IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 17, 2019, accepted February 1, 2019, date of publication February 5, 2019, date of current version March 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2897644

A Novel Hybrid Model Based on TVIW-PSO-GSA Algorithm and Support Vector Machine for Classification Problems

HONGXIN XUE[1,2], YANPING BAI[2], HONGPING HU[2], TING XU[2], AND HAIJIAN LIANG[3]

[1]School of Information and Communication Engineering, North University of China, Taiyuan 030051, China
[2]Department of Mathematics, School of Science, North University of China, Taiyuan 030051, China
[3]National Key Laboratory for Electronic Measurement Technology, Key Laboratory of Instrumentation Science and Dynamic Measurement Ministry of Educations, School of Information and Communication Engineering, North University of China, Taiyuan 030051, China

Corresponding author: Yanping Bai (baiyp666@163.com)

This work was supported in part by the Shanxi Natural Science Foundation under Grant 201701D121012, Grant 201801D121026, and Grant 201701D221121, in part by the National Nature Science Foundation of China under Grant 61774137, in part by the Shanxi Scholarship Council of China under Grant 2016-088, and in part by the North University of China Postgraduate Scientific and Technological Innovation Project under Grant 20181549.

ABSTRACT The increasingly serious haze problem in China has brought about a growing public awareness of air quality. Precise air quality index (AQI) forecasts play an important role in both controlling air pollution and promoting the sustainable development of human society. However, the randomness, non-stationarity, and irregularity of the AQI series make its classifications very difficult. This paper introduces a time-varying inertia weighting (TVIW) strategy based on a combination of gravitation search algorithm (GSA) and particle swarm optimization (PSO) called the TVIW-PSO-GSA. The TVIW-PSO-GSA is utilized to optimize the penalty parameter C and kernel function parameter γ of a support vector machine (SVM) to create a hybrid TVIW-PSO-GSA-SVM algorithm. Twenty-three benchmark functions, five UCI datasets, and an AQI hierarchical classification example are tested to find that the convergence speed and performance of the TVI-PSO-GSA exceed those of other algorithms, and the TVIW-PSO-GSA-SVM algorithm also achieves higher classification accuracy and efficiency than the PSO-GSA-SVM, GSA-SVM, GA-SVM, or PSO-SVM, which indicates that the TVIW-PSO-GSA-SVM reliably and accurately classifies AQI and UCI datasets.

INDEX TERMS Intelligent optimization algorithm, GSA, PSO, time-varying inertia weighting strategy, SVM.

I. INTRODUCTION

Rapid advancements in society, economy, industrialization, urbanization, and modernization of transportation have brought about the excessive consumption of fossil fuels (coal, oil, and gas) thus deteriorating the quality of air, particularly in urban areas [1]. Many studies have shown that long-term exposure to high concentrations of atmospheric pollutants can severely harm human health, such as by greatly aggravating the risk of asthma and/or causing bronchial inflammation, pulmonary dysfunction, cardiovascular diseases, and cerebrovascular diseases [2]–[4]. Accurate air quality level forecasts are beneficial in allowing relevant departments to execute timely control measures, adjust pollutant emissions, and reduce the occurrence of major disasters.

The associate editor coordinating the review of this manuscript and approving it for publication was Jinsong Wu.

The SVM was first proposed by Vapnik in 1995. It is a machine learning algorithm based on the VC dimension theory of statistical theory and the structural risk minimization principle [5]. SVM serves to create a classification hyper-plane as the decision surface; this hyper-plane separates positive from negative samples and maximizes the isolation edge between them [6]. SVM has not only enriched statistical theory itself, but also allowed for advancements in text categorization [7]–[9], image analysis [10]–[12], handwriting recognition [13], [14], face recognition [15]–[17], fault diagnosis [18], [19], biological sciences [20]–[22], and other applications.

Parameter settings directly affect the SVM's classification accuracy [23], [24]. Parameters to be optimized include the penalty parameter C and kernel function parameter γ for the Radial Basis Function (RBF) [25]. Various heuristic

optimization algorithms have been proposed in recent years to optimize SVM parameters for improved classification and prediction accuracy, including the Grid Search (GS) algorithm [26], [27], Genetic Algorithm (GA) [28], [29], Particle Swarm Optimization (PSO) algorithm [30], [31], Fruit Fly Optimization Algorithm (FOA) [32], [33], Ant Colony Optimization (ACO) algorithm [34], and Gravitational Search Algorithm (GSA) [35], [36].

The GA is a search algorithm based on the natural, biological evolutionary process of ''survival of the fittest'' [37]. It works based on organic coordination between efficiency and stability in solving optimization problems without sacrificing robustness. GAs have been used to solve many challenging problems across various research fields [38]–[40]. To operate a GA, first, a population of a certain size is randomly generated for the problem to be solved. The adaptive value of each individual is calculated and a fitness assessment is performed for all individuals in the group. Second, by selecting, crossing, and mutating a group of individuals, a set of individuals more adaptable to the environment is produced. Finally, based on the new generation, the three operations to select, cross, and mutate are carried out. After several generations of evolution the optimal solution to the problem is identified and the set termination conditions are satisfied [41].

The GSA is an emerging swarm intelligence optimization algorithm which is based on Newton's law of gravitation and the interaction between particles [42]. It has several advantages over other evolutionary algorithms—it is straightforward, easily implemented, computationally efficient, has few control parameters to adjust, and features high practicability [43], [44]. It has been successfully applied to solve complex problems in both the scientific research world and industrial application fields such as data clustering [45], decision function estimation [46], fault diagnosis [47], virtual enterprise [48] and others. Despite its powerful search ability, GSA has notable shortcomings such as poor local optimization ability [49], [50] and premature convergence [51] in the optimization process. There has been a great deal of research on improvements to the GSA in recent years. Liu *et al.* [52], for example, proposed an improved GSA that integrates niche technology and a GA crossover operator to solve vehicle routing problems. Xu and Wang [53] enhanced GSA performance from several different perspectives. Zhang and Ma [54] proposed an unsupervised color image segmentation method based on an improved GSA.

However, these improved GSA algorithm only uses the influence of the current position and does not consider the swarm information exchange between particles, which results in weak development capability. In this study, we first improved the memory of particles in GSA by used a memory (g_{best}) to save the best solution has found so far and tend towards it, so that make it is accessible anytime. Secondly, introduced the TVIW strategy into the GSA velocity update formula for trade-off the exploration and exploitation abilities of particles. We refer to the proposed algorithm from here on as ''TVIW-PSO-GSA''. The 23 benchmark functions were used to test the optimization performance of TVIW-PSO-GSA by comparison against PSO-GSA, GSA, GA, and PSO. Finally, this study used the TVIW-PSO-GSA algorithm to optimize the penalty parameter C and kernel function parameter γ of an SVM; this algorithm is referred to as ''TVIW-PSO-GSA-SVM''. It was applied to five UCI datasets and an AQI hierarchical classification problems.

II. THE TVIW-PSO-GSA ALGORITHM

A. PARTICLE SWARM OPTIMIZATION ALGORITHM

The PSO algorithm is a global optimization algorithm first proposed by Kenney and Eberhart in 1995. It works based on the migration and clustering behavior of birds during predation [55], [56]. Suppose that in a D-dimensional search space, a population composed of m particles is $X = \{x_1, x_2, \cdots, x_m\}$. The position of the i-th particle is $x_i = (x_{i1}, x_{i2}, \cdots, x_{iD})$ and the corresponding flying velocity is $V_i = (v_{i1}, v_{i2}, \cdots, v_{iD})$. The individual best of the particle is $P_i = (p_{i1}, p_{i2}, \cdots, p_{iD})$ and the global best is $P_g = (p_{g1}, p_{g2}, \cdots, p_{gD})$. In each iteration, particles update their velocity and position by the following formula [57]:

$$V_i(t+1) = \omega V_i(t) + c_1 r_1 (P_i - X_i(t)) + c_2 r_2 (P_g - X_i(t)). \tag{1}$$

$$X_i(t+1) = X_i(t) + V_i(t+1). \tag{2}$$

where ω is the inertia weight which controls the effect of the front velocity on the current velocity; t is the current number of iterations; c_1 and c_2 are non-negative acceleration factors; r_1 and r_2 are random numbers between [0, 1]; $V_i(t)$ is the velocity of the i-th particle at the t-th iteration; and $X_i(t)$ is the position of the i-th particle at the t-th iteration.

B. GRAVITATIONAL SEARCH ALGORITHM

The GSA was proposed in 2009 by Rashedi *et al.* of the Shahid Bahonar University of Kerman, Iran [42]–[44]. It is a meta-heuristic intelligent optimization algorithm based on Newton's law of gravitation and the interaction between particles. Suppose that N particles are contained in a D-dimensional search space. The position and velocity of the i-th particle are expressed as:

$$X_i = (x_i^1, \cdots, x_i^k, \cdots, x_i^D) \qquad i = 1, 2, \cdots, N. \tag{3}$$

$$V_i = (v_i^1, \cdots, v_i^k, \cdots, v_i^D) \qquad i = 1, 2, \cdots, N. \tag{4}$$

where x_i^k and v_i^k represent the position and velocity components of the i-th particle on the k-th dimension, respectively. At t time, the gravitational force of particle j on particle i in the k-th dimensional space is defined as follows:

$$F_{ij}^k(t) = G(t) \frac{M_{pi}(t) \times M_{aj}(t)}{R_{ij}(t) + \varepsilon} \left(x_j^k(t) - x_i^k(t) \right). \tag{5}$$

where $M_{aj}(t)$ is the active gravitational mass related to particle j, $M_{pi}(t)$ is the passive gravitational mass related to particle i, $R_{ij}(t)$ is the Euclidean distance between the i-th and j-th particles, that is $R_{ij}(t) = \|X_i(t), X_j(t)\|_2$, ε is a small

constant that ensures a non-zero denominator and $G(t)$ is the gravitational constant at iteration t, that is:

$$G(t) = G_0 e^{-\alpha \frac{t}{T}}. \tag{6}$$

where G_0 is the value of G at time t_0, α is the attenuation rate of the gravitational constant G, t is the current number of iterations, and T is the maximum number of iterations.

On the k-th dimension, the resultant force of particle i is defined as follows:

$$F_i^k(t) = \sum_{j=1, j\neq i}^{N} rand_j \cdot F_{ij}^k(t). \tag{7}$$

where $rand_j$ is a random number between [0, 1].

According to Newton's second law, the acceleration $a_i^k(t)$ of particle i in the k-dimensional space at time t is defined as follows:

$$a_i^k(t) = \frac{F_i^k(t)}{M_{ii}(t)}. \tag{8}$$

where M_{ii} is the inertia mass of the i-th particle.

In the GSA algorithm, each iteration of the particle updates its velocity and position according to the following formulas:

$$V_i^k(t+1) = rand_i \times V_i^k(t) + a_i^k(t). \tag{9}$$

$$X_i^k(t+1) = X_i^k(t) + V_i^k(t+1). \tag{10}$$

where $rand_i$ is a random variable that obeys uniform distribution between [0, 1] and is used for the sake of randomness in the search process, $X_i^k(t)$ and $V_i^k(t)$ represent the position and velocity of the particle i in the k-dimensional space at the current time.

The inertia mass of each particle is calculated according to its fitness value. Its updated formula is as follows:

$$\begin{cases} m_i(t) = \dfrac{fit_i(t) - worst(t)}{best(t) - worst(t)} \\ M_i(t) = \dfrac{m_i(t)}{\sum_{j=1}^{N} m_j(t)} \end{cases} \tag{11}$$

where $fit_i(t)$ is the fitness function value of particle X_i at t iteration, $best(t)$ and $worst(t)$ represent the optimal fitness function value and worst fitness function value of all particles at t iteration, respectively, and $M_i(t)$ is the mass of the i-th particle at iteration t.

For the minimization problem, $best(t)$ and $worst(t)$ are defined as follows:

$$best(t) = \min_{i\in\{1,\cdots,N\}} fit_i(t). \tag{12}$$

$$worst(t) = \max_{i\in\{1,\cdots,N\}} fit_i(t). \tag{13}$$

For the maximization problem, $best(t)$ and $worst(t)$ are defined as follows:

$$best(t) = \max_{i\in\{1,\cdots,N\}} fit_i(t). \tag{14}$$

$$worst(t) = \min_{i\in\{1,\cdots,N\}} fit_i(t). \tag{15}$$

The position and velocity of each particle are continually updated according to the GSA principle. The algorithm is terminated when the global optimal solution $best(t)$ reaches the preset accuracy or the maximum number of iterations has been reached.

C. TVIW-PSO-GSA

Although the GSA algorithm has strong optimization ability, it has notable shortcomings such as premature convergence, susceptibility to falling into local optimum, and slow convergence speed [49]–[51]. The GSA algorithm only uses the influence of the current position to update the position and does not consider the swarm information exchange between particles, which results in weak development capability. PSO-GSA integrates the group information exchange function of the PSO algorithm with the local search capability of GSA; the improved GSA algorithm thus not only complies with the laws of motion, but also the group communication function of PSO. The velocity update formula for PSO-GSA is:

$$V_i(t+1) = \omega \times V_i(t) + c_1' r_1' ac_i(t) + c_2' r_2' (g_{best} - X_i(t)). \tag{16}$$

where c_1' and c_2' are constants between [0, 1], r_1' and r_2' are random numbers between [0, 1], g_{best} is the best solution so far, which helps them to exploit the global best; ω is the inertia weight, $V_i(t)$ is the velocity of particle i at iteration t, and $ac_i(t)$ is the acceleration of particle i at iteration t.

The right-hand side of Eq. (16) has three parts: 1) the velocity of particles, 2) the acceleration of particles, and 3) the group-wide information-sharing among particles. The third part alter the velocity of particles. Without the third part, the particles will "fly" in the same direction until they reach the boundary. The GSA will not find an acceptable solution until there is an acceptable solution on the flight path. Without the first and second parts, the "flight" speed of particles only depend on their best position in history and there is no memory for speed. By adding the first and second parts, the particles have a tendency to expand the search space. They have the ability to explore new areas.

The inertia weight ω is mainly used to balance the global search capability and local development capability of particles in Eq. (16). A larger inertia weight results in overly rapid particle velocity and deviation from the search area of the optimal solution. A smaller inertia weight gives the particle stronger local search ability, but necessitates a longer search time for the global optimal solution. To trade-off the exploration and exploitation abilities of particles in the optimization process, the TVIW strategy is proposed here to improve the iterative efficiency and search accuracy of the algorithm:

$$\omega(t) = \omega_{max} \cdot \left(\frac{\omega_{min}}{\omega_{max}}\right)^{\frac{t}{T}}. \tag{17}$$

where ω_{max} is the maximum inertia weight, ω_{min} is the minimum inertia weight, t is the current iteration number, and T is

the maximum number of iterations-when $t = 0, \omega(t) = \omega_{max}$, and when $t = T, \omega(t) = \omega_{min}$.

The particle has strong global search ability at the beginning of the iterative process and the approximate region of the optimal solution can be determined relatively quickly. At the final iteration, a small ω decelerates the searching velocity of particles and reinforces their local development ability. The particle swarm can find the global optimal solution in the feasible solution region by fine-tuning the search. To distinguish it from the original PSO-GSA, the improved PSO-GSA algorithm based on the TVIW strategy is referred to from here on as "TVIW-PSO-GSA".

After the velocity update, the particle location is updated as follows:

$$X_i(t+1) = X(t) + V_i(t+1). \tag{18}$$

III. THE TVIW-PSO-GSA-SVM ALGORITHM

A. SUPPORT VECTOR MACHINE

SVM essentially functions by finding an optimal hyper-plane which separates positive from negative samples, then maximizing the sum (interval) of the minimum distances from two types of samples to the optimal hyper-plane [58], [59]. Consider a given training set:

$$T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}. \tag{19}$$

where $x_i \in R^n$, $y_i \in \{1, -1\}(i = 1, 2, \cdots, n)$, x_i is the input, y_i is the label for x_i, and n is the number of samples. All training samples meet the following qualifications:

$$y_i((\omega, x_i) + b) - 1 + \xi_i \geq 0. \tag{20}$$

where ξ_i is a slack variable, $\xi_i \geq 0, i = 1, 2, \cdots, n$. Maximizing the boundary of SVM by Eq. (20) is equivalent to solving the following optimization problem:

$$\begin{aligned} &min\ \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}\xi_i \\ &s.t.\ y_i((\omega, x_i) + b) - 1 + \xi_i \geq 0, \\ &\quad \xi_i \geq 0, \quad i = 1, 2, \cdots, n. \end{aligned} \tag{21}$$

where $\xi_i \geq 0$ is the relaxation term, the constant $C > 0$ is used to control the penalty level for misclassified samples that exceed the error ε.

The dual form of the above quadratic programm is:

$$\begin{aligned} &\min_{\alpha} \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{j=1}^{i}\sum_{i=1}^{n} y_i y_j \alpha_i \alpha_j K(x_i, x_j) \\ &s.t.\ \sum_{i=1}^{n} y_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq C,\ i = 1, 2, \cdots, n. \end{aligned} \tag{22}$$

where α_i represents Lagrange multipliers and kernel function $K(x_i, x_j) = \Phi(x_i)^T \cdot \Phi(x_j)$. In this study, RBFs—which are widely used and perform well—were chosen as kernel functions.

$$K(x, x_i) = exp(-\gamma\|x - x_i\|^2), \quad \gamma > 0. \tag{23}$$

where γ is the kernel function parameter.

After solving the above problem, the optimal discriminant function is as follows:

$$f(x) = sgn\left(\sum_{i=1}^{n} y_i \alpha_i K(x, x_i) + b\right). \tag{24}$$

B. TVIW-PSO-GSA-SVM

The selection of penalty parameter C and kernel function parameter γ is very important when using the SVM for classification prediction [23], [24]. In this study, we applied the proposed TVIW-PSO-GSA algorithm to optimize the C and γ of an SVM: the TVIW-PSO-GSA-SVM. The fitness function discussed here is calculated as follows:

$$MSE = \frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{m}\left(y_i^k - \overline{y_i^k}\right)^2. \tag{25}$$

where n is the number of training samples; y_i^k and $\overline{y_i^k}$ are the actual output and desired output of the i-th input unit when the k-th training sample is used, respectively.

The basic steps of TVIW-PSO-GSA-SVM algorithm are as follows.

Step 1: Parameter initialization, where a particle is constituted by C and γ; initialize the particle swarm $\{C, \gamma\}$. Determine the population size, initialize the position, velocity, and the upper and lower limits of weight for all particles, and set the number of iterations of the algorithm.

Step 2: Train SVM and calculate fitness function Eq. (25).

Step 3: Update $G(t)$, $best(t)$, $worst(t)$, and $M_i(t)$.

Step 4: Calculate the resultant force on the particles.

Step 5: Calculate the particle's acceleration and update the velocity, position, and inertia weight of the particle according to Eqs. (16), (18), and (17), respectively.

Step 6: Determine whether the optimal condition is met (i.e., if the number of iterations is maximum or the preset accuracy is achieved). If it is satisfied, the optimization process is ended, the optimal parameters $\{C^{best}, \gamma^{best}\}$ are obtained, and Step 7 begins. If not, go to Step 2 and continue the next optimization.

Step 7: Establish the SVM classification model with $\{C^{best}, \gamma^{best}\}$ via the training samples and verify with test data.

A flowchart of the TVIW-PSO-GSA-SVM algorithm process is shown in Figure 1.

IV. RESULTS AND DISCUSSION

Due to the stochastic nature of meta-heuristic and evolutionary algorithms, several test cases must be used to ensure that the superior results of any optimization are not simple happenstance. We employed several test functions with different characteristics to assess the optimization performance of TVIW-PSO-GSA, then applied TVIW-PSO-GSA-SVM to five UCI classification datasets and a real air quality classification problem to further verify said performance.

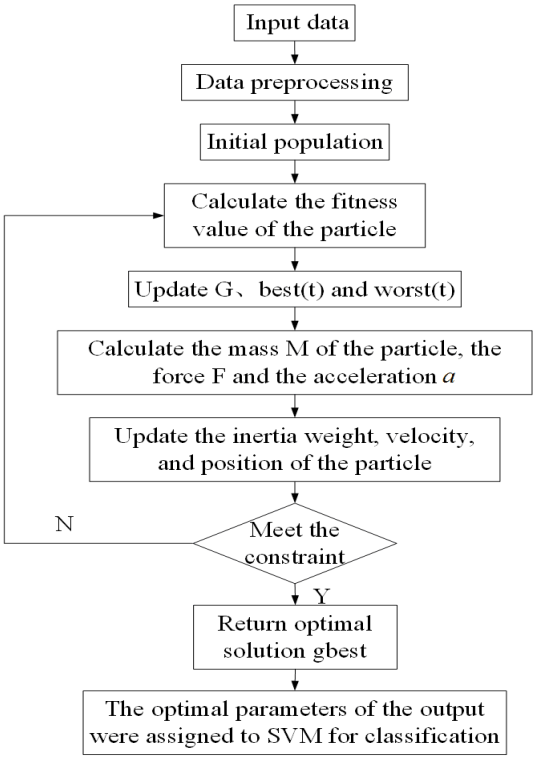


FIGURE 1. TVIW-PSO-GSA-SVM algorithm.

A. PARAMETER SETTING

For TVIW-PSO-GSA, the parameters c_1' and c_2' in Eq. (16) are positive constants that are used to adjust the step size of particles. High c_1' and c_2' values indicate that particles may cross the target region. Here, c_1' and c_2' were set to 1.5 according to Clerc's constriction factor [60]. r_1' and r_2' are uniformly distributed in the interval [0, 1], and were used to give a randomized characteristic to the search. The inertia weight ω is employed to control the impact of the previous history of velocities on the current one. If $\omega \ll 1$, then previous motion states do not significantly influence current behavior; if ω is larger, there will be a large search space, but it is difficult for the particles to change motion directions and converge to an optimal position. In other words, a larger ω facilitates searching new areas and a smaller ω facilitates a fine search. In this study, the inertia weight ω was decreased according to TVIW strategy (form 0.9-0.2) and thereby gradually changed from an exploration to exploitation. The value of α in Eq. (6) was taken from the literature [42]. G_0 is the initial value of the gravitational constant, which is a constant in the GSA algorithm; Rashedi *et al.* [42] sets the value to 100. In this paper, after several experiments, we found that the accuracy of algorithms are relatively high when G_0 is 10.

For PSO-GSA and GSA, for purposes of comparison, the value of $c_1', c_2', r_1', r_2', G_0$ and α were taken the same values as that of TVIW-PSO-GSA and ω was set to 1. For PSO, the ability of each particle in a particle swarm to reach its global optimum position was determined by the parameters c_1 and c_2 in Eq. (1). According to Clerc's [60] constriction factor, c_1 and c_2 were set to 1.5. The parameters r_1 and r_2 are used to maintain the diversity of the population and uniformly distributed in the range [0, 1]. ω is the inertia weight which reflects the influence of the particle's inertia on the velocity; it was used to control the balance between global and local development capabilities. The value of ω was taken from the literature [61].

The population sizes and maximum iteration of TVIW-PSO-GSA, PSO-GSA, GSA, GA, and PSO for the function optimization problem were set to 70 and 1000, respectively. The population sizes and maximum iteration of TVIW-PSO-GSA-SVM, PSO-GSA-SVM, GSA-SVM, GA-SVM, and PSO-SVM for the data classification problem were equal to 20 and 100, respectively.

B. FUNCTION OPTIMIZATION PROBLEM

We applied TVIW-PSO-GSA to 23 standard benchmark functions to evaluate its performance [62]–[64]. The cost function, range, and minimum of these functions are listed in the Appendix; n is the dimension of the function, f_{min} is the minimum value of the function, and S is the range of variation of optimization variables ($S \subseteq R^n$). Functions $F_1 - F_7$ are single-modal (i.e., containing only one extreme point). They are mainly used to investigate the convergence characteristics and optimization accuracy of a given algorithm. Functions $F_8 - F_{23}$ are multimodal functions (i.e., containing more than one extreme point), which are mainly used to check whether an algorithm can avoid precocity and find global optimal solutions. Functions $F_8 - F_{13}$ are multimodal functions wherein the number of local minima increases exponentially with the problem dimension. They are generally considered the most difficult class of problems for optimization algorithms. Functions $F_{14} - F_{23}$ are low-dimensional functions which have only a few local minima.

The TVIW-PSO-GSA algorithm was, as discussed above, compared to PSO-GSA, GSA, GA, and PSO for performance verification. To avoid random interference, each test function was run independently 30 times. The Average (*ave*) and Standard Deviation (*std*) of each test function are shown in Table 1. The best results are marked in bold type.

Table 1 shows that the TVIW-PSO-GSA algorithm outperforms others on the majority of the test cases. The minimum value of TVIW-PSO-GSA is better than the other algorithms on all single-modal test functions, which reflects a few notable advantages—most importantly, that TVIW-PSO-GSA has the highest accuracy in searching for optimal solutions based on the average in experiment. TVIW-PSO-GSA is also the best in terms of stability corresponding to the *std* and thus better explorative capability. The TVIW-PSO-GSA algorithm also outperforms all of the other algorithms on the majority of the multi-modal test functions ($F_8, F_{10}, F_{13}, F_{15}, F_{16}, F_{18}, F_{19}$ and $F_{21} - F_{23}$), which suggests that it effectively balances exploration and exploitation phases. This performance is a result of the TVIW strategy utilized to update the velocity. In fact, on most test problems, the

TABLE 1. Comparison of optimization results obtained for the benchmark functions.

F	TVIW-PSO-GSA		PSO-GSA		GSA		GA		PSO	
	ave	*std*	*ave*	*std*	*ave*	*std*	*ave*	*std*	*ave*	*std*
F1	**7.3966E-18**	**1.2170E-17**	4.7366E-12	2.6512E-11	0.2619	0.4904	0.0022	0.0143	0.0007	0.0020
F2	**6.5612E-09**	**4.2974E-09**	3.8897E-06	1.4606E-05	0.7906	0.6496	0.0044	0.0127	0.0620	0.0959
F3	**2.4733E-17**	**3.9187E-17**	1.2092E-08	1.0780E-07	36.7312	131.7310	392.3513	2382.1587	0.0090	0.0547
F4	**1.6126E-09**	**1.3164E-09**	1.0747E-05	4.4195E-05	0.7475	2.7602	0.3785	1.2192	0.0312	0.0732
F5	**0.6356**	**1.4534**	4.5670	68.0080	18.8701	34.2360	20.4708	85.9694	26.4893	264.9570
F6	**8.8194E-18**	**1.1194E-17**	6.7909E-12	5.5450E-11	0.2884	0.7134	0.0019	0.0164	0.0006	0.0023
F7	**0.0063**	**0.0233**	0.0078	0.0269	0.0605	0.1056	0.0074	0.0214	0.0086	0.0239
F8	**-1853.4401**	**826.5595**	-2386.3339	1766.0537	-1634.0806	976.1411	-1297.1051	1350.2251	-1815.5994	1520.1985
F9	2.5192	27.1017	7.9265	22.3138	20.4724	22.4018	**0.0007**	**0.0055**	8.0268	19.6089
F10	**3.4100E-09**	**2.9585E-09**	2.2009E-06	9.2770E-06	0.9115	0.7854	0.0164	0.0450	1.4405	5.8309
F11	0.9329	2.5612	1.0345	5.6389	28.3154	53.3819	**0.0865**	**0.1797**	3.5323	8.1573
F12	0.1234	2.7754	0.4986	6.3240	0.5135	3.8999	**3.0045E-05**	**2.5580E-04**	0.7698	5.9455
F13	**2.1215E-18**	**6.9749E-18**	1.4705E-12	1.0918E-11	0.0276	0.0375	0.0005	0.0109	0.0025	0.0224
F14	1.0974	1.6334	2.0862	8.6356	2.3409	7.8678	**0.9980**	**3.1002E-12**	1.2950	5.0621
F15	**0.0005**	**0.0017**	0.0010	0.0197	0.0024	0.0075	0.0080	0.0575	0.0013	0.0013
F16	**-1.0316**	**2.7644E-15**	-1.0316	2.7465E-15	-1.0315	0.0007	-0.9718	0.4125	-1.0316	6.2342E-06
F17	0.3979	0.0000	**0.3979**	**5.8390E-15**	0.3979	0.0003	55.6022	0.0003	0.3979	1.1887E-05
F18	**3.0000**	**1.1833E-14**	3.0000	5.4696E-13	3.0064	0.0371	3.0000	3.0034E-06	3.0001	0.0003
F19	**-3.8628**	**1.2251E-14**	-3.8470	0.3387	-3.8618	0.0053	-3.3488	1.9099	-3.8618	0.0108
F20	-3.1701	1.7972	-2.7676	2.4609	**-3.2273**	**0.2983**	-1.7236	3.4300	-3.1746	0.7329
F21	**-7.4016**	**15.3635**	-6.7461	20.0804	-5.6314	19.8307	-2.9124	6.3642	-5.4782	18.6161
F22	**-8.6206**	**17.7462**	-6.0692	20.7572	-8.2891	18.3150	-3.4414	7.9349	-6.8842	20.7536
F23	**-9.3422**	**12.7181**	-7.6919	20.4835	-8.8042	16.7260	-3.6741	7.3888	-7.4714	19.6347

TVIW-PSO-GSA algorithm is most or second-most efficient due to its comprehensive exploration mechanism, which leads to effective global optimization in the algorithm.

Figure 2 shows the convergence curves of TVIW-PSO-GSA, PSO-GSA, GSA, GA, and PSO algorithms in optimizing the test functions ($F_1 - F_{23}$). The ''average best-so-far'' indicates the average of the best solution obtained currently in each iteration over 30 runs. These figures confirm that the TVIW-PSO-GSA has superior convergence speed and performance to the other algorithms. For functions $F_{20} - F_{22}$, the ''average best so far'' of the proposed algorithm does not represent the best performance during iterations 0-150. This is because TVIW-PSO-GSA is a stochastic search algorithm, so if the initial search points are closer to the global optimum, the convergence speed is relatively quick; if the initial search points are relatively far away from global optimum, the convergence speed is relatively slow. The characteristics of the function itself also have an impact on search speed. In the TVIW strategy, the iterations T increase and ω gradually decreases, which also affect the convergence speed. For functions $F_{20} - F_{22}$, although the convergence speed of TVIW-PSO-GSA is not the fastest in the early iterations, it produced the best or suboptimal results of the iteration. In summary, the proposed TVIW strategy lends the TVIW-PSO-GSA a good balance of exploration and exploitation in gaining competitive superiority on the most of the test functions.

C. DATA CLASSIFICATION PROBLEM

As discussed above, we also applied GA, PSO, GSA, PSO-GSA, and TVIW-PSO-GSA to optimize the penalty parameter C and kernel function parameters γ of SVM; these mechanisms are called GA-SVM, PSO-SVM, GSA-SVM, PSO-GSA-SVM, and TVIW-PSO-GSA-SVM, respectively. Two benchmark problems were used to compare the abilities of these algorithms in training SVM: the UCI data classification problem and an air quality grade classification problem.

1) UCI DATA CLASSIFICATION PROBLEM

We selected five data sets from the UCI machine learning database [65] for clustering to compare our series of algorithms (Table 2). The ''No. of data points'' represents the number of instances, the ''No. of features'' represents the number of attributes, the ''No. of classes'' represents the cluster number, the ''No. of train sets'' represents the number of training samples corresponding to each data set, and the ''No. of test sets'' represents the number of test samples corresponding to each data set. In each generation, all train set samples were used to train the SVM and the test sample to test the SVM. The classification accuracy of all five algorithms is shown in Table 3. The best results are indicated in bold type.

As shown in Table 3, the classification accuracy of TVIW-PSO-GSA-SVM exceeds that of PSO-GSA-SVM, GSA-SVM, GA-SVM, or PSO-SVM on all UCI data.

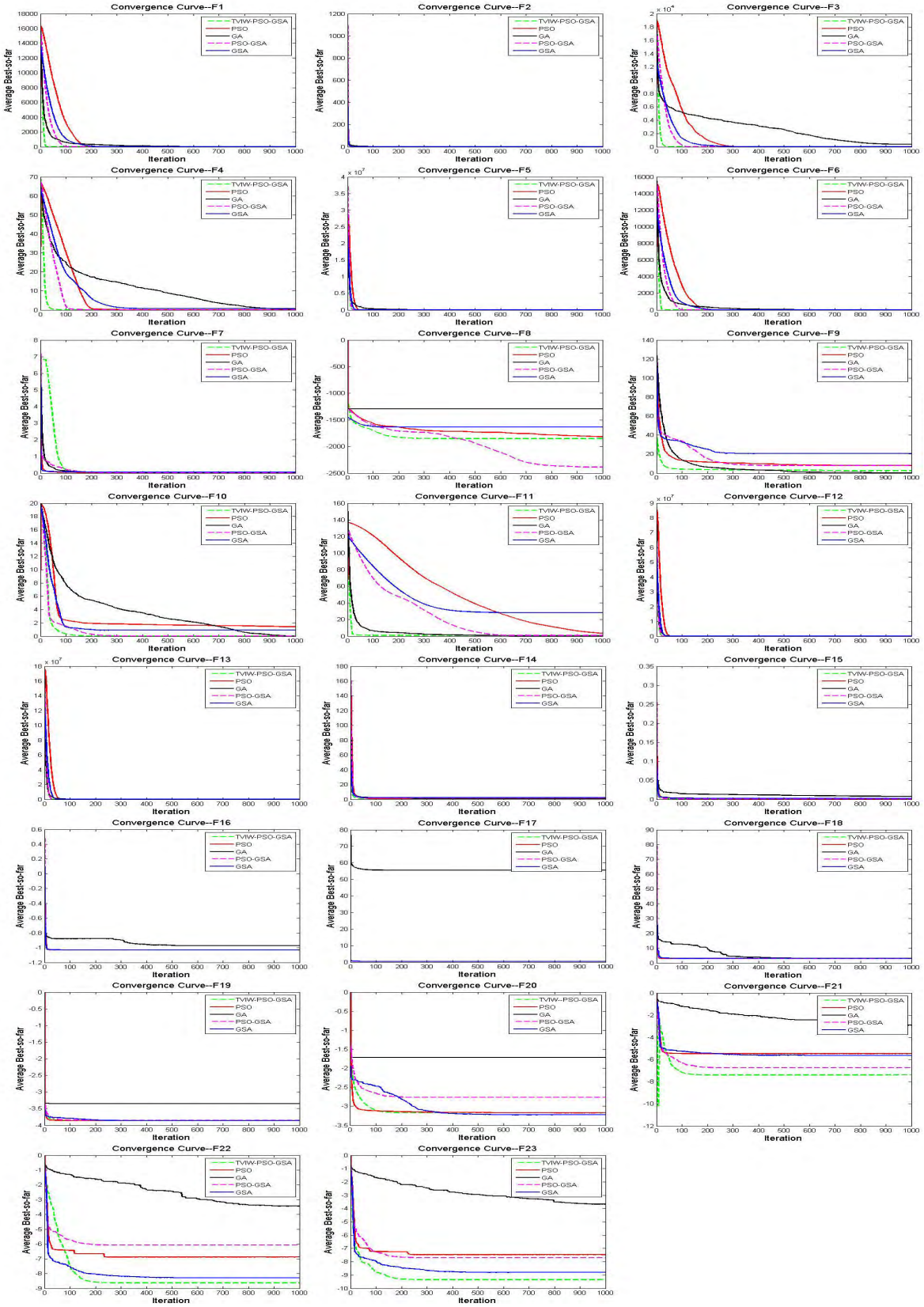


FIGURE 2. Convergence curves of five algorithms on benchmark functions $F_1 - F_{23}$.

TABLE 2. UCI data set.

Dataset	No. of data points	No. of features	No. of classes	No. of Train sets	No. of Test sets
Iris	150	4	3	130	20
Glass	214	10	7	180	34
Wine	178	13	3	160	18
Wireless	2000	7	4	1600	400
Blood	748	5	2	700	48

TABLE 3. Classification results on UCI test set data.

Method \ Dataset	Iris		Glass		Wine		Wireless		Blood	
TVIW-PSO-GSA-SVM	100%	20/20	97.06%	33/34	100%	18/18	99.75%	399/400	91.67%	44/48
PSO-GSA-SVM	90%	18/20	94.12%	32/34	94.44%	17/18	98.50%	394/400	89.58%	43/48
GSA-SVM	85%	17/20	82.35%	28/34	78%	14/18	98.75%	395/400	87.50%	42/48
GA-SVM	80%	16/20	73.53%	25/34	100%	18/18	94.25%	377/400	81.25%	39/48
PSO-SVM	95%	19/20	76.47%	26/34	94.44%	17/18	95.00%	380/400	79.17%	38/48

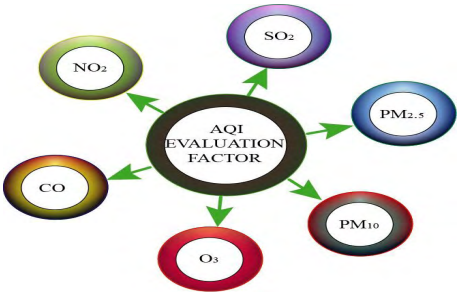


FIGURE 3. AQI main evaluation factors.

TABLE 4. AQI range and rank distribution.

AQI Ranges	AQI category	AQI Grades
$0 \sim 50$	Excellent	I
$51 \sim 100$	Good	II
$101 \sim 150$	Micro-pollution	III
$151 \sim 200$	Moderate pollution	IV
$201 \sim 300$	Heavily pollution	V
>300	Severe pollution	VI

These results prove that TVIW-PSO-GSA-SVM is capable of solving the UCI data classification problem more reliably and accurately than PSO-GSA-SVM, GSA-SVM, GA-SVM, or PSO-SVM. TVIW-PSO-GSA-SVM achieves this reliability because it is much less likely to be trapped in local minima than the other algorithms.

2) AQI CLASSIFICATION PROBLEM

The emission of air pollutants in urban areas, and specifically whether their detrimental effects on air quality can be improved, have become very popular research topics [66]. Air pollution forecasting is significant in regards to public health and pollution control [67]. The AQI is widely used in international quantitative evaluations of air quality. The main pollutants involved in AQI evaluation are shown in Figure 3. The air pollution index ranges from 0 to 500. The pollutant concentration limits average daily values of national air quality standards under grades I, II, III, IV, V, and VI corresponding to 0-50, 51-100, 101-150, 151-200, 201-300, and >300 ranges [68]. The distribution of AQI grades is shown in Table 4.

The test samples we used in this study are real-time air quality data for Taiyuan as-collected from the data center of the Ministry of Environmental Protection of the People's Republic of China (http://datacenter.mep.gov.cn/). Pollutants fine particulate matter ($PM2.5$), particulate matter ($PM10$), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), ozone (O_3), and carbon monoxide (CO) were used as evaluation indexes. The sampling period was from December 2, 2013, to May 24, 2018, at a sampling frequency of once daily. The data total was 1,632 groups. There are 101 data sets in grade I, 830 data sets in grade II, 474 data sets in grade III, 133 data sets in grade IV, 76 data sets in grade V, and 18 data sets in grade VI. A distribution histogram of the data is shown in Figure 4. We selected 1,400 data at random to train the SVM model and the remaining 232 to test the model.

We used TVIW-PSO-GSA, PSO-GSA, GSA, GA, and PSO to optimize SVM model parameters and build an air quality classification model. The results of the five models were compared in terms of air quality classification as shown in Table 5. A chart of the air quality classifications of the five model test sets is shown in Figure 5.

Table 5 and Figure 5 together indicate that TVIW-PSO-GSA-SVM has the highest classification accuracy at up to 99.14% while the classification accuracy of PSO-GSA-SVM, GSA-SVM, GA-SVM, and PSO-SVM is 96.12%, 90.52%, 56.03%, and 58.19%, respectively. The TVIW-PSO-GSA-SVM has the highest classification accuracy with the introduction of TVIW, but also improves the capability of the SVM to avoid local minima in this AQI classification problem.

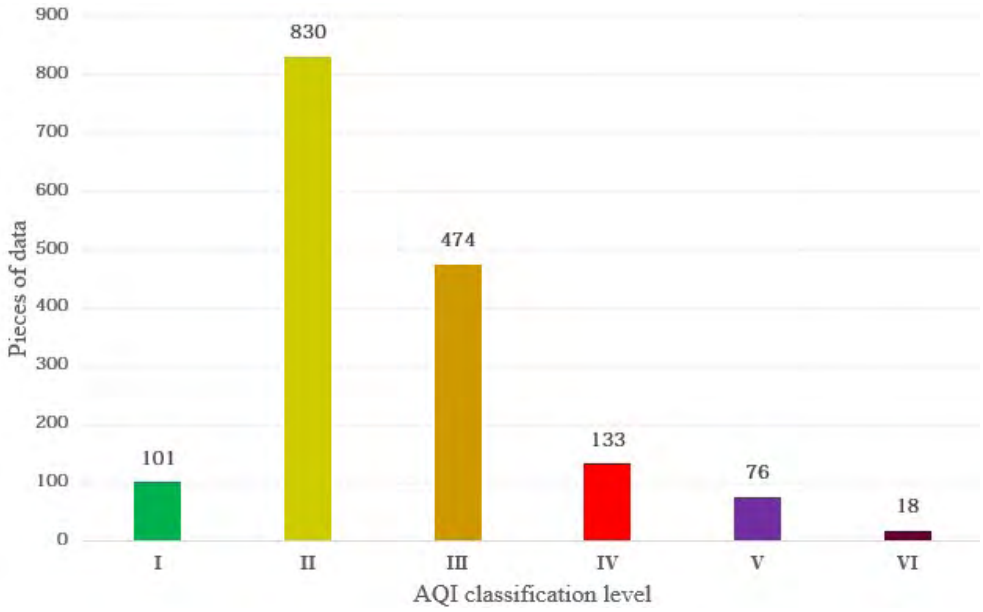


FIGURE 4. Classification distribution graph of AQI data.

TABLE 5. Classification results on test sets for AQI grades.

Method	TVIW-PSO-GSA-SVM	PSO-GSA-SVM	GSA-SVM	GA-SVM	PSO-SVM
Classification Accuracy	99.14%	96.12%	90.52%	56.03%	58.19%
	230/232	223/232	210/232	130/232	135/232

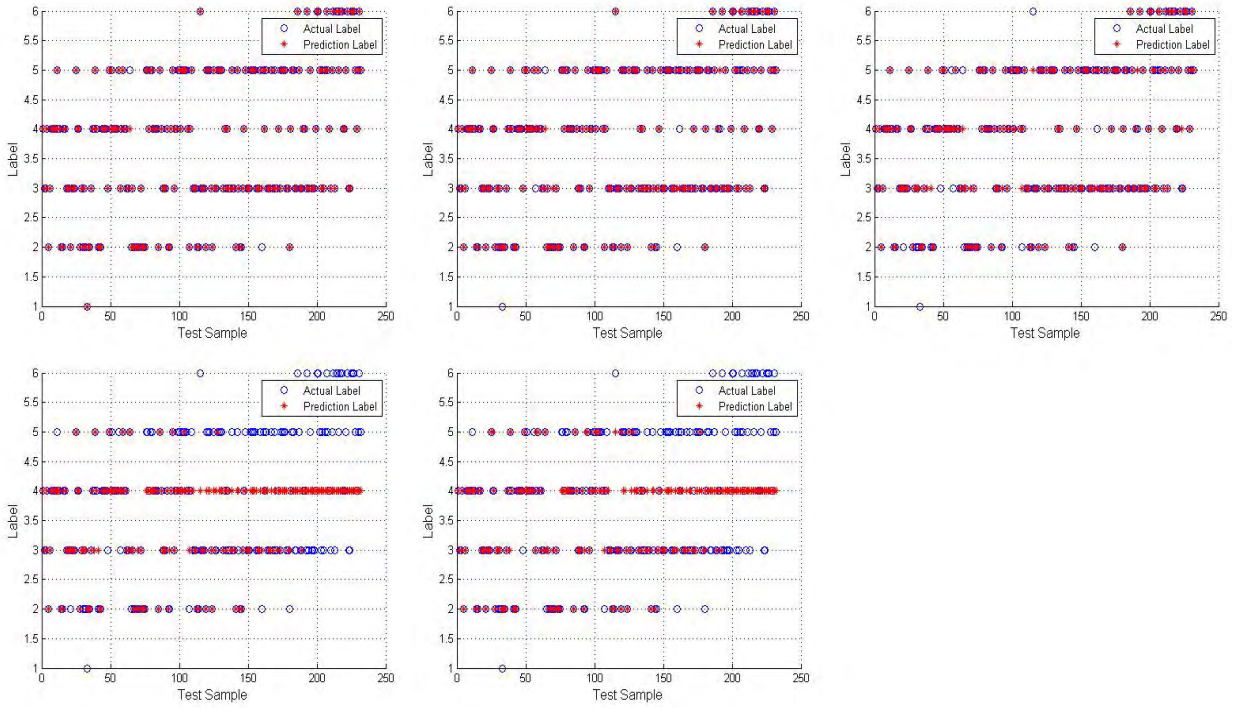


FIGURE 5. Five model classification results on test sets for AQI grades. (a) TVIW-PSO-GSA-SVM model. (b) PSO-GSA-SVM model. (c) GSA-SVM model. (d) GA-SVM model. (e) PSO-SVM model.

Tables 3 and 5 show that in the wine problem, the GA method provides the same results as TVIW-PSO-GSA-SVM, but its poor performance in the AQI problem. This is because the characteristics and distribution of wine data are not susceptible to the influence of human factors. AQI is affected by the factors both anthropogenic and meteorological; it is a

TABLE 6. Benchmark function

Test function	n	S	f_{min}
$F_1(x) = \sum_{i=1}^{n} x_i^2$	30	[-100,100]	0
$F_2(x) = \sum_{i=1}^{n} \mid x_i \mid + \prod_{i=1}^{n} \mid x_i \mid$	30	[-10,10]	0
$F_3(x) = \sum_{i=1}^{n} (\sum_{j=1}^{i} x_j)^2$	30	[-100,100]	0
$F_4(x) = \max_i \{ \mid x_i \mid, 1 \le x \le n \}$	30	[-100,100]	0
$F_5(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	30	[-30,30]	0
$F_6(x) = \sum_{i=1}^{n} (\lfloor x_i + 0.5 \rfloor)^2$	30	[-100,100]	0
$F_7(x) = \sum_{i=1}^{n} i x_i^4 + random[0,1)$	30	[-1.28,1.28]	0
$F_8(x) = \sum_{i=1}^{n} [-x_i sin(\sqrt{\mid x_i \mid})]$	30	[-500,500]	-418.9829×5
$F_9(x) = 10n + \sum_{i=1}^{n} x_i^2 - 10cos(2\pi x_i)$	30	[-5.12,5.12]	0
$F_{10}(x) = -20exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right) - exp\left(\frac{1}{n}\sum_{i=1}^{n} cos(2\pi x_i)\right) + 20 + e$	30	[-32,32]	0
$F_{11}(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} cos(\frac{x_i}{\sqrt{i}}) + 1$	30	[-600,600]	0
$F_{12}(x) = \frac{\pi}{n}\left[10\sin^2(\pi y_1) + \sum_{i=1}^{n-1}(y_i - 1)^2[1 + 10\sin^2(\pi y_{i+1})] + (y_n - 1)^2\right] + \sum_{i=1}^{n} u(x_i, 10, 100, 4)$ $y_i = 1 + \frac{1}{4}(x_i + 1)$ $u(x_i, a, k, m) = \begin{cases} k(x_i - a)^m, & x_i > a \\ 0, & -a \le x_i \le a \\ k(-x_i - a)^m, & x_i \le -a \end{cases}$	30	[-50,50]	0
$F_{13}(x) = 0.1\{\sin^2(3\pi x_1) + \sum_{i=1}^{n-1}(x_i - 1)^2[1 + \sin^2(3\pi x_{i+1})] + (x_n - 1)^2[1 + \sin^2(2\pi x_n)]\}$ $+ \sum_{i=1}^{n} u(x_i, 5, 100, 4)$	30	[-50,50]	0
$F_{14}(x) = \left(\frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^{2}(x_i - a_{ij})^6}\right)^{-1}$	2	[-65.536,65.536]	1
$F_{15}(x) = \sum_{i=1}^{11}\left[a_i - \frac{x_1(b_i^2 + b_i x_2)}{b_i^2 + b_i x_3 + x_4}\right]^2$	4	[-5,5]	0.00030
$F_{16}(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$	2	[-5,5]	-1.0316
$F_{17}(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos x_1 + 10$	2	[-5,10]×[10,15]	0.398
$F_{18}(x) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + ex_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)]$ $\times [30 + (2x_1 - 3x_2)^2 \times (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$	2	[-2,2]	3
$F_{19}(x) = -\sum_{i=1}^{4} c_i exp\left(-\sum_{i=1}^{4} a_{ij}(x_j - p_{ij})^2\right)$	3	[0,1]	-3.86
$F_{20}(x) = -\sum_{i=1}^{4} c_i exp\left(-\sum_{i=1}^{4} a_{ij}(x_j - p_{ij})^2\right)$	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

complex multidimensional nonlinear classification problem. Therefore, the classification results are not ideal.

From the above analysis, it can be concluded that TVIW-PSO-GSA-SVM also improves the capability of the SVM to avoid local minima in this data classification problem. Introducing the TVIW strategy allows the TVIW-PSO-GSA algorithm to search the approximate range of SVM parameters C and γ which made the classification accuracy was enhanced by a larger ω at the beginning of the iteration (i.e., by the high inertia weight for coarse global exploration). As the iterations progressed, the value of ω decreased and the fine search of SVM parameters was enhanced. In short,

the inertia weight decreased for finer local explorations in later iterations. The search was terminated once finding the C and γ with the highest classification accuracy (or the maximum number of iterations).

V. CONCLUSIONS

GSA has a strong exploration ability, but its slow search process is problematic. PSO has a fairly fast speed of approach to the optimal solution and can effectively optimize system parameters, but is prone to premature convergence-especially when dealing with multimodal search problems. By combining the advantages of PSO and GSA, we built an improved PSO-GSA algorithm based on a TVIW strategy in this study.

We selected 23 benchmark functions to test the optimization capability of TVIW-PSO-GSA and compared its optimization results against those of PSO-GSA, GSA, GA, and PSO to find that it has superior accuracy and stability for both single-modal and multimodal functions. In effect, introducing the TVIW strategy into the GSA velocity update formula significantly improves the explorative ability and developmental ability of particles in the system while providing stronger global optimization ability and local optimization ability to the algorithm. We also used the proposed TVIW-PSO-GSA algorithm to optimize SVM parameters, which were tested on five UCI datasets and AQI data. Compared to PSO-GSA-SVM, GSA-SVM, GA-SVM, and PSO-SVM, the TVIW-PSO-GSA-SVM approach has better classification accuracy, efficiency, and overall effectiveness.

APPENDIX
BENCHMARK FUNCTION

See Table 6.

REFERENCES

[1] J. Wu, S. Guo, H. Huang, W. Liu, and Y. Xiang, "Information and communications technologies for sustainable development goals: State-of-the-art, needs and perspectives," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2389–2406, 3rd Quart., 2018.

[2] J. Wu, S. Guo, J. Li, and D. Zeng, "Big data meet green challenges: Big data toward green applications," *IEEE Syst. J.*, vol. 10, no. 3, pp. 888–900, Sep. 2016.

[3] J. Wu, S. Guo, J. Li, and D. Zeng, "Big data meet green challenges: Greening big data," *IEEE Syst. J.*, vol. 10, no. 3, pp. 873–887, Sep. 2016.

[4] R. Atat, L. Liu, J. Wu, G. Li, C. Ye, and Y. Yang, "Big data meet cyber-physical systems: A panoramic survey," *IEEE Access*, vol. 6, pp. 73603–73636, 2018.

[5] X. Xie, "Improvement on projection twin support vector machine," *Neural Comput. Appl.*, vol. 30, no. 2, pp. 371–387, Oct. 2017.

[6] Y. Huang and L. Zhao, "Review on landslide susceptibility mapping using support vector machines," *CATENA*, vol. 165, pp. 520–529, Jun. 2018.

[7] M. Y. Jiang, X. Y. Wang, Z. F. Zhang, Q. H. Wang, J. Q. Jiang, and Z. L. Pei, "Chinese text classification study base on the improved dnn-svm algorithm," *Basic Clin. Pharmacol.*, vol. 123, no. 3, pp. 59–60, Sep. 2018.

[8] G. Feng, S. Li, T. Sun, and B. Zhang, "A probabilistic model derived term weighting scheme for text classification," *Pattern Recognit. Lett.*, vol. 110, pp. 23–29, Jul. 2018.

[9] S.-B. Kim, K.-S. Han, H.-C. Rim, and S. H. Myaeng, "Some effective techniques for Naive Bayes text classification," *IEEE Trans. Knowl. Data Eng.*, vol. 18, no. 11, pp. 1457–1466, Nov. 2006.

[10] O. Chapelle, P. Haffner, and V. N. Vapnik, "Support vector machines for histogram-based image classification," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 1055–1064, Sep. 1999.

[11] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for semisupervised classification of remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3363–3373, Nov. 2006.

[12] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification via kernel sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 217–231, Jan. 2013.

[13] B. K. Prasad and G. Sanyal, "Multiple hidden Markov model post processed with support vector machine to recognize English handwritten numerals," *Int. J. Artif. Intell. Tools*, vol. 27, no. 5, Aug. 2018, Art. no. 1850019.

[14] B. K. Prasad and G. Sanyal, "Novel features and a cascaded classifier based Arabic numerals recognition system," *Multidimens Syst. Signal Process.*, vol. 29., no. 1, pp. 321–338, Jan. 2016.

[15] H.-H. Tsai and Y.-C. Chang, "Facial expression recognition using a combination of multiple facial features and support vector machine," *Soft Comput.*, vol. 22, no. 13, pp. 4389–4405, Jul. 2017.

[16] I. Kotsia and I. Pitas, "Facial expression recognition in image sequences using geometric deformation features and support vector machines," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 172–187, Jan. 2007.

[17] T.-K. Kim and J. Kittler, "Locally linear discriminant analysis for multimodally distributed classes for face recognition with a single model image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 318–327, Mar. 2005.

[18] A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2560–2574, 2007.

[19] Z. Liu *et al.*, "Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals," *Appl. Acoust.*, vol. 89, no. 3, pp. 16–27 Mar. 2015.

[20] R. Gaudiuso *et al.*, "Using LIBS to diagnose melanoma in biomedical fluids deposited on solid substrates: Limits of direct spectral analysis and capability of machine learning," *Spectrochim. Acta. B, At. Spectrosc.*, vol. 146, pp. 106–114, Aug. 2018.

[21] Y. L. Chen and Q. Z. Li, "Prediction of apoptosis protein subcellular location using improved hybrid approach and pseudo-amino acid composition," *J. Theor. Biol.*, vol. 248, no. 2, pp. 377–381, Sep. 2007.

[22] M.-F. Lucas, A. Gaufriau, S. Pascual, C. Doncarli, and D. Farina, "Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization," *Biomed. Signal Process. Control*, vol. 3, no. 2, pp. 169–174, 2008.

[23] M. Li, Y. Liu, and J. Wang, "A new parameter optimization algorithm of SVM," in *Proc. Int. Conf. Adv. Intell. Awareness Internet (AIAI)*, Shenzhen, China, 2011, pp. 174–178.

[24] C.-L. Huang and J.-F. Dun, "A distributed PSO–SVM hybrid system with feature selection and parameter optimization," *Appl. Soft Comput.*, vol. 8, no. 4, pp. 1381–1391, Sep. 2008.

[25] P. J. G. Nieto, E. Garcia-Gonzalo, J. R. A. Fernandez, and C. D. Muniz, "A hybrid PSO optimized SVM-based model for predicting a successful growth cycle of the *Spirulina platensis* from raceway experiments data," *J. Comput. Appl. Math.*, vol. 291, pp. 293–303, Jan. 2016.

[26] J. Zhang, W. Song, B. Jiang, and M. Li, "Measurement of lumber moisture content based on PCA and GS-SVM," *J. Forestry Res.*, vol. 29, no. 2, pp. 557–564, Mar. 2018.

[27] O. Devos, C. Ruckebusch, A. Durand, L. Duponchel, and J.-P. Huvenne, "Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation," *Chemometrics Intell. Lab. Syst.*, vol. 96, no. 1, pp. 27–33, Mar. 2009.

[28] M. Zhao, C. Fu, L. Ji, K. Tang, and M. Zhou, "Feature selection and parameter optimization for support vector machines," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 5197–5204, May 2011.

[29] K. S. Sajan, V. Kumar, and B. Tyagi, "Genetic algorithm based support vector machine for on-line voltage stability monitoring," *Int. J. Elect. Power Energy Syst.*, vol. 73, pp. 200–208, Dec. 2015.

[30] X. Li, S.-D. Yang, and J.-X. Qi, "A new support vector machine optimized by improved particle swarm optimization and its application," *J. Central South Univ. Technol.*, vol. 13, no. 5, pp. 568–572, Oct. 2006.

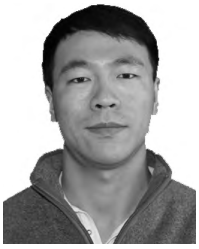
[31] L. J. Tang *et al.*, "Radial basis function network-based transform for a nonlinear support vector machine as optimized by a particle swarm optimization algorithm with application to QSAR studies," *J. Chem. Inf. Model.*, vol. 47, no. 4, pp. 1438–1445, Feb. 2007.
[32] S. Ding, X. Zhang, and J. Yu, "Twin support vector machines based on fruit fly optimization algorithm," *Int. J. Mach. Learn. Cybern.*, vol. 7, no. 2, pp. 193–203, Apr. 2016.
[33] H. Jiang *et al.*, "Construction of pancreatic cancer classifier based on SVM optimized by improved FOA," *Biomed. Res. Int.*, vol. 2015, no. 4, Aug. 2015, Art. no. 781023.
[34] O. Kadri, L. H. Mouss, and M. D. Mouss, "Fault diagnosis of rotary kiln using SVM and binary ACO," *J. Mech. Sci. Technol.*, vol. 26, no. 2, pp. 601–608, Feb. 2012.
[35] S. Sarafrazi and H. Nezamabadi-Pour, "Facing the classification of binary problems with a GSA-SVM hybrid system," *Math. Comput. Model.*, vol. 57, nos. 1–2, pp. 270–278, Jan. 2013.
[36] W. Zhang, P. Niu, G. Li, and P. Li, "Forecasting of turbine heat rate with online least squares support vector machine based on gravitational search algorithm," *Knowl-Based Syst.*, vol. 39, no. 2, pp. 34–44, Feb. 2013.
[37] K. S. Tang, K. F. Man, S. Kwong, and Q. He, "Genetic algorithms and their applications," *IEEE Signal Process. Mag.*, vol. 13, no. 6, pp. 22–37, Nov. 1996.
[38] M. Paulinas and A. Ušinskas, "A survey of genetic algorithms applications for image enhancement and segmentation," *Inf. Technol. Control*, vol. 36, no. 3, pp. 278–284, 2007.
[39] K. G. Khoo and P. N. Suganthan, "Structural pattern recognition using genetic algorithms with specialized operators," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 33, no. 1, pp. 156–165, Feb. 2003.
[40] C. Y. Tang, Y. L. Wu, and Y. H. Lai, "Fundamental matrix estimation using evolutionary algorithms with multi-objective functions," *J. Inf. Sci. Eng.*, vol. 24, no. 3, pp. 785–800, May 2008.
[41] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
[42] E. Rashedi, H. Nezamabadi-Pour, and S. Saryazdi, "GSA: A gravitational search algorithm," *J. Inf. Sci.*, vol. 179, no. 13, pp. 2232–2248, 2009.
[43] E. Rashedi, H. Nezamabadi-Pour, and S. Saryazdi, "Filter modeling using gravitational search algorithm," *Eng. Appl. Artif. Intell.*, vol. 24, no. 1, pp. 117–122, Feb. 2011.
[44] S. Duman, U. Güvenç, Y. Sönmez, and N. Yörükeren, "Optimal power flow using gravitational search algorithm," *Energ. Convers. Manage.*, vol. 59, no. 59, pp. 86–95, Jul. 2012.
[45] M. Yin, Y. Hu, F. Yang, X. Li, and W. Gu, "A novel hybrid K-harmonic means and gravitational search algorithm approach for clustering," *Expert Syst. Appl.*, vol. 38, no. 8, pp. 9319–9324, Aug. 2011.
[46] H. Askari and S.-H. Zahiri, "Decision function estimation using intelligent gravitational search algorithm," *Int. J. Mach. Learn. Cybern.*, vol. 3, no. 2, pp. 163–172, Jun. 2012.
[47] C. Li and J. Zhou, "Semi-supervised weighted kernel clustering based on gravitational search for fault diagnosis," *ISA Trans.*, vol. 53, no. 5, pp. 1534–1543, Sep. 2014.
[48] J. H. Xiao, Y. Niu, P. Chen, S. C. H. Leung, and F. Xing, "An improved gravitational search algorithm for green partner selection in virtual enterprises," *Neurocomputing*, vol. 217, pp. 103–109, Dec. 2016.
[49] C. Li and J. Zhou, "Parameters identification of hydraulic turbine governing system using improved gravitational search algorithm," *Energ. Convers. Manage.*, vol. 52, no. 1, pp. 374–381, Jan. 2011.
[50] B. Shaw, V. Mukherjee, and S. P. Ghoshal, "A novel opposition-based gravitational search algorithm for combined economic and emission dispatch problems of power systems," *Int. J. Elect. Power*, vol. 35, no. 1, pp. 21–33, Feb. 2012.
[51] S. Sarafrazi, H. Nezamabadi-Pour, and S. Saryazdi, "Disruption: A new operator in gravitational search algorithm," *Sci. Iranica*, vol. 18, no. 3, pp. 539–548, Jun. 2011.
[52] N. Liu, Y. Qiu, and B. Zhang, "An improved gravitation search algorithm forvehicle routing problem," *Mod. Transp.*, vol. 3, pp. 39–48, Apr. 2014.
[53] Y. Xu and S. T. Wang, "Enhanced version of gravitational search algorithm: Weighted GSA," *Comput. Eng. Appl.*, vol. 47, no. 35, pp. 188–192, Apr. 2011.
[54] C. Zhang and L. Ma, "Color image segmentation research based on improved GSA," *Video Eng.*, vol. 38, no. 13, pp. 39–42, Aug. 2014.
[55] J. J. Liang, A. K. Qin, P. N. Suganthan, and S. Baskar, "Comprehensive learning particle swarm optimizer for global optimization of multimodal functions," *IEEE Trans. Evol. Comput.*, vol. 10, no. 3, pp. 281–295, Jun. 2006.
[56] Z.-H. Zhan, J. Zhang, Y. Li, and H. S.-H. Chung, "Adaptive particle swarm optimization," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 6, pp. 1362–1381, Dec. 2009.
[57] G. Abbas, J. Gu, U. Farooq, A. Raza, M. U. Asad, and M. E. El-Hawary, "Solution of an economic dispatch problem through particle swarm optimization: A detailed survey—Part II," *IEEE Access*, vol. 5, pp. 24426–24445, 2017.
[58] J. Tang, Y. Tian, P. Zhang , and X. Liu, "Multiview privileged support vector machines," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3463–3477, Aug. 2018.
[59] C. Venkatesan, P. Karthigaikumar, A. Paul, S. Satheeskumaran, and R. Kumar, "ECG signal preprocessing and SVM classifier-based abnormality detection in remote healthcare applications," *IEEE Access*, vol. 6, pp. 9767–9773, 2018.
[60] M. Clerc, "The swarm and the queen: Towards a deterministic and adaptive particle swarm optimization," in *Proc. Congr. Evol. Comput.*, Washington, DC, USA, 1999, pp. 1951–1957.
[61] Y. Shi and R. C. Eberhart, "Parameter Selection in particle swarm optimization," in *Proc. 7th Int. Conf. Evol. Program. VII (EP)*. London, U.K.: Springer-Verlag, Mar. 1998, pp. 591–600.
[62] J. G. Digalakis and K. G. Margaritis, "On benchmarking functions for genetic algorithms," *Int. J. Comput. Math.*, vol. 77, no. 4, pp. 481–506, Sep. 2001.
[63] X. Yao, Y. Liu, and G. Lin, "Evolutionary programming made faster," *IEEE Trans. Evol. Comput.*, vol. 3, no. 2, pp. 82–102, Jul. 1999.
[64] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–57, May 2016.
[65] *UCI Machine Learning Database*. Accessed: Sep. 20, 2018. [Online]. Available: http://archive.ics.uci.edu/ml/datasets.html
[66] L. Yang, M. Ye, and B.-J. He, "CFD simulation research on residential indoor air quality," *Sci. Total Environ.*, vol. 472, pp. 1137–1144, Feb. 2014.
[67] B. T. Li, "The Ordinal classification and regression model and its application in judgement the rank of air quality," M.S. thesis, Fac. Sci., Yunnan Univ. Sci. Technol., Kunming, China, May 2015.
[68] L. Y. Wang, "Air quality evaluation of 31 main cities in China and the main pollutant concentration prediction," M.S. thesis, College Math., Chongqing Normal Univ., Chongqing, China, Jun. 2014.

HONGXIN XUE received the M.S. degree from the Department of Mathematics, North University of China, Taiyuan, Shanxi, China, in 2015, where she is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering. Her current research interests include artificial intelligence, signal processing, support vector machines, and intelligent optimization algorithm.

YANPING BAI received the Ph.D. degree from the North University of China, Taiyuan, Shanxi, China, in 2005, where she is currently a Professor and a Doctoral Tutor with the School of Information and Communication Engineering. Her research interests include optimization algorithm, artificial intelligence, signal processing, image processing, and MEMS reliability research.

HONGPING HU received the Ph.D. degree from the North University of China, Taiyuan, Shanxi, China, in 2009, where she is currently a Professor and a Master Tutor with the Department of Mathematics. Her research interests include combinatorial mathematics, artificial intelligence, and image processing.

TING XU was born in 1980. She is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering, North University of China, where she is a Lecturer with the Department of Mathematics. Her current research interests include artificial intelligence, pattern recognition, and signal processing.



HAIJIAN LIANG received the M.S. degree from the School of Information and Communication Engineering, North University of China, Taiyuan, Shanxi, China, in 2015, where he is currently pursuing the Ph.D. degree. His current research interests include temperature measurement and signal processing.