 30 | 128 | MDEN-12 | 137 | 33 | 104 |
| SHsNH2 | 158 | 32 | 126 | SHBint2 | 173 | 69 | 104 |
| ATSm1 | 158 | 33 | 125 | minHBd | 49 | 151 | 102 |
| SHdNH | 155 | 31 | 124 | nBase | 162 | 61 | 101 |
| minHCsats | 12 | 136 | 124 | minaaaC | 43 | 144 | 101 |
| SdNH | 154 | 31 | 123 | minHBint6 | 2 | 102 | 100 |
| nHdNH | 154 | 31 | 123 | SHBint8 | 156 | 57 | 99 |
| ndNH | 154 | 31 | 123 | SHsOH | 81 | 180 | 99 |
| ATSc5 | 142 | 20 | 122 | nHBint2 | 163 | 66 | 97 |
| nHBint7 | 157 | 37 | 120 | SsssCH | 179 | 82 | 97 |
| nHBint8 | 162 | 46 | 116 | SaaaC | 43 | 139 | 96 |
| MLFER_BO | 179 | 68 | 111 | maxsCH3 | 43 | 138 | 95 |
| SHBint3 | 167 | 56 | 111 | maxaaaC | 46 | 141 | 95 |
| ndssC | 168 | 58 | 110 | nsssCH | 194 | 100 | 94 |
| XLogP | 73 | 183 | 110 | minHsOH | 68 | 161 | 93 |
| nHBint10 | 165 | 55 | 110 | maxssssC | 96 | 4 | 92 |
| maxHdNH | 141 | 31 | 110 | SdO | 200 | 108 | 92 |

Table 2: Weight ranking of the first 15 feature variables in graph model extraction.

| Feature number | Feature name | Weight |
|---|---|---|
| 659 | MDEC-23 | 0.110333 |
| 587 | LipoaffinityIndex | 0.100936 |
| 406 | minsssN | 0.075303 |
| 476 | maxHsOH | 0.036822 |
| 531 | maxssO | 0.036599 |
| 357 | minHsOH | 0.031988 |
| 56 | C1SP2 | 0.025563 |
| 39 | BCUTc-1l | 0.024249 |
| 673 | MLFER_A | 0.016138 |
| 652 | MLogP | 0.016063 |
| 79 | VC-5 | 0.012582 |
| 639 | nHBAcc | 0.010918 |
| 351 | minHBint5 | 0.010303 |
| 410 | minsOH | 0.009069 |
| 103 | CrippenLogP | 0.007243 |

Table 3: Error results of two feature extraction algorithms.

| Arithmetic | Evaluation index | | |
|---|---|---|---|
| | MAE | MSE | RMSE |
| Graph model | 0.5217 | 0.5044 | 0.7102 |
| PCA | 0.5571 | 0.5936 | 0.7704 |

least square algorithm and transform the constrained optimization problem into an unconstrained problem by introducing additional variables. For this reason, we construct sequential least square programming (sequential least square programming optimization algorithm, SLSQP). SLSQP efficiently preserves the nonlinear relationship between features. When solving the parameters, this method can consider the constraints other than the objective function at the same time, which meets the need of considering the ADMET properties of anticancer compounds in Question 3. The basic description of the square programming problem is as follows:

$$\begin{aligned} &\min F(x) \\ &\text{subject to } C_j(x) = 0, \quad j = 1, \cdots, MEQ \\ &C_j(x) \geq 0, \quad j = MEQ + 1, \cdots, M \\ &XL \leq x \leq XU, \quad I = 1, \cdots, N \end{aligned} \tag{4}$$

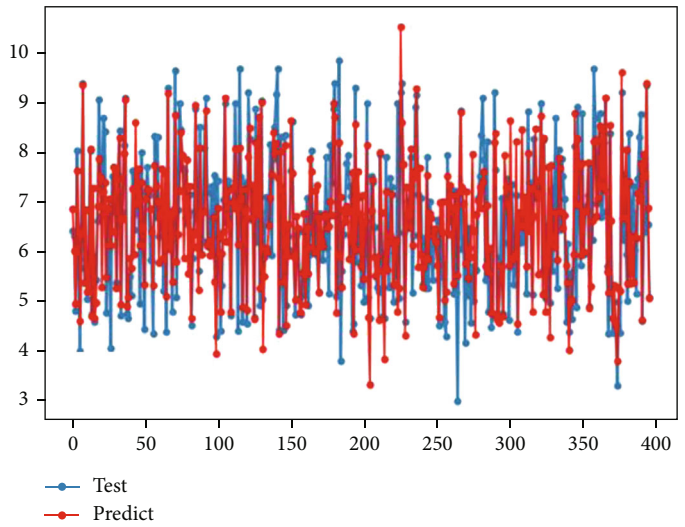



Figure 3: Optimize the fitting effect of the model for inhibiting the biological activity of ER$\alpha$.

In equation (4), $F(x)$ is the objective function. $C_j(x) = 0$ is the equality constraint. $C_j(x) \geq 0$ is the inequality constraint. $XL$ and $XU$ are the lower and upper bounds of the variable $x$. The solution process of the algorithm is as follows:

(i) Step 1: given the initial point $x^0$ and convergence accuracy $\epsilon$, set the parameter $k = 0$

(ii) Step 2: $F(x)$ is added to the Lagrangian operator at $x^0$ for Taylor expansion, and the current optimal solution $s^k$ is calculated

(iii) Step 3: $s^k$ is taken as the search direction of the next iteration, and the next iteration point $x^k$ is obtained by one-dimensional search of $F(x)$ according to constraints

(iv) Step 4: if $x^{k+1}$ satisfies the termination criterion of a given accuracy, $x^{k+1}$ is taken as the optimal solution and $F(x^{k+1})$ as the optimal cost of the objective function to terminate the calculation

## 3. Experimental Results and Analysis

In this section, we introduce the experimental environment, the source of the data set, and the specific experimental results. Depending on the four research questions designed, we have carried out comparative experiments and analysis.

The program is clearly understood by Python3.6 programming, and the program runs on a microcomputer with CPU 2.40 GHz and 8 GB memory.

The data set used throughout this paper is the D problem data set provided by the 18th Huawei Cup Mathematical Modeling Competition. The data set contains a large number of 729 attribute data of anti-breast cancer compounds and the corresponding ADMET property data.

*3.1. Question 1: Feature Selection of Anticancer Drugs Based on Graph Model.* The similarity between the characteristic variables is computed according to the degree change between the minimum spanning tree nodes. The positive and negative degree changes of some of the nodes are given in Table 1. If they are greater than the threshold $DNI_{\min}$, the two nodes are connected, and the value is 1 in the adjacency matrix. Otherwise, the value is 0. Finally, the minimum spanning tree of all nodes is obtained, and the degrees of all nodes are calculated and arranged in a descending order. The size of the value is invoked as the additional weight of the feature.

First, standardize the integrated data; then, use the graph model to solve the weight coefficients of 729 feature components; and select the first 15 feature components according to the weight, as shown in Table 2.

In order to make a quantitative comparison with PCA algorithm, MAE (mean absolute error), MSE (mean squared error), and RMSE (root mean square error) are selected to evaluate the effect of feature extraction. Among them, MAE is the average of the absolute error which can represent the actual situation of the predicted error. MSE is the expected value of the square of the difference between the estimated value and the true value of the parameter. RMSE is the arithmetic square root of MSE. They all can evaluate the change degree of the data, and the smaller their values are, the better the accuracy of the prediction model which provides a description of the experimental data. The calculation formula of each index is as follows:

$$\begin{aligned} \mathrm{MAE} &= \frac{1}{m}\sum_{i=1}^{m} |(y_i - \widehat{y}_i)|, \\ \mathrm{MSE} &= \frac{1}{m}\sum_{i=1}^{m} (y_i - \widehat{y}_i)^2, \\ \mathrm{RMSE} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\widehat{y}_i - y_i)^2}. \end{aligned} \tag{5}$$

The comparison of the three indicators of the algorithm and PCA for the data set of anti-breast cancer compounds is shown in Table 3.

It can be seen in Table 3 that the error index of the graph model is smaller and better than that of the principal component analysis. In MAE, the error rate of the graph model is 6.4% lower than that of PCA. In MSE, the error rate of the graph model is 15% lower than that of PCA. In RMSE, the error rate of the graph model is 7.8% lower than that of PCA. This shows the superiority of the graph model in extracting essential feature indexes and provides excellent characteristic variables for the construction of quantitative analysis model of biological activity of compounds against ER$\alpha$.

*3.2. Question 2: Quantitative Prediction of the Biological Activity of Anticancer Substances against ER$\alpha$.* In this paper, the kernel function is used to fit the low-dimensional features extracted from problem 1 many times by high-dimensional mapping to fit the nonlinear function. The fitting effect of the model is the best when the number of variables is 2 (for example, $x^2{}_i$); that is, the new 135 features can be constructed from the kernel function through 15 feature variables and compared with Adaboost regression and Lasso regression. The fitting effect diagram and model evaluation comparison table are shown in Figure 3 and Table 4.

As can be seen from Figure 3, the nonlinear function fitted by high-dimensional mapping of features by using kernel functions has a good fitting effect on the test set data, and most of the predicted data sets are consistent with the test set data. The fitting degree score of the function is 0.6231.

As can be seen from Table 4, compared with Adaboost and Lasso, the sequential least square programming constructed in this paper reduces the error rate by 22.4% and 32.2% on MAE, 23.8% and 44.3% in MSE, and 12.7% and 25.4% in RMSE and increases the fitting degree by 19.4% and 48.0%, respectively. This shows that the nonlinear inhibition ER$\alpha$ bioactivity optimization model constructed by sequential least square programming has good fitting effect and a small error.

Therefore, this paper constructs the following objective function:

$$\max F(x_1, x_2, \cdots, x_{15}) = \sum_{i=1}^{i=15} k_i x_i + \left( \sum_{i=1}^{i=15} k_i x_i \right)^2 + b, \quad i = 1, 2, \cdots, 15, \tag{6}$$

where $x_i$ is the characteristic variable, $k_i$ is the regression coefficient, and $b$ is the intercept of the function. The polynomial coefficients $\widehat{k_m}(m = 1, 2, \cdots, 135)$ (partial display) are shown in Table 5.

As can be seen from Table 5, the regression coefficients of these 135 features can be divided into three types: greater than 0, less than 0, and equal to 0. From the mathematical point of view, we can see that there is an inflection point in the model; that is, there is a local optimal solution. Among them, the characteristic of regression coefficient greater than 0 was positively correlated with the inhibition of cancer cell activity, the characteristic of regression coefficient less than 0 was negatively correlated with the inhibition of cancer cell activity, and the characteristic of regression coefficient equal to 0 had no effect on the inhibition of cancer cell activity.

Table 4: Comparison table of model evaluation results.

| Arithmetic | Evaluation index | | | |
|---|---|---|---|---|
| | MAE | MSE | RMSE | Score |
| SLSQP | 0.6587 | 0.7715 | 0.8783 | 0.6231 |
| Adaboost | 0.8492 | 1.0127 | 1.0063 | 0.5023 |
| Lasso | 0.9721 | 1.3849 | 1.1767 | 0.3236 |

Table 5: Regression coefficient $\widehat{k_m}$ result (partial display).

| Regression coefficient $\widehat{k_m}$ | Regression coefficient value |
|---|---|
| $\widehat{k_1}$ | 1.5541*e*-07 |
| $\widehat{k_2}$ | -0.0029 |
| $\widehat{k_3}$ | 0.2517 |
| $\widehat{k_4}$ | -0.0141 |
| $\widehat{k_5}$ | 0 |
| $\widehat{k_6}$ | -0.3898 |
| $\widehat{k_7}$ | 0 |
| $\widehat{k_8}$ | 0.1859 |
| $\widehat{k_9}$ | 0 |
| $\widehat{k_{10}}$ | 0 |
| $\widehat{k_{11}}$ | 0 |
| $\widehat{k_{12}}$ | 0.4203 |
| $\widehat{k_{13}}$ | -0.0173 |
| $\widehat{k_{14}}$ | 0 |
| ⋮ | ⋮ |
| $\widehat{k_{131}}$ | -0.1225 |
| $\widehat{k_{132}}$ | -0.0051 |
| $\widehat{k_{133}}$ | -0.0038 |
| $\widehat{k_{134}}$ | -0.0008 |
| $\widehat{k_{135}}$ | 0.0001 |

*3.3. Question 3: Classification Prediction Results of ADMET Properties of Anticancer Substances Based on Support Vector Machine (SVM).* Considering the high-dimensional attributes of anticancer compounds, firstly, PCA and graph model are used to extract features with strong distinguishing ability, and then, SVM algorithm is used to classify the features, and then, the optimal classification prediction model MST-SVM of compound ADMET is constructed and compared with PCA-SVM model. The operation flow chart of Question 3 is shown in Figure 4.

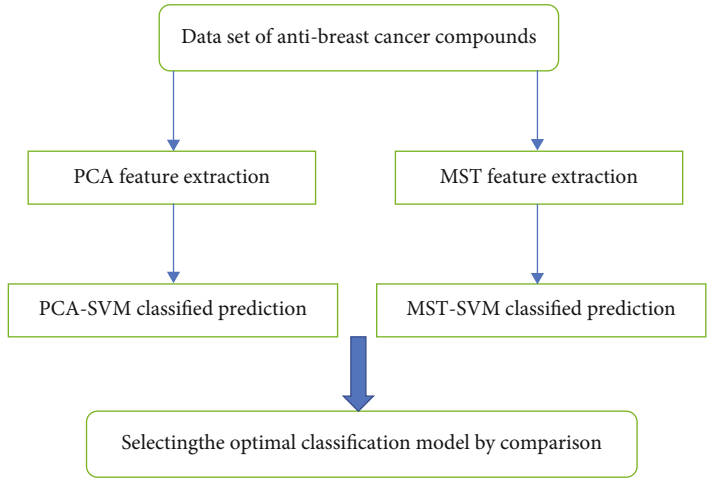



Figure 4: Question 3 operation flow chart.

Based on the 15 important features extracted by the graph model, the 15-dimensional features are input into the SVM classification model combined with the compound ADMET properties (0 or 1) for classification prediction. The classification results are shown in Figures 5–9.

Among them, blue dots and orange dots in all the classification effect maps represent 0 and 1, respectively, that is, the ADMET properties of the compound. From the classification effective images of the above five property classification models, we can see that the classification effect is obvious, and the positive and negative samples can be well distinguished. This shows that the classification prediction model of ADMET properties of MST-SVM compounds has a good classification effect. In order to make a more accurate quantitative analysis, we further introduce four indicators, namely, accuracy, accuracy, recall, and F1-score, to evaluate the classification effect, as shown in Table 6.

As can show in Table 6, the classification prediction models based on the properties of Caco-2, CYP3A4 and MN are 0.8580, 0.9379 and 1.0000 in recall, respectively. This shows that there are small false counterexamples in the classification model, and the model can achieve good results in predicting correct counterexamples. The classification prediction models based on CYP3A4 and hERG properties are 0.8947 and 0.8643 in precision, respectively. This shows that there are few correct examples in the prediction of the classification model, and the model can achieve good results in predicting the correct instances. At the same time, in accuracy, all the five classification models have higher accuracy scores. This shows that the graph model-support vector machine classification prediction model set out in the present paper can accurately judge whether the candidate drugs conform to the ADMET properties.On the basis of the above, the classification results of the ADMET properties of compounds by this method and PCA-SVM are compared as showed in Table 7:

It can be seen from Table 7 that the score of the classification prediction model of ADMET properties of PCA-SVM compounds is better than that of the MST-SVM classification prediction model in accuracy and F1-score. However,

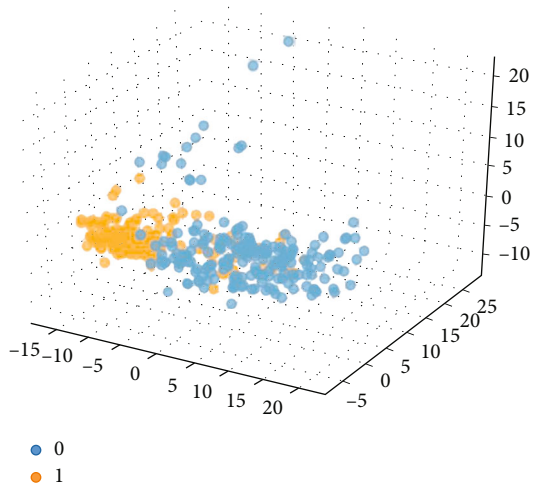



Figure 5: Classification effect diagram based on property Caco-2.

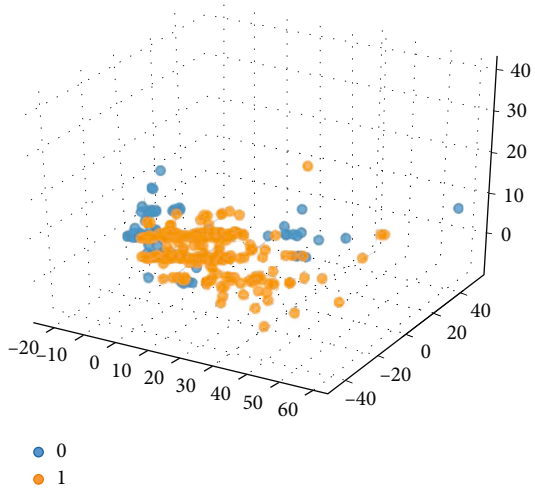



Figure 6: Classification effect diagram based on property CYP3A4.

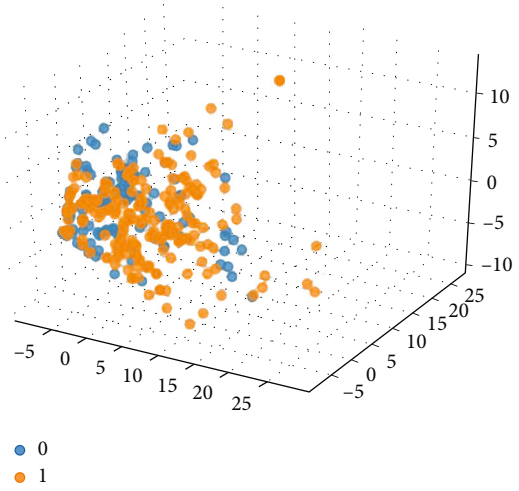



Figure 7: Classification effect diagram based on property hERG.

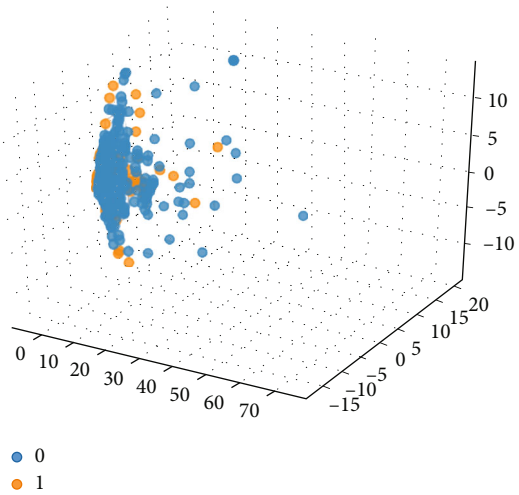



Figure 8: Classification effect diagram based on property HOB.

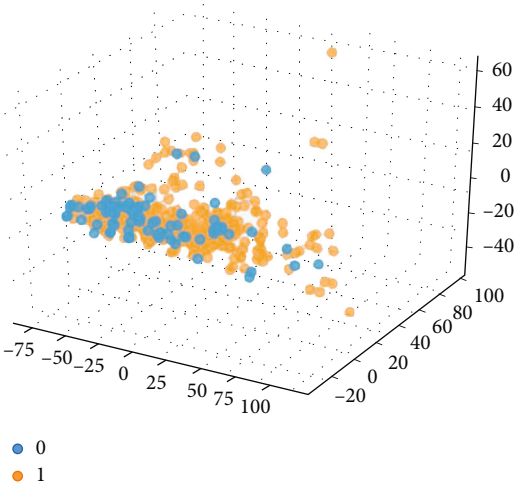



Figure 9: Classification effect diagram based on property MN.

Table 6: Graph model-SVM various accuracy results of classified prediction models.

| ADMET | Accuracy | Recall | Precision | F1-score |
|---|---|---|---|---|
| Caco-2 | 0.8127 | 0.8580 | 0.7616 | 0.8141 |
| CYP3A4 | 0.8734 | 0.9379 | 0.8947 | 0.8704 |
| hERG | 0.8025 | 0.7713 | 0.8643 | 0.8034 |
| HOB | 0.7392 | 0.0381 | 0.6667 | 0.6420 |
| MN | 0.7443 | 1.0000 | 0.7443 | 0.6352 |

the PCA-SVM model performs poorly in the recall and precision indexes. It is worth noting that in the HOB property classification, the recall and precision of the PCA-SVM model are both 0, indicating that the stability of the model in this property classification is weak. It can be seen that the high score of accuracy in additional property classification of the PCA-SVM model cannot accurately evaluate the classification effect. On the contrary, the MST-SVM model performs well in each evaluation index, and the overall score of recall and precision is better than that of the PCA-SVM model, with an average increase of 19.5% and 12.41% in recall and precision, indicating that its classification effect is stable. Therefore, based on the above analysis, this paper chooses the MST-SVM classification prediction model which is more stable.

*3.4. Question 4: Sequential Least Square Programming (SLSQP) Is Used to Solve the Quantitative Prediction Model of the Bioactivity of Anticancer Substances against ERα.* In this paper, the compound is required to optimize the inhibition of ERα biological activity (pIC50 value) under the premise of satisfying ADMET properties (at least three properties), so that the pIC50 value is the best (the higher the better), and the corresponding characteristic variables are obtained. In this paper, the equation (6) is taken as the objective function, and 15 important characteristic variables are numerically constrained.

$$\begin{cases} 10 \le X_1 \le 50, \\ 3 \le X_2 \le 12, \\ 0 \le X_3 \le 0.00001, \\ 0.5 \le X_4 \le 0.7, \\ 5 \le X_5 \le 6.5, \\ 0.4 \le X_6 \le 0.6, \\ 0 \le X_7 \le 2, \\ -0.36 \le X_8 \le -0.35, \\ 0.5 \le X_9 \le 1, \\ 0.04 \le X_{10} \le 0.3, \\ 0 \le X_{11} \le 3, \\ 2 \le X_{12} \le 3, \\ -0.1 \le X_{13} \le -0.008, \\ 3.6 \le X_{14} \le 15, \\ 60 \le X_{15} \le 80. \end{cases} \tag{7}$$

Table 7: Accuracy comparison of ADMET property classification prediction model in two classification algorithms.

| | Accuracy | | Recall | | Precision | | F1-score | |
|---|---|---|---|---|---|---|---|---|
| | PCA-SVM | MST-SVM | PCA-SVM | MST-SVM | PCA-SVM | MST-SVM | PCA-SVM | MST-SVM |
| Caco-2 | 0.8633 | 0.8127 | 0.9136 | 0.8580 | 0.7878 | 0.7916 | 0.8643 | 0.8141 |
| CYP3A4 | 0.8785 | 0.8734 | 0.9000 | 0.9379 | 0.9321 | 0.8947 | 0.8802 | 0.8704 |
| hERG | 0.8228 | 0.8025 | 0.8430 | 0.7713 | 0.8430 | 0.8643 | 0.8228 | 0.8034 |
| HOB | 0.7342 | 0.7392 | 0.0000 | 0.7381 | 0.0000 | 0.6667 | 0.6212 | 0.6420 |
| MN | 0.7722 | 0.7443 | 0.9966 | 1.0000 | 0.7670 | 0.7443 | 0.6990 | 0.6352 |

By solving the model composed of equations (6) and (7), the optimal pIC50 value is obtained. Then, the ADMET property of the compound is tested; that is to say, the SVM classification prediction model is used to make a classification prediction according to the 15 variables $x_i$ to be obtained in this problem. If the variables can satisfy more than 3 ADMET properties, then the values of the variables $x_i$ and pIC50 are directly output as the final optimization scheme, and if they are not satisfied, the new variable $\dot{x}_i$ is put into the constraint to judge.

The mathematical expression of ADMET property constraints is as follows:

$$
\begin{aligned}
M_{\text{Caco-2}} &\begin{cases} 1, & \text{the property is good,} \\ 0, & \text{the property is poor,} \end{cases} \\
M_{\text{CYP3A4}} &\begin{cases} 1, & \text{the property is poor,} \\ 0, & \text{the property is good,} \end{cases} \\
M_{\text{hERG}} &\begin{cases} 1, & \text{the property is poor,} \\ 0, & \text{the property is good,} \end{cases} \\
M_{\text{HOB}} &\begin{cases} 1, & \text{the property is good,} \\ 0, & \text{the property is poor,} \end{cases} \\
M_{\text{MN}} &\begin{cases} 1, & \text{the property is poor,} \\ 0, & \text{the property is good.} \end{cases}
\end{aligned} \tag{8}
$$

Finally, we obtained the optimal inhibitory activity value of ER$\alpha$ and the corresponding characteristic variable $x_i$ value. The results are shown in Table 8.

Depending on the above table, under the constraint of the ADMET property of the compound, the values of 15 characteristic variables, the optimal bioactivity value (IC50_nM), and the inhibitory bioactivity value (pIC50) were obtained. Among them, the distinguishing variable of positive value was positively correlated with the biological activity of ER$\alpha$. The characteristic variable of negative value was negatively correlated with the biological activity of ER$\alpha$, and the characteristic variable of 0 value had no effect on the biological activity of ER$\alpha$. Finally, the functional relationship between

Table 8: Solution results of bioactivity optimization model for inhibition of ER$\alpha$.

| | | |
|---|---|---|
| Characteristic variable $x_i$ | MDEC-23 | 10 |
| | LipoaffinityIndex | 12 |
| | minsssN | 0 |
| | maxHsOH | 0.5 |
| | maxssO | 6.5 |
| | minHsOH | 0.6 |
| | C1SP2 | 0 |
| | BCUTc-1l | -0.36 |
| | MLFER_A | 1 |
| | VC-5 | 0.3 |
| | nHBAcc | 3 |
| | minHBint5 | 3 |
| | ATSc3 | -0.1 |
| | MDEC-33 | 12.9 |
| | TopoPSA | 80 |
| Biological activity | IC50_nM | 34.6 |
| Inhibitory bioactivity | pIC50 | 7.46 |

inhibitory activity value and biological activity value is presented in the following equation:

$$F_2 = 9 - \log_{10} F_1. \tag{9}$$

In equation (9), $F_1$ is the value of biological activity and $F_2$ is the value of inhibitory activity.

It is found that 15 characteristic variable inputs satisfy more than 3 ADMET properties when they are put into the constraint. This demonstrates that the low-dimensional features screened by the graph model not only have strong ability to describe and distinguish but also perform better through ADMET properties. It can be seen that the quantitative analysis model of ER$\alpha$ biological activity and the classification model based on ADMET properties of the support vector machine can quickly and accurately screen effective compounds from anti-breast cancer drug candidates. The running time of the experimental program is 0.1369 s, and the number of iterations is 22. From the analysis of time complexity and iterative process, sequential least square programming (SLSQP) algorithm constructed in this paper is

better than the traditional intelligent optimization algorithm (such as ant colony algorithm, genetic algorithm).

## 4. Conclusions

In view of the increasing number of breast cancer patients, various kinds of anti-breast cancer candidate drugs, and great pressure on doctors to use anti-breast cancer drugs, this paper is aimed at the problem of screening anti-breast cancer candidate drugs that propose an optimal modeling method of anti-breast cancer candidate drugs based on graph model feature extraction. Compared with the traditional feature extraction methods (such as principal component analysis and random forest), the graph model feature extraction method proposed in this paper addresses the problem of large error and low accuracy of the existing methods in the evaluation index. At the same time, the classification prediction model constructed in this paper is utilized to effectively detect whether the drug will have adverse reactions to the human body when screening candidate drugs. Therefore, through the method of this paper, we can efficiently and intelligently predict whether the compound can become a candidate drug for the treatment of breast cancer and assist human doctors to accurately select effective anti-breast cancer drugs for breast cancer patients, which is of great significance to improve the cure rate of breast cancer.

## Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Acknowledgments

This work was supported in part by the Natural Science Foundation of Hunan Province (2020JJ4248, 2020JJ4251), scientific research project supported by Education Department of Hunan Province (21A0370, 20B142, and 19A270), China University Industry Research and Innovation Foundation—New Generation Information Technology Program (2020ITA09005), and the Philosophy and Social Science Foundation of Hunan Province under Grant 18YBA258.

## References

[1] X. Nguyen, M. Hooper, J. P. Borlagdan, and A. Palumbo, "A review of fam-trastuzumab deruxtecan-nxki in HER2-positive breast cancer," *Annals of Pharmacotherapy*, vol. 55, no. 11, pp. 1410–1418, 2021.

[2] K. Tassenberg, B. Nenchev, J. Strickland, S. Perry, and D. Weston, "Den Map single crystal solidification structure feature extraction: automation and application," *Materials Characterization*, vol. 171, article 110763, 2021.

[3] Q. Xue, B. Xu, C. He et al., "Feature extraction using hierarchical dispersion entropy for rolling bearing fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021.

[4] M. Yang, H. Liu, W. Shan et al., "Nanowatt acoustic inference sensing exploiting nonlinear analog feature extraction," *IEEE Journal of Solid-State Circuits*, vol. 56, no. 10, pp. 3123–3133, 2021.

[5] Y. Xue, Y. Wang, and B. Sun, "Asymmetric probability distribution function-based distillation curve reconstruction and feature extraction for industrial oil-refining processes," *Energy & Fuels*, vol. 34, no. 2, pp. 2533–2544, 2020.

[6] L. Chengyu, L. Yu-Chen, H. Hsien-Da, and W. Wang, "Biogenesis mechanisms of circular RNA can be categorized through feature extraction of a machine learning model," *Bioinformatics*, vol. 35, no. 23, pp. 4867–4870, 2020.

[7] X. Zhu, K. Guo, S. Ren, B. Hu, M. Hu, and H. Fang, "Lightweight image super-resolution with expectation-maximization attention mechanism," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 3, pp. 1273–1284, 2022.

[8] D. Gajjar Normi, M. Dhekelia Tejas, and G. B. Shah, "In search of RdRp and Mpro inhibitors against SARS CoV-2: molecular docking, molecular dynamic simulations and ADMET analysis," *Journal of Molecular Structure*, vol. 1239, p. 130488, 2021.

[9] A. Bhakhar Kaushikkumar, D. Gajjar Normi, B. Bodiwala Kunjan, K. Sureja Dipen, and T. M. Dhameliya, "Identification of anti-mycobacterial agents against mmpL3: virtual screening, ADMET analysis and MD simulations," *Journal of Molecular Structure*, vol. 1244, article 130941, 2021.

[10] D. N. ElHuda, B. Pobitra, D. P. Kishore et al., "ADMET profiling in drug discovery and development: perspectives of in silico, in vitro and integrated approaches," *Current Drug Metabolism*, vol. 22, no. 7, pp. 503–522, 2021.

[11] K. Guo, C. Shen, B. Hu, M. Hu, and X. Kui, "RSNet: relation separation network for few-shot similar class recognition," *IEEE Transactions on Multimedia*, 2022.

[12] X. Wang, J. Cao, D. Yang, Z. Qin, and R. Buyya, "Online cloud resource prediction via scalable window waveform sampling on classified workloads," *Future Generation Computer Systems*, vol. 117, pp. 338–358, 2021.

[13] B. Wang, S. Ding, X. Liu, X. Li, and G. Li, "Predictive classification of ICU readmission using weight decay random forest," *Future Generation Computer Systems*, vol. 124, pp. 351–360, 2021.

[14] Z. Yuezhou, X. Henri, and G. Leo, "Predictive classification models and targets identification for betulin derivatives as Leishmania donovani inhibitors," *Journal of Cheminformatics*, vol. 10, no. 1, 2020.

[15] J. Steckenrider and T. Furukawa, "Continuum detection and predictive-corrective classification of crack networks," in *2019 22th International Conference on Information Fusion (FUSION)*, Ottawa, ON, Canada, 2020.

[16] Z. Chen, K. Chen, Y. Lou, J. Zhu, W. Mao, and Z. Song, "Machine learning applied to near-infrared spectra for clinical pleural effusion classification," *Scientific Reports*, vol. 11, no. 1, 2021.

[17] J. Barth, D. Catenulate, C. Yang, and J. Cao, "Classification of wines using principal component analysis," *Economics*, vol. 16, no. 1, pp. 56–67, 2021.

[18] R. Lamge, V. Karmaran, G. Hakke, and S. Pansambal, "Skin disease detection and classification using image processing algorithm," *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology*, vol. 12, Supplement 1, pp. 169–171, 2020.

[19] M. Schultz, S. Reitmann, and S. Alam, "Predictive classification and understanding of weather impact on airport performance through machine learning," *Transportation Research Part C Emerging Technologies*, vol. 131, no. 10, article 103119, 2021.