

Digital image technology based on PCA and SVM for detection and recognition of foreign bodies in lyophilized powder

Zhihong Zhang^a, Yingchun Luo^{b,*}, Xudong Deng^a and Weinan Luo^a

^a*College of Computer Engineering and Applied Mathematics, Changsha University, Changsha, Hunan, 410200, China*

^b*Department of Ultrasound, The Maternal and Child Health Care Hospital of Hunan Province, Changsha, Hunan, 410008, China*

Abstract.

BACKGROUND: Digital image technology has made great progress in the field of foreign body detection and classification, which is of great help to drug purity extraction and impurity analysis and classification.

OBJECTIVE: The detection and classification of foreign bodies in lyophilized powder are important. The method which can obtain a higher accuracy of recognition needs to be proposed.

METHODS: We used digital image technology to detect and classify foreign bodies in lyophilized powder, and studied the process of image preprocessing, median filtering, Wiener filtering and average filtering balance to better detect and classify foreign bodies in lyophilized powder.

RESULTS: Through industrial small sample data simulation, test results show that in the process of image preprocessing, 3×3 median filtering is best. In the aspect of foreign body recognition, the recognition based on principal component analysis (PCA) and support vector machine (SVM) algorithm and the recognition based on PCA and Third-Nearest Neighbor classification algorithm are compared and results show that the PCA+SVM algorithm is better.

CONCLUSION: We demonstrated that integrating PCA and SVM to classify foreign bodies in lyophilized powder. Our proposed method is effective for the prediction of essential proteins.

Keywords: Foreign bodies, median filtering, PCA, SVM

1. Introduction

Lyophilized powder for injection freezes the original liquid drug solution into a solid state in a sterile environment, and then sublimates and dries the water contained in the solid state through vacuum extraction, and finally forms a sterile powder injection. Although lyophilized powder for injection has the advantages of long-term preservation of products, there are often some visible foreign matters (fiber, hair, glass debris, etc.) in injection products. Because of this, the safety and quality of injection products are largely affected, which is a great hidden danger for the medical and health industry. At present, in order to reduce the economic losses and negative image impact caused by “drug recall”, most domestic

*Corresponding author: Yingchun Luo, Department of Ultrasound, The Maternal and Child Health Care Hospital of Hunan Province, Changsha, Hunan, 410008, China. Tel.: +86 13875811747; E-mail: 807859088@qq.com.

pharmaceutical enterprises conduct manual testing in dark rooms. Practice has proved that this method has many disadvantages: low efficiency, poor quality, workers prone to visual fatigue. This method not only consumes a lot of labor, but also increases the production cost. In comparison, machine vision detection has the characteristics of long-term continuous, fatigue free and rapid work. The application of this technology to foreign matter inspection in medical products not only improves the detection and identification speed, but also reduces the production cost, which has a high practical value. For the application of foreign matter inspection and recognition technology in medical products, there have been many studies abroad. Researchers who have successfully translated taxonomic identification techniques into medical products include Brevetti C.E.A, Seidenader and Eisai. Famous scholars include Ishii [1], Schimizu and Kato [2,3]. However, the research on classification recognition algorithm is not mature, and the application of its research results in the actual production line cannot achieve the desired effect. With the development of computer technology [4,5], domestic studies on the detection and classification of foreign matter inspection in pharmaceutical liquids have also made some achievements. Xiao et al. [6] mainly carried out studies on detection and tracking of moving foreign matter in ampoule bottle solutions. The image is obtained by means of rotating the emergency stop, the moving target is tracked by means of particle filter combined with template matching, and the target particle is extracted by difference method. Ge et al. [7–9] also carried out relevant studies. They first used sequence image difference method to extract the region of interest, and then used the trained pulse-coupled neural network to effectively identify the characteristics of foreign matter in oral liquid medicine. Now, there are few studies on the classification and recognition of foreign matter of defective lyophilized powder for injection by machine learning. In order to realize intelligent classification and identification of defective products in lyophilized powder for injection, we did some research on this. Filtering denoising methods were studied in the preprocessing of foreign matter images of lyophilized powder for injection. Fast PCA algorithm was selected for feature extraction of lyophilized powder for injection foreign matter image, and third-order neighborhood algorithm and SVM algorithm were selected for classification and recognition of lyophilized powder for injection foreign matter image.

2. Lyophilized powder foreign matter image preprocessing

In the early stage of detection and classification of the foreign matter image of lyophilized powder, some preprocessing of the foreign matter image was carried out, because there were some gaps in the light direction and light intensity of the foreign matter image of small samples collected by industry, which would cause great interference to the experimental results.

Median filter [10] is a typical and well-known nonlinear signal processing algorithm, and its implementation mainly relies on efficient sorting algorithm. The core idea of this algorithm is to replace the original pixel value with the median value of each pixel point field. Experimental results show that this method can effectively eliminate isolated noise points. At the same time, under certain conditions, the median filter can avoid the problem of fuzzy image details. For the median filter of different dimensions, the implementation method is also different. The following mainly introduces the core steps of one-dimensional median filter and two-dimensional median filter. One-dimensional median filter is realized by Eq. (1) to calculate the median value of the pixel Y_k :

$$Y_k = \text{med}\{x_{k-N}, x_{k-N+1}, \dots, x_k, \dots, x_{k+N}\} \quad (1)$$

where. x_k is the value of k th pixel, N is the one-dimensional window size.

Table 1
Pixel arrangement of window 3×3

	Row zero	First row	Second row
Line zero	P_0	P_1	P_2
First line	P_3	P_4	P_5
Second line	P_6	P_7	P_8

In the implementation of the two-dimensional median filter, the selection of window size and shape is particularly important, that is, because of the different pixels of various object images and different application scenarios, it is necessary to select unique size and shape in the design of sliding window. In order to reduce the time complexity of median filtering and improve the efficiency of the algorithm, a fast parallel algorithm with lower time complexity was selected to replace the typical bubble sorting algorithm. The main idea of the fast parallel method is described below (with the size window as the research object). Table 1 shows the arrangement of window pixels.

The maximum, minimum and median values of each column in Table 1 are calculated respectively, and the calculation formula is as follows:

Maximum group: $Max0 = \max[P_0, P_3, P_6]$, $Max1 = \max[P_1, P_4, P_7]$, $Max2 = \max[P_2, P_5, P_8]$

The median group: $Med0 = \text{med}[P_0, P_3, P_6]$, $Med1 = \text{med}[P_1, P_4, P_7]$, $Med2 = \text{med}[P_2, P_5, P_8]$

Minimum group: $Min0 = \text{med}[P_0, P_3, P_6]$, $Min1 = \text{med}[P_1, P_4, P_7]$, $Min2 = \text{med}[P_2, P_5, P_8]$

According to the above series of operational formulas and the corresponding data, the following conclusions can be obtained after analysis: the median value in the median group, the maximum value in the minimum group, and the minimum value in the maximum group is the median value of these 9 elements.

3. Feature extraction of lyophilized powder foreign mater image

In 1901, Pearson first proposed principal component analysis (PCA) algorithm [11]. The principal component is also called the principal element. The core idea of this algorithm is to map the high dimensional vector to the low dimensional vector space, then the original high dimensional vector can be represented by this low dimensional vector. An appropriate eigenvector matrix is introduced in the process of dimensional transformation. In fact, in the process of dimensional transformation, the image will lose some secondary information, but it does not affect the extraction of the main features of the image.

3.1. PCA fast algorithm

The operation time of PCA algorithm is usually wasted on the calculation of the eigenvalues and eigenvectors of the covariance matrix of the sample data set. If the data set matrix \mathbf{X} has a size $n \times d$, then, when calculating the covariance matrix s_t of the data set, it is equivalent to calculating a square matrix of $d \times d$. If the dimension d is of a large order of magnitude, the operation time complexity will become very high, and a large amount of time is needed to calculate all the eigenvalues of the covariance matrix. In order to improve the time efficiency of the algorithm, a fast PCA algorithm was introduced in the experiment [12].

Suppose $\mathbf{Z}_{n \times d}$ matrix is the matrix obtained by subtracting m from each sample of the sample matrix \mathbf{X} , then the distribution matrix \mathbf{S} is $(\mathbf{Z}^T \mathbf{Z})_{d \times d}$. Now let's introduce a new matrix \mathbf{R}

$$\mathbf{R} = (\mathbf{Z}\mathbf{Z}^T)_{n \times n} \quad (2)$$

Therefore, the dimension of the matrix is $n \times n$. In general, the dimension d of the sample image is often larger than the number n of samples, and the size of \mathbf{S} is often larger than \mathbf{R} , so it is easier to calculate the eigenvalue of \mathbf{R} and the time complexity is lower. In fact, the eigenvalues corresponding to \mathbf{S} and \mathbf{R} are the same, so only the eigenvalues of \mathbf{R} with a small order of magnitude are required to obtain the eigenvalues of \mathbf{S} .

Set the n dimensional column vector \mathbf{v} to the eigenvector corresponding to \mathbf{R} :

$$(\mathbf{Z}^T \mathbf{Z})\mathbf{v} = \lambda \mathbf{v} \quad (3)$$

Multiply both sides of the equation by \mathbf{Z}^T to get the following equation:

$$(\mathbf{Z}^T \mathbf{Z})(\mathbf{Z}^T \vec{v}) = \lambda(\mathbf{Z}^T \vec{v}) \quad (4)$$

$\mathbf{Z}^T \mathbf{v}$ is the eigenvalue of the matrix $\mathbf{S} = (\mathbf{Z}^T \mathbf{Z})_{d \times d}$ in equation.

Through the above argument, the eigenvalues of matrices with a larger order of magnitude and their corresponding eigenvectors can be obtained based on the eigenvalues of matrices with a smaller order of magnitude and their corresponding eigenvectors. The fast PCA algorithm is used to reduce the time complexity and optimize the algorithm.

The fast PCA algorithm was used to reduce the dimension of the defective lyophilized powder image and extract the principal component characteristic information of the foreign matter according to the training set image. In the multi-classification (fiber, hair, glass debris) experiment, the number of principal component characteristics was set to 15, and the transformation of principal component characteristics was realized through projection. By processing image information with the above method, the feature vectors of each foreign matter sample (fiber, hair, glass debris) in the sample image are trained to decrease from high dimension to 15 dimension. In the following series of experiments, the foreign matter is identified with a 15-dimensional basis.

The core algorithm of fast PCA is as follows:

1. Feature centralization.
2. The original data set \mathbf{A} is a 250×150 matrix. The deviation of each column of the matrix is calculated and assigned to the matrix \mathbf{B} .
3. Calculate the covariance matrix \mathbf{C} of \mathbf{B} .
4. Find the eigenvalue of covariance \mathbf{C} and its corresponding eigenvector.
5. Select the eigenvalues with larger values and their corresponding eigenvectors to form a new data set.

4. Classification and recognition of foreign matter images of lyophilized powder

4.1. Third-Nearest Neighbor classification algorithm

In this paper, the third-Nearest Neighbor classification algorithm was used. The third-Nearest Neighbor classification algorithm has mature theory, simple algorithm and can be used for nonlinear classification. However, this algorithm also has some disadvantages: unbalanced sample effect is not good, large amount of prediction calculation, large memory consumption.

K-Nearest Neighbor (KNN) classification algorithm first selects the different eigenvalues corresponding to the images to be tested and the images in the training set, calculates the distance between them, and

classifies them based on the minimum distance. The idea of KNN algorithm: the label and data of training set images are determined, then the distance between the corresponding features of training samples and the features of test samples is calculated, K pieces of data with the smallest interval are selected, and the category to which the K pieces of data belong is counted. The class with the largest number of times is the category to which the test sample belongs. The KNN algorithm process is as follows:

1. Calculate the distance between the training set image and the test set image features.
2. Select K data points with the smallest distance.
3. Count the categories of K data points.
4. The category to which the test sample belongs, that is, the category with the highest frequency in which K points are located.

Calculating the absolute value of the difference between pixels is an important part of the third-order neighborhood recognition and classification algorithm. The formula to solve the distance is as follows:

$$L(x, y) = |x_i - y_i| \quad (5)$$

The nearest similar image is obtained from Eq. (5), and its category is determined according to the frequency of occurrence of the category.

The main task of the third-Nearest Neighbor classification algorithm is to identify three training set images. At the same time, the classes of the selected three training set images are set as one, two and three respectively. If one and two do not belong to the same category, and two and three do not belong to the same category, it can be concluded that the test graph belongs to one. If one and two belong to the same category, then the test graph belongs to one, but the test graph is similar to two. If two and three belong to the same category, then the test graph belongs to two, but the test graph is similar to three.

4.2. SVM classifier

Vapnik, Boser and Guyon first proposed the SVM [13–15] algorithm. The classification model of SVM is based on linear separable. For non-linear scenarios, it is often used to project from low-dimensional space to high-dimensional feature space for information processing, in which the non-linear mapping algorithm is involved, so as to transform the non-linear data set into the linear separable data set. Therefore, linear algorithm can be used to analyze the original data set in high dimensional space.

SVM aims at the case of linear separability. For the training sample set (x_i, y_i) , and $x_i \in R_N, y_i \in \{-1, 1\}, i = 1, 2, 3 \dots n$ is to find a special hyperplane, which can well separate the data of these different categories from each other and maximize the spacing between these different categories.

Suppose the equation of this particular hyperplane is $w \cdot x + v = 0$, if $w \cdot x + b > 0$ then this class is set to the “1” class, if $w \cdot x + b < 0$ then this class is set to the “-1” class. In fact, the maximization of the two types of interval is equivalent to the minimization:

$$J(w) = \frac{1}{2} \|w\| \quad (6)$$

Constraint conditions: $y_i(w \cdot x + b) \geq 1, \forall i \in \{1, 2, 3 \dots n\}$.

Lagrange multiplier is added here, so that Eq. (5) can be written as Wolf dual form to obtain the following equation:

$$L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (7)$$

Solve for α and use α to determine the plane parameters w and b . According to the above, the optimal classification function of SVM can be determined as follows:

$$f(x) = \text{sgn} \left(\sum_{i=1}^n \alpha_i y_i (x_i \cdot x) + b \right) \quad (8)$$

For the case of linear indivisibility, the original image sample needs to be projected into a high-dimensional space first. And the kernel function is chosen to be linearly separable in high dimensional space. Thus, the objective function can be converted into the following equation:

$$J(w, b, \varepsilon) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \varepsilon_i \quad (9)$$

In this equation, ε represents the relaxation variable, and C represents the penalty factor, which must meet the following conditions:

$$y_i(w \cdot x + b) \geq 1 - \varepsilon_i, \varepsilon_i \geq 0, \forall i \in \{1, 2, 3 \dots n\} \quad (10)$$

The optimal classification function is:

$$f(x) = \text{sgn} \left(\sum_{i=1}^n \alpha_i y_i K(x_i \cdot x) + b \right) \quad (11)$$

where $K(x_i \cdot x)$ is the kernel, such as Linear kernel function, Polynomial kernel function, Radial basis kernel function and Radial basis kernel function.

In the lyophilized powder foreign matter image samples, the common three types of visible foreign matter are glass debris foreign matter, fiber foreign bodies and hair foreign matter. There are 10 pictures of foreign matter in each category of lyophilized powder for injection. The first 5 pictures of foreign matter in each category are set as the training set data, and 15 training set data are obtained. Each graph is represented by a row matrix, so 15 graphs form a 15 by 250 by 150 matrix. The remaining images are set as test data Eq. (5). When we read the training sample images, we have identified the labels `real_labels` and `picture_matrix` for all training images.

The basis of the foreign matter image in the low-dimensional space is the feature image of the training set, so all images, after dimensionality reduction, can be expressed linearly by the feature image of the training set, which is the basis of image recognition.

The implementation steps of SVM algorithm are as follows:

1. Set training set and test set;
2. Set labels for each type of foreign matter data set;
3. Use the training set for training, so as to obtain the training classifier `svm_model`;
4. `Accuracy_rate` is measured and calculated for the test set according to the `svm_model`.

5. Experiment and analysis

5.1. Experimental environment

Matlab R2014b and Professor Lin Zhiren's LIBSVM toolbox [16] were used to realize the experiment. Industrial pictures of defective lyophilized powder for injection were used as data sets. The data contains fiber, hair, glass debris and other foreign bodies. The pictures of foreign bodies collected by the industry

Table 2
PCA+third-Nearest Neighbor classification algorithm results

Image preprocessing method	Image recognition rate
3×3 Median filtering of templates	60% (9/15)
5×5 Median filtering of templates	33.3% (5/15)
7×7 Median filtering of templates	40% (6/15)
9×9 Median filtering of templates	13.3% (2/15)
3×3 Wiener filtering of templates	60% (9/15)
3×3 Average filtering of templates	33.3% (5/15)

Table 3
PCA+SVM algorithm results

Image preprocessing method	Image recognition rate
3×3 Median filtering of templates	73.3% (11/15)
5×5 Median filtering of templates	33.3% (5/15)
7×7 Median filtering of templates	40% (6/15)
9×9 Median filtering of templates	13.3% (2/15)
3×3 Wiener filtering of templates	60% (9/15)
3×3 Average filtering of templates	73.3% (11/15)

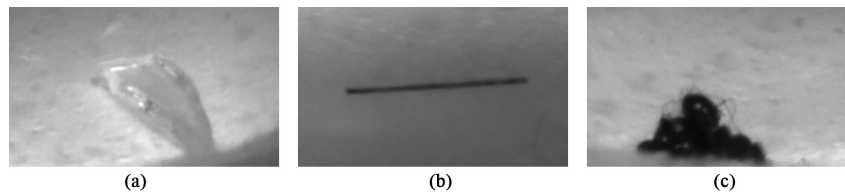


Fig. 1. Foreign matters: a. Glass debris, b. Hair, c. Fiber.

are shown in Fig. 1. Due to the different shooting angles and light, the form and size of visible foreign bodies in the same category are also different. In the experimental data set, each kind of producing injection has 10 foreign matter images, and put the pictures of the each kind of foreign matter before 5 set as the training set, the rest picture set to test set, after producing a foreign matter image preprocessing, image feature extraction and producing a foreign producing foreign matter image classification and recognition of the experiment, to achieve more than two classes and the class of visible foreign matter classification and recognition. In the process of lyophilized powder foreign matter image preprocessing, first of all, each image was converted to gray, because the foreign matter image collected by the industry is colored. Then, median filtering, Wiener filtering and average filtering were carried out to denoise the image.

5.2. Results and analysis

Two links, image preprocessing and image classification and recognition, were compared. The first group of comparative experiments compared the effects of various image preprocessing methods when the classifier was multi-classification (fiber, hair, glass debris) and the image classification and recognition algorithm were the same. The feature extraction algorithm of lyophilized powder foreign matter image is PCA, and the classification and recognition algorithm of foreign matter image is third-order neighborhood classification and recognition algorithm and SVM, respectively. Experimental results are shown in Tables 2 and 3.

As can be seen in Tables 2 and 3, in view of the industrial collection of small sample producing injection

Table 4
Classification (fiber, hair) identification results

Method	Image recognition rate
PCA+third-Nearest Neighbor algorithm	90% (9/10)
PCA+SVM algorithm	90% (9/10)

Table 5
Multi-classification (fiber, hair, glass debris) identification results

Method	Image recognition rate
PCA+third-Nearest Neighbor algorithm	60% (9/15)
PCA+SVM algorithm	73.3% (11/15)

foreign matter images, the classifier for classification of classifier and image recognition algorithm under the condition of same, producing defective product image preprocessing effect is best with 3×3 median filtering and Wiener template, however 9×9 of the template effect of median filtering is the worst. Through the first group of comparative experiments, in the following series of experiments, the image preprocessing method of the defective lyophilized powder selects the median filtering of the 3×3 template. The second group of comparison experiments, in the case that the classifier is a multi-classifier and the image preprocessing method is the same, compare the effects of various classification recognition algorithms. The experimental results are shown in Tables 4 and 5.

As can be seen in Tables 4 and 5, in view of the industrial collection of small sample producing injection foreign matter image, image preprocessing method in classifier for multiple classifier and the same situation, based on PCA and SVM algorithm of recognition rate than based on PCA and the third-Nearest Neighbor classification algorithm of high recognition rate. Through the above multiple comparative experiments, the median filter of 3×3 template was selected in the image preprocessing method of lyophilized powder defective products, and the classification and recognition effect was the best based on principal component analysis and support vector machine algorithm.

6. Conclusion

Lyophilized powder injection of defective products often contains some visible foreign matter, such as glass fragments, hair, etc. and in order to ensure medical safety, these visible foreign matters need to be detected and classified in industrial production. Image preprocessing of defective lyophilized powder was carried out, and the denoising methods of median filtering, Wiener filtering and average filtering were studied. In the phase of foreign matter classification and recognition, fast PCA algorithm, third-Nearest Neighbor classification algorithm and SVM algorithm were studied. All the above methods have been tested with small sample data collected by the industry. Through multiple groups of comparative experiments. In the case that the classifier is a multi-classifier and the image classification recognition algorithm is the same, the best image preprocessing effect is the median filtering of the 3×3 template. If the classifier is a multi-classifier and the image preprocessing method is the same, the recognition rate based on PCA and SVM algorithm is higher than that based on PCA and third-Nearest Neighbor classification algorithm. The clearer the collected pictures of foreign matter in freeze-dried powder are, the better the effect of classification and identification. In this paper, the key technology of machine vision automatic detection of visible foreign matter in lyophilized powder injection has been studied, and many practical application problems have been solved, and some research results have been obtained. The experimental results show that the more pictures of lyophilized powder injections are collected, the higher the picture clarity and the higher the accuracy of recognition.

Acknowledgments

This project was supported by the National Natural Science Foundation of Hunan Province (grant nos 2018JJ3565 and 2018JJ2459), the Major Science and Technology Projects in Hunan Province (grant nos 2019SK1013 and 2019SK1010), the Technology Innovation Guidance Program Clinical Medical Technology Innovation Guide Project of Hunan Province (grant no. 2018SK50504), the Science and Technology Plan of Changsha (grant nos kc1809035 and k1509003-11), and the Scientific Research Fund of the Hunan Provincial Education Department (grant no. 19A048).

Conflict of interest

None to report.

References

- [1] Ishii A, Mizuta T, Todo S. Detection of foreign substances mixed in a plastic bottle of medicinal solution using real-time video image processing. In Fourteenth International Conference on Pattern Recognition, 1998, p. 1646. doi: 10.1109/ICPR.1998.712034.
- [2] Shimizu I, Kato F, et al. A technique for making holograms easily and for measuring simultaneously the behavior of particles of different sizes and/or shapes. *Measurement Science and Technology*, 2004, 15(4): 656. doi: 10.1088/0957-0233/15/4/007.
- [3] Kato F, Miyakawa T, Shimizu I. Developmental Research of Visualization of Spatial Distribution and the Behavior of Particles in the Room. In 20th Annual Tech. Meeting on Air Cleaning and Contamination Control. Tokyo, Japan, 2002, p. 127.
- [4] Peng B, Wang Y, Hall TJ, et al. A GPU-accelerated 3-D coupled subsample estimation algorithm for volumetric breast strain elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2017, 64(4): 694. doi: 10.1109/TUFFC.2017.2661821.
- [5] Peng B, Wang Y, Yang W, et al. Relative elastic modulus imaging using sector ultrasound data for abdominal applications: An evaluation of strategies and feasibility. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2016, 63(9): 1432. doi: 10.1109/TUFFC.2016.2589270.
- [6] Xiao F, Wang Y, Zhang S, et al. Study on real-time visual detection system of visible foreign bodies in ampoule preparations. *Computer Measurement and Control*, 2010, 18(2): 295.
- [7] Ge J, Wang Y, Zhang H, et al. Research on intelligent light detector based on improved PCNN. *Journal of Instrumentation*, 2009, 30(9): 1867.
- [8] Lu J, Wang Y, Yu H, et al. Design of intelligent online detection system for visible foreign matter in large infusion. *Computer Measurement and Control*, 2008, 16(12): 1802.
- [9] Lu C. Iris recognition system based on feature fusion and optimization of extreme learning machine algorithm. *Computer Application and Software*, 2016, 33(7): 326.
- [10] Chen G. A hierarchical median filtering algorithm based on noise connection components. *Computer Applications and Software*, 2016, 33(10): 321.
- [11] He L. Multi-dimension principal component analysis based on face recognition. *Journal of New Industrialization*, 2012, 2(1): 59.
- [12] Virmani J, Dey N, Kumar V. PCA-PNN and PCA-SVM Based CAD Systems for Breast Density Classification. 2016, 96: 159. doi: 10.1007/978-3-319-21212-8_7.
- [13] Burges CJC. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 1998, 2(2): 121. doi: 10.1023/a:1009715923555.
- [14] Combes C, Vapnik V. Support vector networks. *Machine Learning*, 1995, 20: 273. doi: 10.1007/bf00994018.
- [15] Deng N, Tian Y. A new method of data mining supports vector machines. Beijing: science press, 2004, 202.
- [16] Chang C, Lin C. A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2011, 2(27): 1. doi: 10.1145/1961189.1961199.