

Research Article

Construction of Pancreatic Cancer Classifier Based on SVM Optimized by Improved FOA

Huiyan Jiang,¹ Di Zhao,¹ Ruiping Zheng,¹ and Xiaoqi Ma²

¹Software College, Northeastern University, Shenyang 110819, China

²School of Science and Technology, Nottingham Trent University, Nottingham NG17 8NU, UK

Correspondence should be addressed to Huiyan Jiang; hyjiang@mail.neu.edu.cn

Received 17 April 2015; Revised 1 August 2015; Accepted 13 August 2015

Academic Editor: Jianhua Ruan

Copyright © 2015 Huiyan Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A novel method is proposed to establish the pancreatic cancer classifier. Firstly, the concept of quantum and fruit fly optimal algorithm (FOA) are introduced, respectively. Then FOA is improved by quantum coding and quantum operation, and a new smell concentration determination function is defined. Finally, the improved FOA is used to optimize the parameters of support vector machine (SVM) and the classifier is established by optimized SVM. In order to verify the effectiveness of the proposed method, SVM and other classification methods have been chosen as the comparing methods. The experimental results show that the proposed method can improve the classifier performance and cost less time.

1. Introduction

Pancreatic cancer is one of the world's top 10 malignant tumors [1]. Its early and accurate diagnosis is difficult. Once the diagnosis is confirmed, the tumor has reached an advanced stage. It is of great significance to improve prognosis for early detection, early diagnosis, and early treatment [2]. With the development of computer science and computer image-processing technology, computer aided detection (CAD) technology is established. CAD systems are increasingly used as an aid by radiologists for detection and interpretation of diseases [3], reducing the burden of doctors and improving the diagnosis accuracy.

Image recognition is one of the most important parts of CAD technology. The recognition process is mainly divided into two phases, namely, feature extraction and selection and classifier construction. In [4], we argue that tensors can describe space information among image features and need less space than vectors. Multilinear principal component analysis (MPCA) method [5] can be used to select the core tensors. In this paper, we also use tensors to represent CT images and MPCA to select core tensors to reduce the tensor dimension.

There are many methods to establish the medical image classifier. Kovalerchuk et al. [6] and Pendharkar et al. [7] used

machine learning and data mining technology in breast cancer detection. In recent years, many researchers have made thorough research on medical image classification. Antonie et al. [8] combined association rule and neural network to mine the texture feature in different regions of breast images and realized the automatic diagnosis of breast cancer. Zhang et al. [9] classified cervix uterus lymphonodus by support vector machine (SVM) and size and shape features. Ramírez et al. [10] proposed to use neural network method in classification of brain images of Alzheimer's disease.

However, the research of pancreatic cancer classification is in a fledging period. Tsai and Kojma [11] proposed the pancreatic tiny anomaly detection method for CT images and introduced the square of logarithm operation in grayscale to enhance the margin of low grayscale. Takada et al. [12] proposed a new pancreatic classification system to distinguish the four parts of pancreas based on the anatomy of pancreas and their own experience. He et al. [13] proposed a novel group search optimizer- (GSO-) based biomarker discovery method for pancreatic cancer diagnosis using mass spectrometry data, compared with a genetic algorithm, evolution strategies, evolutionary programming, and a particle swarm optimizer and achieved better classification performance than other algorithms.

Theoretically, imaging examination of any body tissues and organs can use CAD technology to improve diagnostic accuracy. However, since the position of pancreas is covert and has complex relationship with other organs, the pancreatic cancer image classification is difficult.

In this paper, we employ SVM [14], which are suitable for solving small-sample learning and nonlinear and high dimension problems, to establish the pancreatic cancer classification, and improve fruit fly optimization algorithm (FOA) [15] to optimize parameters of SVM. We provide a new fitness function which is more in line with the actual clinical needs in K -fold cross-validation to assess the classifier performance. Using the above strategies, the classification performance can be improved. Experimental results on pancreatic regions of abdominal CT images demonstrate the feasibility and efficiency of the proposed method.

This paper is organized as follows. Section 2 introduces the background of this researching, including support vector machine, fruit fly optimization algorithm, and the concept of quantum. Section 3 illustrates the method of construction of SVM classifier based on improved FOA. Section 4 presents the experimental data and the evaluation criterions, showing the results of the pancreatic cancer classification based on the improved FOA and other comparative methods. It also discusses the experiment results. Section 5 concludes the work in this paper.

2. Background

We introduce SVM and FOA in this part; the concept of quantum is shown in [4].

2.1. Support Vector Machine. Support vector machine (SVM) [14] is built on statistical learning theory. It is suitable for small-sample learning and nonlinear and high dimension problem. SVM is based on the principle of structural risk minimization and has strong generalization ability. It studies optimal separating hyperplane in the high dimension feature space for sample classification.

SVM mainly aims at binary classification. For linear separable problem, we consider samples as (x_i, y_i) . $x_i \in R^n$ is the feature set of medical images, $y_i \in \{+1, -1\}$ is the label of samples, and $i = 1, \dots, l$ is the number of samples. The optimal separating hyperplane is $f(x) = \omega^T x + b = 0$. The functional margin which is the distance from a sample point to separating hyperplane is $\hat{\gamma}_i = y_i(\omega^T x_i + b) = y_i f(x_i)$. The geometrical margin $\tilde{\gamma} = \gamma \hat{\gamma} = \hat{\gamma} / \|\omega\|$ is obtained by normalizing ω , and it is simplified as $\tilde{\gamma} = 1 / \|\omega\|$. The objective is to obtain the maximum value of $\tilde{\gamma}$. It is equivalent to obtain the minimum value of $\|\omega\|$. Finally, the problem translates into the quadratic programming problem as in (1), where C is penalty coefficient, and ξ_i is slack variable.

The Lagrange duality translation is conducted for (1). And (1) translates into dual problem as (2):

$$\begin{aligned} \min \quad & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l \xi_i \\ \text{s.t.} \quad & y_i (\omega^T x + b) \geq 1 - \xi_i, \quad i = 1, \dots, l \end{aligned} \quad (1)$$

$$\max_{\alpha} \quad \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle \quad (2)$$

$$\text{s.t.} \quad 0 \leq \alpha_i \leq C, \quad i = 1, \dots, l, \quad \sum_{i=1}^l \alpha_i y_i = 0.$$

The optimal separating function is shown as

$$f(x) = \text{sgn} \left(\sum_{i=1}^l \alpha_i y_i \langle x_i, x \rangle + b \right). \quad (3)$$

For nonlinear problem, the kernel function is used to translate nonlinear problem in low dimensional space into linear problem in high dimensional space. The optimal separating function is shown as

$$f(x) = \text{sgn} \left(\sum_{i=1}^l \alpha_i y_i K \langle x_i, x \rangle + b \right). \quad (4)$$

The staple kernel functions are shown as (5)~(8). In this paper, the radial basis function (RBF) as (7) is used.

Linear kernel is

$$f(x) = x'x. \quad (5)$$

Polynomial kernel is

$$f(x) = (\gamma \cdot x' \cdot x + b)^d. \quad (6)$$

RBF kernel is

$$f(x) = \exp \left(-\gamma \cdot \|x - x'\|^2 \right). \quad (7)$$

Sigmoid kernel is

$$f(x) = \tanh (\gamma \cdot x' \cdot x + b). \quad (8)$$

The main influencing factor of recognition performance is the parameters used in SVM. Presently the staple methods to select optimal parameters include grid search [16], genetic algorithm (GA) [17], and particle swarm optimization (PSO) [18] algorithm. In [19], Dorigo et al. proposed ant colony optimization (ACO) algorithm to select optimal parameters value, achieving better classification performance while taking more time. In [20, 21], Xu et al. and Tiwari and Vidyarthi proposed quantum genetic algorithm (QGA) to optimize SVM parameters and verified that quantum operation can increase the scope of the search space and has good searching ability. In [4], Jiang et al. used quantum simulated annealing (QSA) algorithm combined QGA and simulated annealing (SA) algorithm [22] to optimize SVM parameters, tested the classification model based upon pancreatic images, and achieved better and stable accuracy.

2.2. Fruit Fly Optimization Algorithm. Fruit fly optimization algorithm (FOA) [15] is a new method based on fruit fly foraging behavior for global optimization. The flowchart of FOA is shown in Figure 1.

The key steps of FOA are shown as follows.

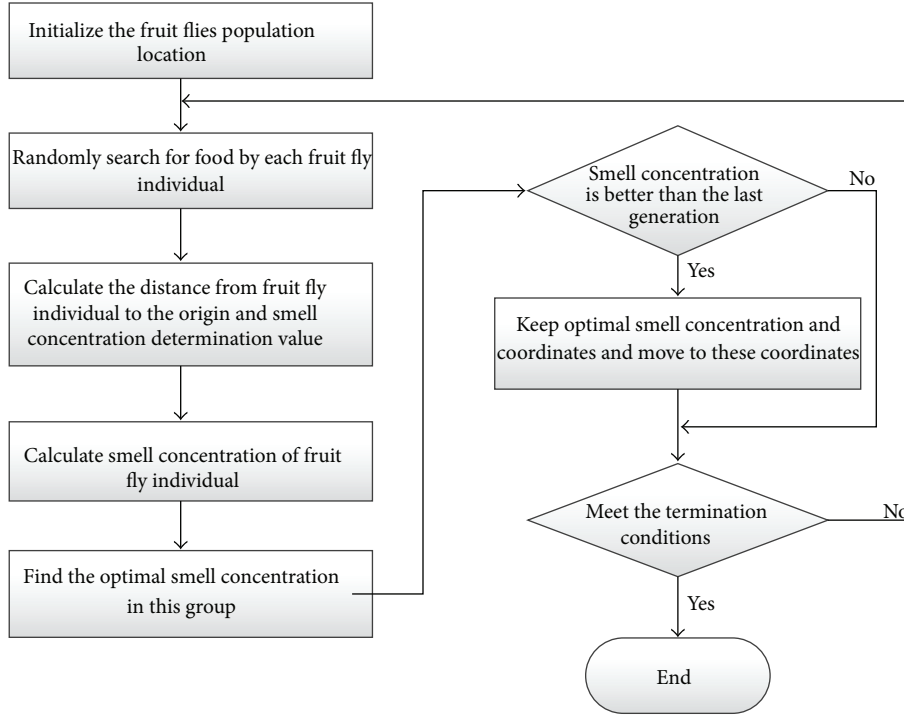


FIGURE 1: The flowchart of FOA.

Step 1. The position of population, (X_axis, Y_axis) , is randomly initialized. X_axis and Y_axis are abscissa value and ordinate value of population's position, respectively.

Step 2. For each fruit fly, the direction and position of flying are randomly evaluated. It is represented as (9). (X_i, Y_i) is the new position of each fruit fly, where $i \in [1, M]$. M is the number of fruit flies in population:

$$\begin{aligned} X_i &= X_axis + \text{Random Value} \\ Y_i &= Y_axis + \text{Random Value.} \end{aligned} \quad (9)$$

Step 3. The distance $(Dist_i)$ from each fruit fly to the origin and the smell concentration determination value (S_i) of each fruit fly are calculated as

$$\begin{aligned} Dist_i &= \sqrt{X_i^2 + Y_i^2} \\ S_i &= \frac{1}{Dist_i}. \end{aligned} \quad (10)$$

Step 4. The smell concentration determination value is used in smell concentration determination function (fitness function) to calculate the smell concentration value as

$$\text{Smell}_i = \text{Function}(S_i). \quad (11)$$

Step 5. The fruit fly which has the best smell concentration is found in population:

$$[\text{bestSmell bestIndex}] = \max(\text{Smell}). \quad (12)$$

Step 6. The best smell concentration and its position (x, y) are saved. The fruit fly population moves to this position by vision.

Step 7. Step 2 to Step 5 are iterated. If the smell concentration is better than previous one, Step 6 is executed.

FOA is one of the intelligent optimization algorithms. It is easy to set up, easy to implement, and fast to optimize. But it also has some problems. In the phase of parameter initialization, FOA uses randomized strategy to determine initial point position. In the phase of fruit fly individual position update, blind search strategy is used. It is slow to converge and easy to fall into extreme values. At present, there are a number of evaluation criteria for classifier performance. In classifier optimization algorithms, classification accuracy and error rate are always used as the fitness function. But those criteria cannot reflect clinical prior knowledge. It is simply to evaluate an operating point and not strong enough when the distribution of class is changed.

3. Methodology

The whole procedure of the proposed method is shown in Figure 2. The detailed process of the proposed method is as follows.

(1) *Feature Extraction.* We extract gray and fractal dimension features from the segmented pancreatic images, and then we normalize those features.

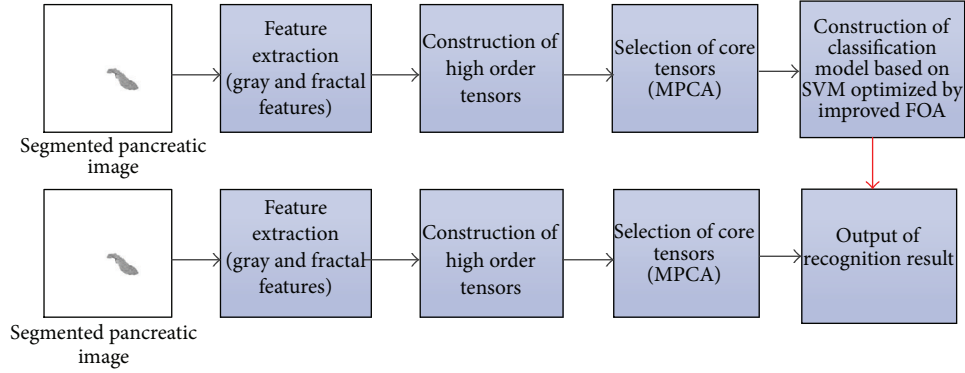


FIGURE 2: The flowchart of the proposed method.

(2) *High Order Tensors Construction*. High order tensors are constructed based on the extracted features to represent pancreatic images.

(3) *Feature Selection*. In this paper we use the MPCA method to extract the eigen tensors for classification.

(4) *Pancreatic Cancer Classification*. After we obtain the eigen tensors by MPCA, we can treat the eigen tensors as input samples, and then we use the approach of SVM optimized by improved FOA to train classification model of pancreatic cancer image.

In the process, high order tensors construction and feature selection are carried out in accordance with [4]. So in this paper, we will no longer discuss them.

3.1. *Improved FOA*. Aiming at the existing problem of FOA, we introduce quantum to FOA and redefine a new fitness function as the smell concentration determination function.

3.1.1. *Quantum Fruit Fly Coding*. In improved FOA (IFOA), quantum phase is used to code fruit flies' position. Compared with FOA which has the same number of fruit flies, the solution search space of quantum fruit flies is the double of the original fruit flies. The quantum fruit flies population position is shown as (13). When initializing, the quantum bit phase angle is $\theta = \pi \cdot (2 \cdot \text{rand} - 1)$, where $\text{rand} \in [0, 1]$, $j \in [1, n]$, and n is the dimension of optimization problem. In this paper, $n = 2$:

$$\begin{aligned} X_axis &= [\theta_{x1}, \theta_{x2}, \dots, \theta_{xj}] \\ Y_axis &= [\theta_{y1}, \theta_{y2}, \dots, \theta_{yj}]. \end{aligned} \quad (13)$$

3.1.2. *Quantum Fruit Fly Smell Concentration Determination Value*. As quantum phase is used to code fruit flies' position, each fruit fly has two solutions, namely, the cosine solution and the sine solution. The distance ($\text{Dist}_j^{(i)}$) from the i th fruit

fly to the origin and the smell concentration determination value ($S_j^{(i)}$) of the i th fruit fly can be calculated as

$$\text{Dist}_j^{(i)} = \begin{bmatrix} \sqrt{\cos^2 \theta_{xj}^{(i)} + \cos^2 \theta_{yj}^{(i)}} \\ \sqrt{\sin^2 \theta_{xj}^{(i)} + \sin^2 \theta_{yj}^{(i)}} \end{bmatrix} \quad (14)$$

$$S_j^{(i)} = \frac{\text{Dist}_j^{(i)}}{\sqrt{2}}, \quad (15)$$

where $i \in [1, M]$, and M is the number of the fruit flies' populations. In (15), Dist is normalized to $[0, 1]$ and then assigned to S . The reason is to facilitate parameters zooming for optimizing SVM.

3.1.3. *Quantum Fruit Fly Smell Concentration Determination Function*. False negative rate (FNR) is known as the rate of missed diagnosis. It is the percentage of actual sickness while identified as disease-free. FNR is complementary with the actual diagnostic sensitivity. False positive rate (FPR) is known as the misdiagnosis rate. It is the percentage of the actual disease-free while identified as sickness. FPR is complementary with the actual diagnostic specificity. In the process of actual disease diagnosis, if diagnosis with high sensitivity is used, the higher is the sensitivity, the less is the rate of missed diagnosis. That is to say, FNR is low. When diagnosis with high specificity is used, the misdiagnosis rate is low. That is to say, FPR is low. Therefore, in improved FOA, the mean of weighted sum of FNR and FPR in k -fold cross-validation is used as the smell concentration determination function. It is shown as

$$\text{Fitness} = \frac{1}{K} \sum_{k=1}^K [w \cdot \text{FNR} + (1 - w) \cdot \text{FPR}]. \quad (16)$$

In (16), K is the parameter of k -fold cross-validation, and w is the weight of FNR. If a fruit fly has small smell concentration value, it is good.

3.1.4. *Quantum Fruit Fly Mutation Operation*. Quantum *not* gate is used to randomly change quantum fruit flies' positions.

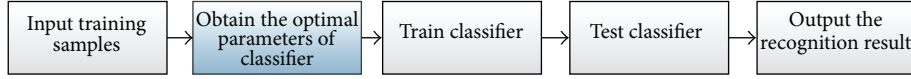


FIGURE 3: The framework of classifier construction.

It not only increases the diversity of the population, but also avoids precocity. The quantum *not* gate based on phase coding is shown as

$$\theta_j^{(i)} = \frac{\pi}{2} - \theta_j^{(i)}. \quad (17)$$

The mutation probability of an individual fruit fly is P_m . If P_m is greater than a random number within (0, 1), the two probability amplitudes of X-coordinate or Y-coordinate of the individual fruit fly randomly selected will be exchanged by quantum *not* gate.

The acceptance probability of mutated new fruit fly position obeys the Boltzmann probability distribution. It is shown as

$$P(\theta^{(i)*} \Rightarrow \theta^{(i)}) = \begin{cases} 1, & F(\theta^{(i)*}) < F(\theta^{(i)}) \\ P_i, & F(\theta^{(i)*}) \geq F(\theta^{(i)}) \end{cases} \quad (18)$$

$$P_i = \left(1 + \exp \left[\frac{F(\theta^{(i)}) - F(\theta^{(i)*})}{l} \right] \right)^{-1}.$$

In (18), $\theta^{(i)*}$ is the mutated new fruit fly position, $\theta^{(i)}$ is the original fruit fly position, $F(\cdot)$ is the smell concentration determination function, and l is iterations. If $F(\theta^{(i)*}) < F(\theta^{(i)})$, the new position will be accepted by probability 1. Otherwise, the new position will be accepted by probability P_i .

3.2. Construction of SVM Classifier Based on Improved FOA. The framework of classifier construction and the flowchart of SVM parameter optimization based on improved FOA are shown in Figures 3 and 4, respectively. The process of classifier construction consists 3 steps, namely, obtaining classifier parameters, training classifier, and testing classifier.

The parameters of SVM, penalty factor C , and RBF kernel function parameter γ have great influence on the performance of classifier. C determines the promotion ability of SVM. The small value of C represents the penalty of empirical error being small, which can lead to “underfitting study.” The large value of C represents the penalty of empirical error being large, which can lead to “overfitting study.” The optimal value of C is different according to different data subspace, and selecting the optimal value of C can make the promotion ability better. SVM can map the input data of low dimensional space into high dimensional space by the kernel function. Vapnik [14] has found that the parameters of kernel function and penalty factor C have great influence on the performance of SVM. So the selection of parameters of penalty factor C and RBF kernel function parameter γ is important.

The detailed process for optimizing SVM parameters, penalty factor C , and RBF kernel function parameter γ is as follows.

Step 1. The population position (X_axis, Y_axis) is initialized by (13).

Step 2. For each fruit fly, the position and the direction of flying are randomly evaluated. It is shown as

$$X_i = X_axis + V_x^{(i)} \quad (19)$$

$$Y_i = Y_axis + V_y^{(i)}.$$

In (19), $V \in [-1, 1]$, $i \in [1, M]$, and M is the number of individuals in population.

Step 3. The distance from each fruit fly to the origin and the smell concentration determination value is calculated as (14) and (15).

Step 4. The smell concentration determination value is zoomed to get C and γ . It is shown as (20). C_m and g_m are zoom multiples of C and γ , which can be obtained by prior knowledge:

$$C = C_m \cdot S_1^{(i)} \quad (20)$$

$$\gamma = g_m \cdot S_2^{(i)}.$$

Step 5. The smell concentration is calculated by (16). We set $K = 5$ and will discuss the value of w in the next section.

Step 6. If the individual fruit fly meets the mutation condition, the mutation operation will be done as (17) and (18).

Step 7. The fruit fly which has the best smell concentration is found as (21). $bestSmell$ is the best smell concentration, $bestIndex$ is the individual fruit fly which has the best smell concentration, and $bestPos$ is the position of best smell concentration of the individual fruit fly:

$$\begin{bmatrix} bestSmell \\ bestIndex \\ bestPos \end{bmatrix} = \min(Smell). \quad (21)$$

Step 8. The axes and position of the best smell concentration are saved. The fruit fly population moves to this position by vision. It is shown as

$$SmellBest = bestSmell \quad (22)$$

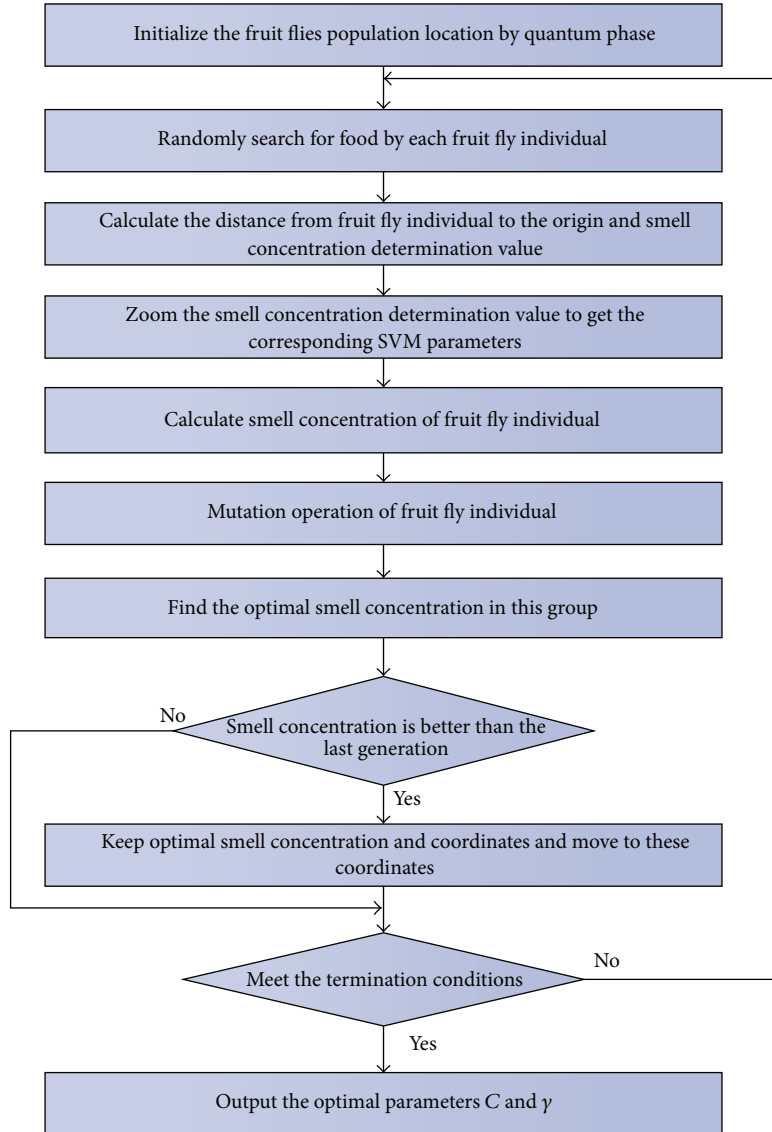


FIGURE 4: The flowchart of SVM parameter optimization based on improved FOA.

$$\text{Posbest} = \text{bestPos} \quad (23)$$

$$X_axis = X(\text{bestIndex}) \quad (24)$$

$$Y_axis = Y(\text{bestIndex}). \quad (25)$$

Step 9. Step 2~Step 7 are iterated. If the smell concentration is better than previous one, Step 8 is executed. If the termination condition is satisfied, the optimum parameters will be returned.

4. Results and Discussion

4.1. Experimental Data. In this paper, abdominal CT images are used in experiments, which are provided by the radiology department of a hospital in Shenyang, China. Their resolution is 512×512 pixels, the scan slice thickness is 2 mm, and the format is DICOM. For the purpose of algorithm simulation,

TABLE 1: Experimental data.

	Training samples	Testing samples	Total
Pancreatic cancer images (positive)	17	16	33
Normal images (negative)	41	40	81
Total	58	56	

the DICOM image is transformed into BMP image. The grayscale is 256 and the resolution is 128×128 . The detailed information of dataset is shown in Table 1.

4.2. Evaluation Criteria. According to the hybrid matrix, which is shown in Table 2, the evaluation criteria are calculated. In this paper, evaluation criteria consist of False Positive

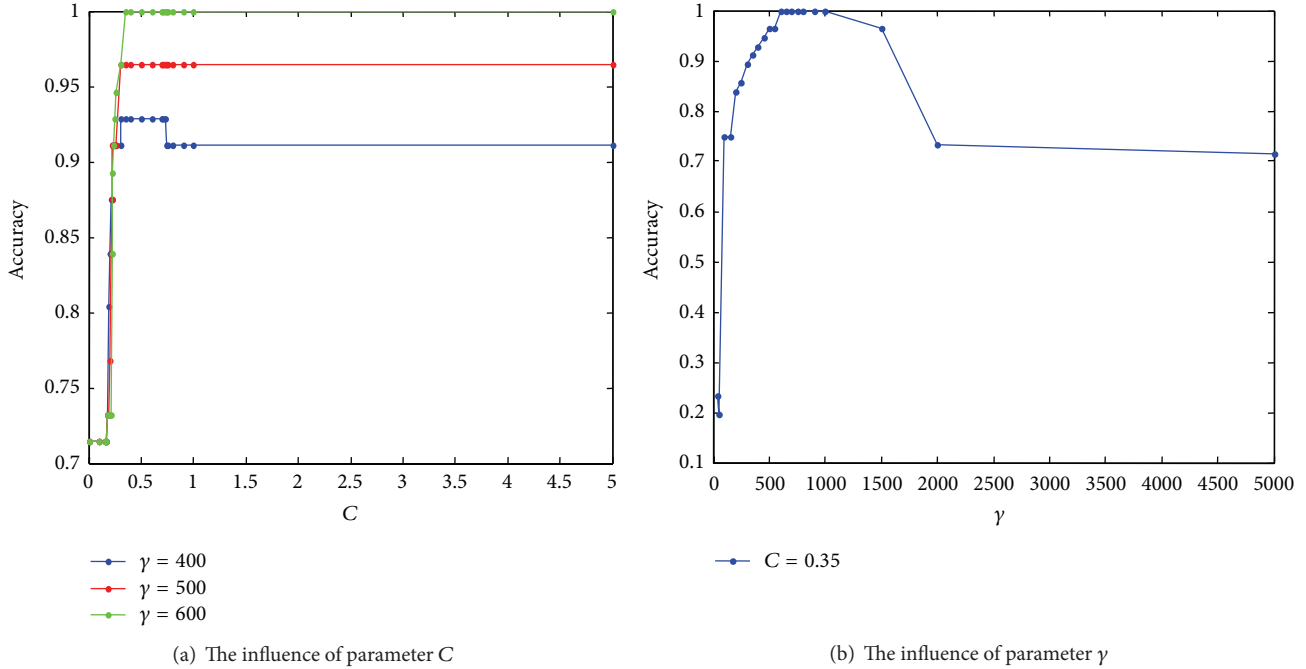


FIGURE 5: The effect of parameters C and γ on the classification accuracy.

TABLE 2: Hybrid matrix.

	Predicted positive example (P')	Predicted negative example (N')
Practical positive example (P)	True positive example (TP)	False negative example (FN)
Practical negative example (N)	False positive example (FP)	True negative example (TN)

Rate (FPR), False Negative Rate (FNR), Accuracy, Precision, F1 value, and the running time of the algorithms. The mean square errors of evaluation criteria in many experiments are also used to evaluate the stability of the algorithm:

$$\begin{aligned}
 \text{FPR} &= \frac{\text{FP}}{\text{N}} \\
 \text{FNR} &= \frac{\text{FN}}{\text{P}} \\
 \text{Accuracy} &= \frac{\text{TP} + \text{TN}}{\text{P} + \text{N}} \\
 \text{Precision} &= \frac{\text{TP}}{\text{P}'} \\
 \text{F1} &= \frac{2}{1/\text{Precision} + 1/\text{Recall}}
 \end{aligned}
 \tag{26}$$

4.3. *Prior Knowledge.* Because of the sensitivity of initial scope for parameters optimization, we make the statistical analysis for the penalty factor C and RBF kernel function

parameter γ, which obtains the prior knowledge of the parameters. The result is shown as in Figure 5.

From Figure 5, we can obtain the initial scope of C and γ by QSA that is [0.1, 1] and [50, 2000], respectively. And the scope of optimal solution is [0.3, 0.5] and [500, 1500], respectively. The scaling of C and γ is 1 and 2000, respectively.

4.4. *Determination of FNR Weight.* In an actual treatment, a patient was ill, but he was diagnosed as disease-free; then the treatment progress would be delayed and the cure opportunity would be reduced. On the contrary, if a patient was disease-free and was diagnosed with illness, patient would undergo further examination to make up the mistake. Therefore, it is believed that FNR is more important than FPR. The weight of FNR should be greater than FPR; that is to say, $0.5 \leq w \leq 1$.

For different values of w, the experiment was run 10 times. Then we compared the mean value of evaluation criterions and their mean square error.

Figures 6 and 7 show the states of optimized parameters C and γ using different w. Figure 8 illustrates the states of FNR and FPR. Figure 9 demonstrates the states of Accuracy, Precision, and F1.

From Figures 6 and 7, it can be seen that the optimized parameters are found to be in line with prior knowledge. According to the mean square error of C, when w = 0.8, the algorithm is the most stable, and w = 0.9 is the second. According to the mean square error of γ, when w = 0.7, the algorithm is the most stable, and w = 0.9 is the second.

The better performance of the algorithm comes when the values of FNR and FPR are smaller. From Figure 8, for FPR, when w is 0.9, its mean value and mean square error are the

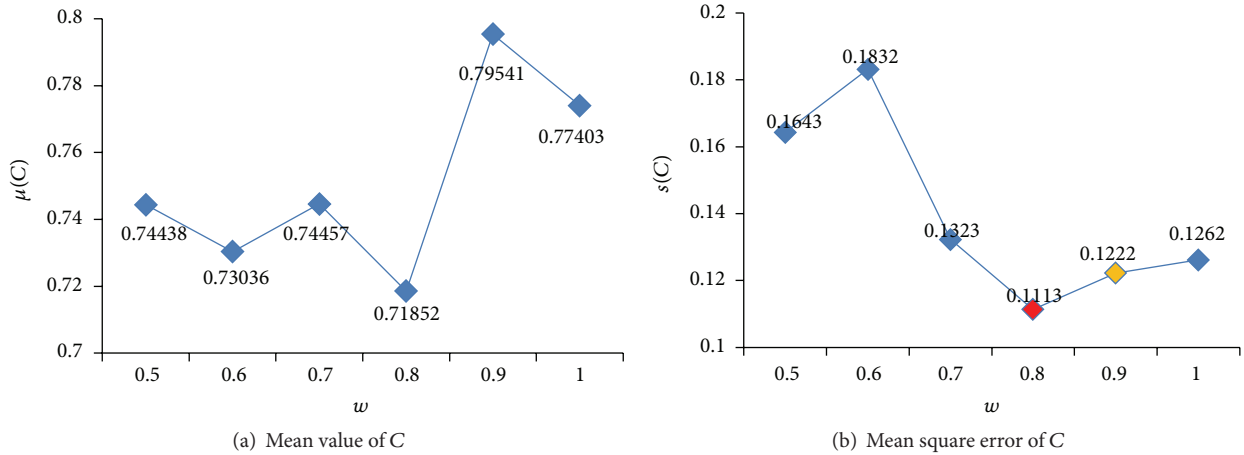


FIGURE 6: The status of parameter C for different w.

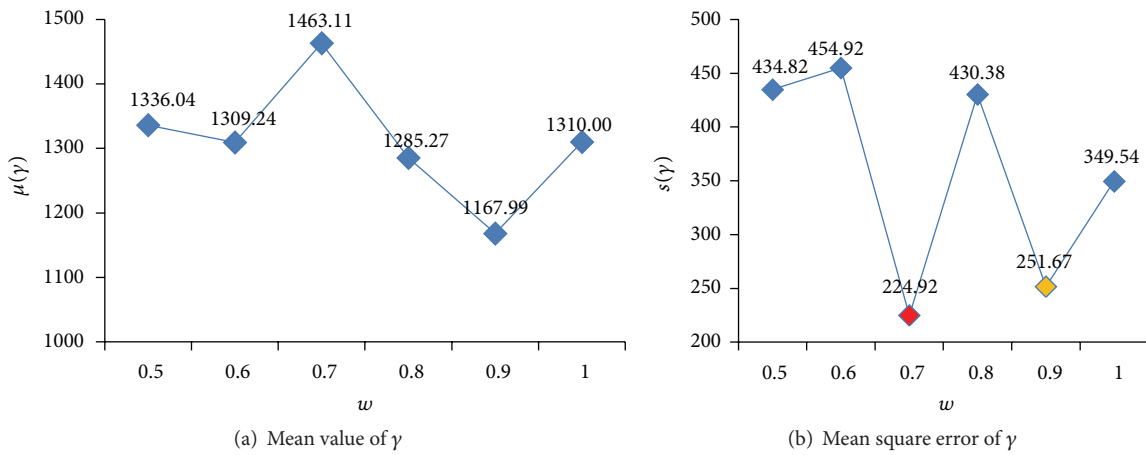


FIGURE 7: The status of parameter γ for different w.

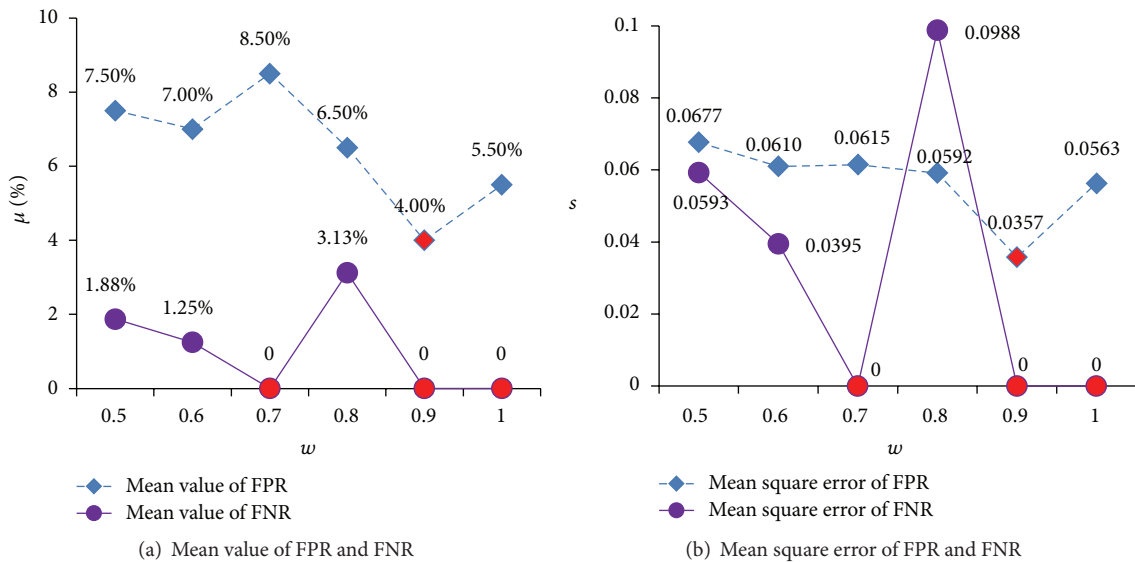


FIGURE 8: FPR and FNR status for different w.

TABLE 3: Experimental results of IFOA-SVM.

	C	γ	FPR	FNR	Accuracy	Precision	F1	Time (s)
1	0.8666	878.79	0	0	100.00%	100.00%	100.00%	31.78
2	0.7089	1003.4	2.50%	0	98.21%	94.12%	96.97%	31.64
3	0.7015	952.7641	2.50%	0	98.21%	94.12%	96.97%	31.35
4	0.8605	1401.3	5.00%	0	96.43%	88.89%	94.12%	30.49
5	0.9922	1314	5.00%	0	96.43%	88.89%	94.12%	31.09
6	0.8319	1364.5	5.00%	0	96.43%	88.89%	94.12%	31.28
7	0.6345	1050.6	2.50%	0	98.21%	94.12%	96.97%	30.7
8	0.9405	820.0642	0	0	100.00%	100.00%	100.00%	30.9
9	0.7671	1384.7	5.00%	0	96.43%	88.89%	94.12%	31.06
10	0.6504	1509.8	12.50%	0	91.07%	76.19%	86.49%	31.68
μ	0.79541	1167.99183	4.00%	0	97.14%	91.41%	95.39%	31.197
s	0.1222	251.6716	0.0357	0	0.0255	0.0688	0.0388	

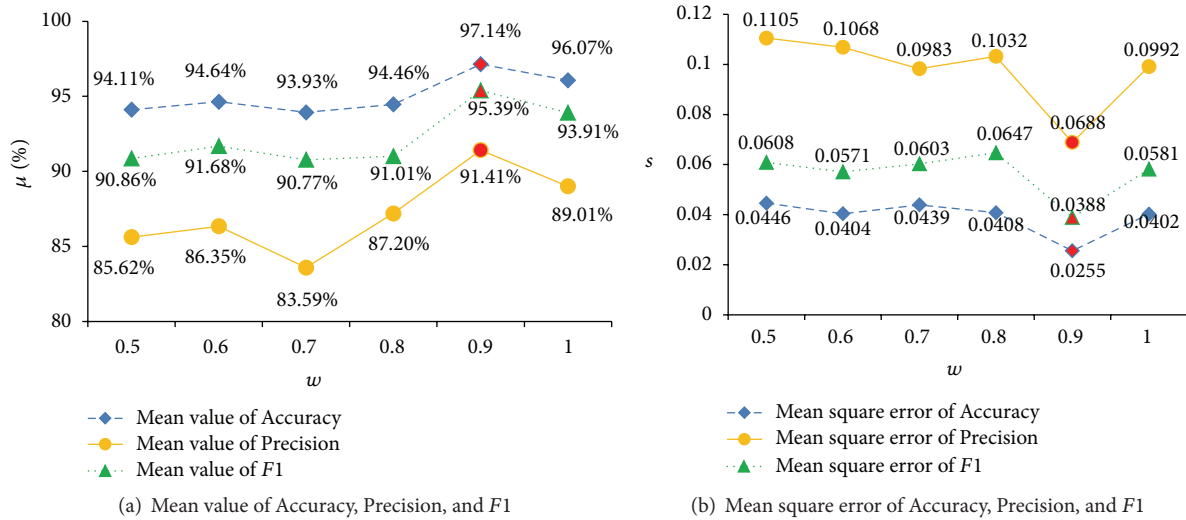


FIGURE 9: Accuracy, Precision, and F1 status for different w .

smallest; for FNR, when w is 0.7, 0.9, or 1, its mean value and mean square error are the smallest. Greater values of Accuracy, Precision, and F1 can lead to better performance of the algorithm. From Figure 8, when $w = 0.9$, its mean values of Accuracy, Precision, and F1 are the greatest, and the mean square error is the smallest. Therefore, the final value of w is determined as 0.9.

4.5. Experimental Results and Analysis. Ten experiments of SVM optimized by improved FOA (IFOA-SVM) are randomly done. The experimental result is shown in Table 3.

From Table 3, it is known that the mean values of C and γ are 0.79541 and 1167.99183, respectively. The average of FPR is 4%, FNR is 0, Accuracy is 97.14%, Precision is 91.41%, F1 is 95.39%, and the time is 31.197 s.

Compared with other classifiers, the performance of IFOA-SVM is better as shown in Figures 10 and 11. In Figure 12, the comparison of running time is shown. Classifier Fisher is the Fisher linear classifier, classifier BPNN

is the BP neural network, SVM is the common SVM, and ACO-SVM, FOA-SVM, and QSA-SVM are the optimized classifier SVM using ant colony algorithm, fruit fly optimal algorithm, and quantum simulated annealing, respectively. IFOA-SVM is the proposed method.

In Figure 10, FNR achieved 100% and FPR is 0 by SVM and ACO-SVM; that is to say, all patients are diagnosed free from diseases. This situation is not allowed in actual diagnosis. FNR of BPNN and Fisher are 88.75% and 56.25%, and FPR are 60% and 49.5%. So BPNN and Fisher lack credibility. FPR and FNR of FOA-SVM are 0 and 35%. It sometimes occurs in missed diagnosis situation. FPR and FNR of QSA-SVM are 6% and 5%. It might occur in missed diagnosis to few patients. FPR and FNR of the proposed IFOA-SVM are 4% and 0. It is better than other methods. IFOA-SVM achieves the best sensibility and stability.

In Figure 11, for the average of Precision, FOA-SVM is the best, which is 100%. IFOA-SVM takes the second place, which is 91.41%. And from Figure 11(b), FOA-SVM is the most stable. For the average of F1, the proposed IFOA-SVM

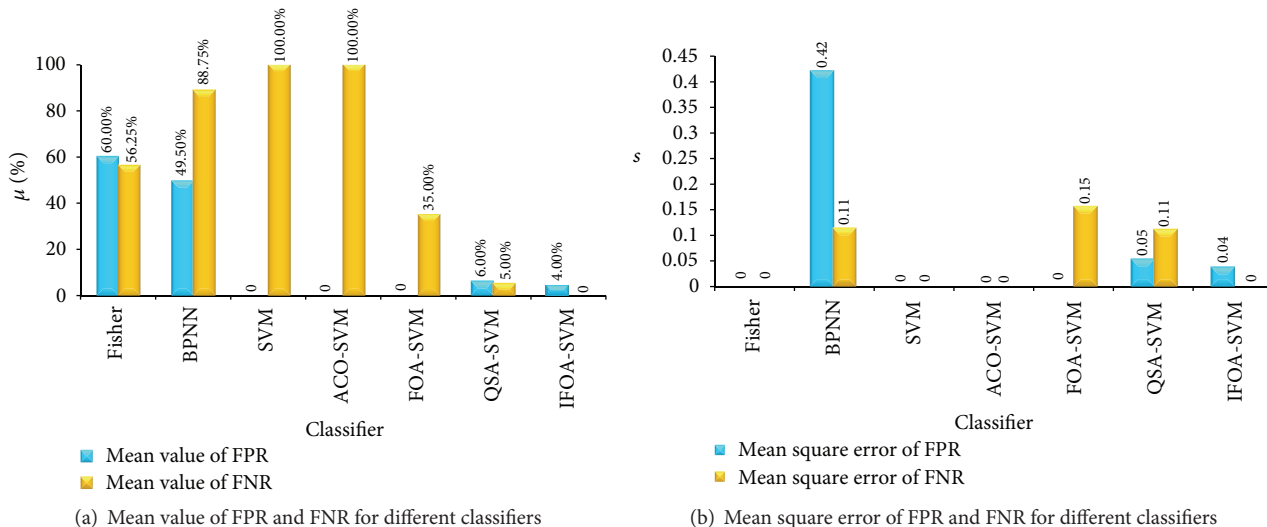


FIGURE 10: FPR and FNR status for different classifiers.

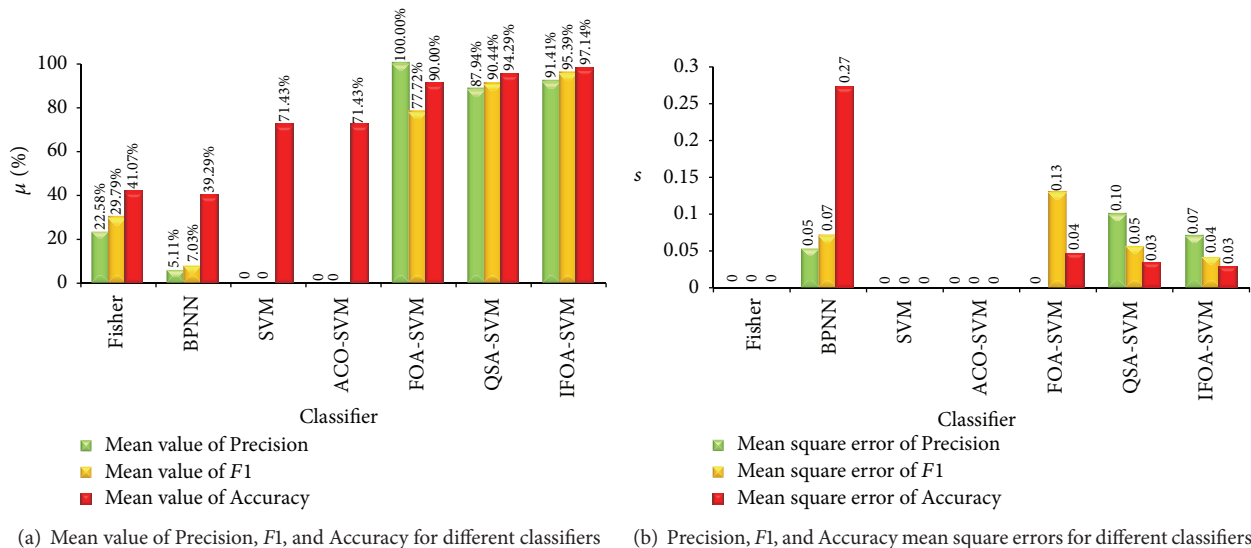


FIGURE 11: Accuracy, Precision, and F1 status for different classifiers.

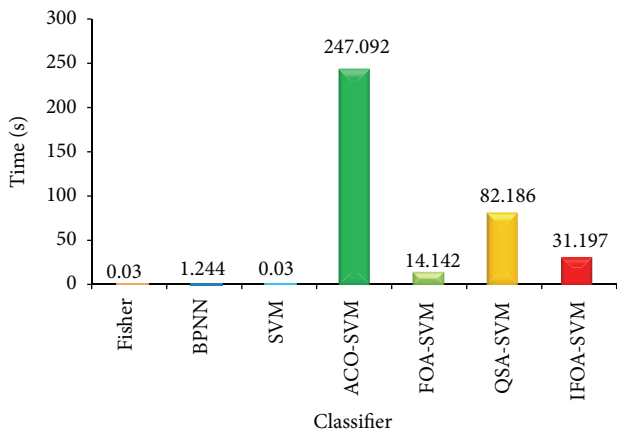


FIGURE 12: Running time for different classifiers.

achieves 95.39%, which is optimal. QSA-SVM is 90.44%, which takes the second place. But IFOA-SVM is more stable than QSA-SVM. For the average of Accuracy, IFOA-SVM is 97.14%, which is the best. QSA-SVM is 94.29%. FOA-SVM is 90%. SVM and ACO-SVM are 71.43%. Fisher and BPNN are less than 50%. Compared with mean square error of Accuracy, IFOA-SVM is the most stable.

In Figure 12, Fisher and SVM cost the least time, which is 0.03 s. BPNN is 1.244 s. FOA-SVM is 14.142 s. IFOA-SVM is 31.197 s. QSA-SVM is 82.186 s. ACO-SVM cost the most time, which is 247.092 s. In actual diagnosis, less time is better. The proposed IFOA-SVM is not the best but is not out of the way.

ACO-SVM, FOA-SVM, QSA-SVM, and the proposed IFOA-SVM can be used to optimize SVM parameters. In Figures 13 and 14, comparative results of mean value and

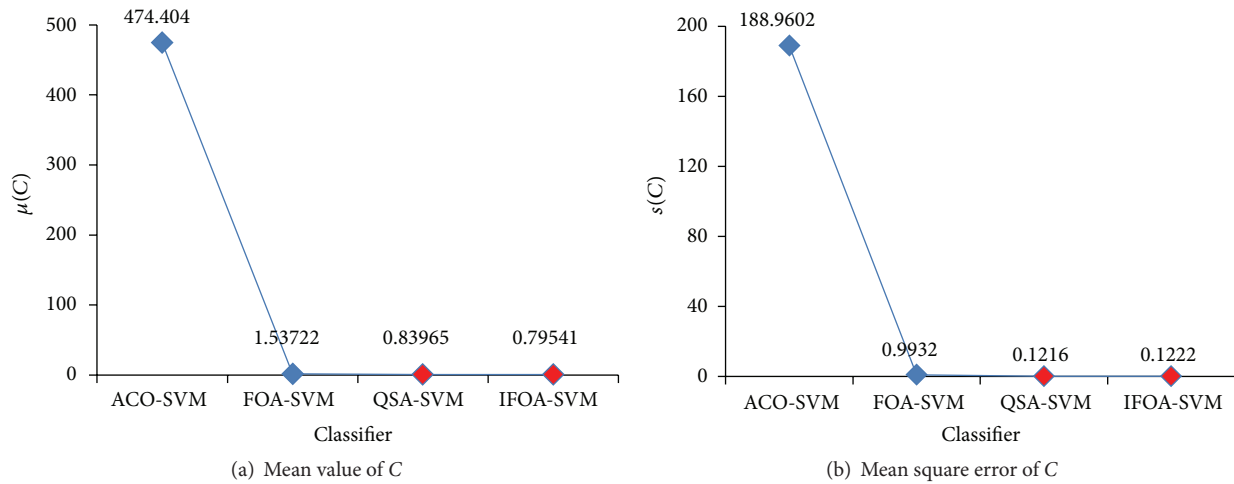


FIGURE 13: Parameter C status for different method.

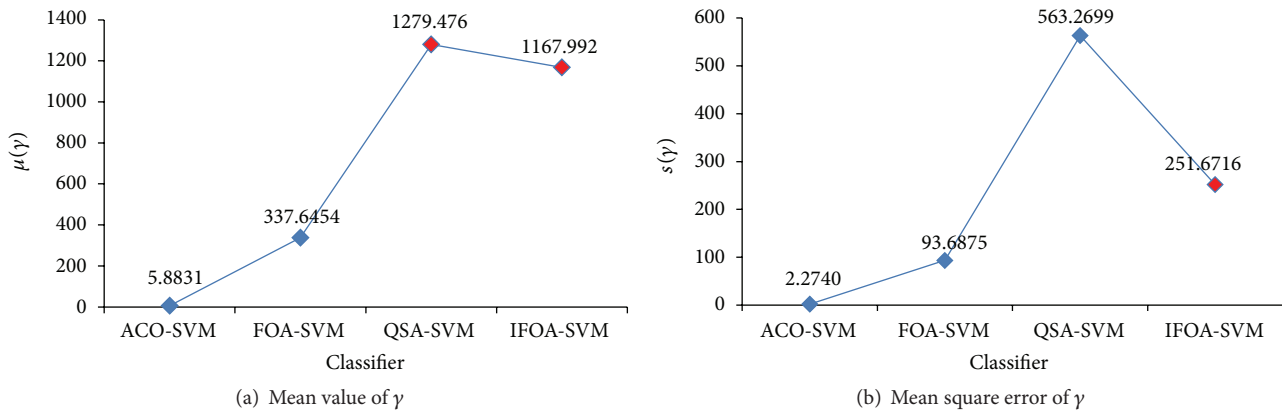


FIGURE 14: Parameter γ status for different method.

mean square error of optimal parameters, C and γ , are shown based on those methods.

In the pancreatic cancer classifier based on ACO-SVM, C is oversized and γ is undersized. By using FOA-SVM, optimal parameters are not in estimation interval of prior knowledge, but in terms of mean square error of optimal parameters FOA is stable. When QSA-SVM and IFOA-SVM are used to optimize SVM parameters, optimal parameters are in estimation interval, and the stability of two methods is similar from mean square error of C . And in terms of mean square error of γ , IFOA-SVM is more stable than QSA-SVM.

5. Conclusion

In this paper, we introduced the concept of quantum to FOA to improve it. A new smell concentration determination function was defined in the improved FOA. The improved FOA was used to optimize the parameters of SVM and a classifier was constructed based on the optimized SVM. As an application, pancreatic cancer classifier was established. The

proposed method achieved better classification performance. The first reason is that quantum coding and quantum operation increased the diversity of the population and avoided precocity. The second reason is that the redefined smell concentration determination function was more suitable to the actual diagnosis requirements. The third reason is the advantages of FOA which are easy to set up, easy to implement, and fast to optimize. Therefore, the proposed method can improve the classification performance of pancreatic cancer images and then assist doctors in diagnosing diseases.

Conflict of Interests

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

Acknowledgment

The research is supported by the National Natural Science Foundation of China (no. 61272176).

References

- [1] A. Jemal, R. Siegel, E. Ward, Y. Hao, J. Xu, and M. J. Thun, "Cancer statistics, 2009," *CA: Cancer Journal for Clinicians*, vol. 59, no. 4, pp. 225–249, 2009.
- [2] Y. Wang and P. Zhao, "Advances in early diagnosis of pancreatic cancer," *Oncology Progress*, vol. 4, no. 4, pp. 327–332, 2006.
- [3] F. Fraioli, G. Serra, and R. Passariello, "CAD (computed-aided detection) and CADx (computer aided diagnosis) systems in identifying and characterising lung nodules on chest CT: overview of research, developments and new prospects," *Radiologia Medica*, vol. 115, no. 3, pp. 385–402, 2010.
- [4] H. Jiang, D. Zhao, T. Feng, S. Liao, and Y. Chen, "Construction of classifier based on MPCA and QSA and its application on classification of pancreatic diseases," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 713174, 7 pages, 2013.
- [5] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "MPCA: multilinear principal component analysis of tensor objects," *IEEE Transactions on Neural Networks*, vol. 19, no. 1, pp. 18–39, 2008.
- [6] B. Kovalerchuk, E. Triantaphyllou, J. F. Ruiz, and J. Clayton, "Fuzzy logic in computer-aided breast cancer diagnosis: analysis of lobulation," *Artificial Intelligence in Medicine*, vol. 11, no. 1, pp. 75–85, 1997.
- [7] P. C. Pendharkar, J. A. Rodger, G. J. Yaverbaum, N. Herman, and M. Benner, "Association, statistical, mathematical and neural approaches for mining breast cancer patterns," *Expert Systems with Applications*, vol. 17, no. 3, pp. 223–232, 1999.
- [8] M. Antonie, O. Zaiane, and A. Coman, "Application of data mining techniques for medical image classification," in *Knowledge Discovery and Data Mining*, pp. 94–101, 2001.
- [9] J. Zhang, Y. Wang, Y. Dong, and Y. Wang, "Ultrasonographic feature selection and pattern classification for cervical lymph nodes using support vector machines," *Computer Methods and Programs in Biomedicine*, vol. 88, no. 1, pp. 75–84, 2007.
- [10] J. Ramírez, R. Chaves, J. Górriz et al., "Functional brain image classification techniques for early Alzheimer disease diagnosis," in *Bioinspired Applications in Artificial and Natural Computation: Third International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2009, Santiago de Compostela, Spain, June 22–26, 2009, Proceedings, Part II*, vol. 5602 of *Lecture Notes in Computer Science*, pp. 150–157, Springer, Berlin, Germany, 2009.
- [11] D. Tsai and K. Kojima, "Enhancement of CT pancreatic features by a simple cascading filter," in *Proceedings of the Nuclear Science Symposium and Medical Imaging Conference, 1993 IEEE Conference Record*, San Francisco, Calif, USA, 1993.
- [12] T. Takada, H. Yasuda, K. Uchiyama, H. Hasegawa, T. Iwagaki, and Y. Yamakawa, "A proposed new pancreatic classification system according to segments: operative procedure for a medial pancreatic segmentectomy," *Journal of Hepato-Biliary-Pancreatic Surgery*, vol. 1, no. 3, pp. 322–325, 1994.
- [13] S. He, H. J. Cooper, D. G. Ward, X. Yao, and J. K. Heath, "Analysis of premalignant pancreatic cancer mass spectrometry data for biomarker selection using a group search optimizer," *Transactions of the Institute of Measurement and Control*, vol. 34, no. 6, pp. 668–676, 2012.
- [14] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.
- [15] 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.
- [16] T. Wang, X. Ye, L. Wang, and H. Li, "Grid search optimized SVM method for dish-like underwater robot attitude prediction," in *Proceedings of the 5th International Joint Conference on Computational Sciences and Optimization (CSO '12)*, pp. 839–843, June 2012.
- [17] M. Zhao, C. Fu, L. Ji, K. Tang, and M. Zhou, "Feature selection and parameter optimization for support vector machines: a new approach based on genetic algorithm with feature chromosomes," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5197–5204, 2011.
- [18] F.-H. Yu and H.-B. Liu, "Structural damage identification by support vector machine and particle swarm algorithm," *Journal of Jilin University Engineering and Technology Edition*, vol. 38, no. 2, pp. 434–438, 2008.
- [19] M. Dorigo, V. Maniezzo, and A. Coloni, "Ant system: optimization by a colony of cooperating agents," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 26, no. 1, pp. 29–41, 1996.
- [20] X. Xu, J. Jiang, J. Jie, H. Wang, and W. Wang, "An improved real coded quantum genetic algorithm and its applications," in *Proceedings of the International Conference on Computational Aspects of Social Networks (CASoN '10)*, pp. 307–310, IEEE, Taiyuan, China, September 2010.
- [21] P. K. Tiwari and D. P. Vidyarthi, "A variant of quantum genetic algorithm and its possible applications," *Advances in Intelligent and Soft Computing*, vol. 130, no. 1, pp. 797–811, 2012.
- [22] S. Xu, H. Zhao, and Y. Xie, "Grey SVM with simulated annealing algorithms in patent application filings forecasting," in *Proceedings of the International Conference on Computational Intelligence and Security Workshops (CIS '07)*, pp. 850–853, IEEE, Harbin, China, December 2007.