Complicated Graphics Model based on Neural Network of Particle Swarm

Zemin Liu
College of Mathematics and Computer Science, Panzhihua University, China

Wei Zong
Massachusetts Institute of Technology, Cambridge, MA 02139, USA
Email: wzong@mit.edu

Abstract—The past decades have seen the great progress pattern recognition and image understanding, motivated by a wide range of real world applications. The previous approaches mostly based on the feature extraction and recognition methods, and usually suffer from the problem of weak discrimination power. In this paper, we propose a complex model based on particle swarm and neural network pattern recognition methods. The proposed approach uses the class label of the input sample as the actual output of the neural network which has the maximum node corresponds to the class. To evaluate the effectiveness of the proposed approach, we experimentally compare the particle swarm neural network based approach with and SVM classification algorithm. The experimental results show that, both the recognition accuracy and training time of the particle swarm and neural network based approach, are better than that of SVM. The particle swarm neural network pattern recognition in complex graphics showed a great advantage in application.

Index Terms—Particle Swarm; Neural Network; Graphic Pattern Recognition

I. INTRODUCTION

Researchers started recognition theory and applications since the 20s of the 20th century. With the rapid development of computer and artificial intelligence techniques, it has been a branch of science until the early days of the sixties. The continuous development of mode recognition techniques obtains a wide range of applications such as character recognition, speech recognition, electroencephalogram, bio-sensor, graphic processing and other fields. The recognition approach of complex graphics mode is one significant computer application based on graphic processing, pattern recognition, graphic interpretation, and classification technique. The general progress of technique brings more and more people to obtain and manage different information by computer technique. From the statistics, about 70% people receive information through visual sense, which means that visual information is the most important way to communicate. The large amount of information with the direct expression of graphics place more and more important place during the target recognition in recent years [1]. Therefore, the pattern recognition of complex graphic research becomes the hot issue [2-3].

In recent years, the fast development of artificial intelligence approaches provides new method especially in the pattern recognition of complex graphics [4]. More and more scholars bring neural network, support vector machine and Bayesian decision theory into the complex patterning of mode recognition [5]. The empirical risk maximum of artificial neural network appears over-fitting and slow learning speed and other disadvantages during the pattern recognition. The recognition performance of support vector machine exerts a tremendous influence from the parameters [6]. The appropriate parameter selection develops advantage of support vector machine. On the contrary, the inappropriate parameter selection brings great difficulty to the graphic pattern recognition. For example, the recognition time gets longer and the recognition accuracy rate becomes lower. The Bayesian classification approach can hardly satisfy the large-scale classification problem. It is difficult to realize the assumption of classification unconditional independence and the required evaluation function selection. Moreover, it increases the learning training complexity. If we are unable to effectively manage the graphic data, the large amount of information will be lost. Moreover, people cannot search the required messages. Therefore, how to realize the effective identity management of graphics become the hot research spot [7-8].

The particle swarm optimization, PSO is a kind of evolving algorithm, which is close to the optimal solution step by step, and then to find the nearest expected number output at the end of training. PSO algorithm was suggested by James Kennedy and Russell Eberhart in 1995[9]. The principle is estimating the bird flock foraging. The PSO is established under organization operation system of group cooperation. During the deepening process, each particle continuous explores in the solution space, and it can memorize the particle best value (pbest) during this period [10-12]. Otherwise, the particle among each other can transmit the optimal particle position, which is called the Global best value (gbest), or the area of the optimal particle is called the local best value (lbest). The algorithm of optimal particle swarm decides each particle exploration direction based
on \( p_{best} \), \( g_{best} \) and \( l_{best} \). Then the particle swarm can close to the target location of the solution space.

It provides the structure of the Probabilistic Neural Network. It is one kind of supervised network architecture. The principle establishes on the Bayesian Decision and non-parametric method to estimate the probability density function (PDF). This PDF has the pattern of Gaussian distribution. The GRNN theory basis is nonlinear regression.

From the above disadvantages, this article provides one research method for complex pattern recognition which based on the neural network of particle swarm. Because the essence of artificial neural network algorithm is based on the gradient descent, it suffers from local minimum, slow convergence rate, and long-term training time. Moreover, the optimization algorithm of particle swarm can solve the above problems. We do not need the differentiable function in the gradient descent, and this algorithm can shorten the training time. Therefore, the combination of particle swarm and neural network can improve local research ability and optimize the global searching ability.

Moreover, we treat the input sample classification as the corresponding classification of the maximal node during judging the actual output of neural network. At last, we compare the recognition algorithm result between the classification of neural network of particle swarm and SVM. The result shows, the accuracy, training time, training error of particle swarm neural network classification is better than the SVM. The neural network of particle swarm places great advantages during the pattern recognition application.

The contributions of the proposed approach in this paper are threefold as follows: (1) we propose to extract data feature using neural networks. To our best knowledge, this is a novel method of this field; (2) we employ particle swarm to train the neural network which is used for feature extraction. The advantage of using particle swarm for training neural network is that particle swarm is able to find a satisfied near-global optimum solution; (3) the extracted feature is delivered to least-square support vector machine (LS-SVM) for recognition. Benefit from the good generalization ability for LS-SVM, the proposed approach in this paper is robust to noise, the number of training set, and show well generalization ability when the number of training samples is limited.

The remainder of this paper is organized as follows. In Section 2 we will propose the main approach, including particle swarm, neural network, and least square support vector machine. Section 3 will evaluate the proposed approach using a set of carefully designed experiments. We draw a conclusion in Section 4.

II. THE PROPOSED SCHEME

We in the section will propose a new approach for image recognition. First we extract a set of features from images, Second, we employ neural networks to map the image feature so that they will have better discrimination ability; Third, we use particle swarm algorithm to solve the neural networks. Fourth, the image features are delivered to least square support vector machine for recognition. The illustration of the proposed approach can be found in the next section.

A. Image Feature Extraction

If we directly operate the large amount of the collected pattern, it is difficult to obtain the feature vector with strong discriminant power. Therefore, we need to extract

---

**Figure 1.** The framework and the optimizing process of the proposed algorithm
features before the classification, which is the significant step during pattern classification. The research of this article pays attention to the color and graphic features extraction as the pattern input.

(1) Color feature extraction. Color feature is one of the most visual perceptual characteristics among the graphic features. Compare with other features, it is easy to extract under the most circumstances, we can obtain the satisfactory result. Therefore, color feature classification receives much attention during the pattern recognition process. Before the color feature extraction we need to define the gray difference. The sum of square difference of RGB value and the average after square root is d, and we use the average value of d as the gray difference D. The formula is in the following.

\[ d = \sqrt{(R - \text{ave})^2 + (G - \text{ave})^2 + (B - \text{ave})^2} \]

\[ \text{ave} = \frac{R + G + B}{3} \]

\[ D = \sum_{i=0}^{N} d_i / N, \ N \text{ is the pixel number in the image.} \]

When RGB valued are equal, the graphic will be the solid gray. The general black-and-white photograph belongs to this type. If the three values are very close, the graphic will on the verge of aromatization. On the contrary, the large difference among three values will appear a certain extent color. Then we express the graphic color in the following formula:

\[ E = \sum_{m=0}^{N} \Delta T \]

\[ E = \sum_{m=0}^{N} 1,0 \text{ If } C_i \text{ is not equal to 0} \]

\[ E = \sum_{m=0}^{N} 0, \text{ Otherwise} \]

In this equation, we map the color space into HSV, which is coincident with human senses. We use the three values to express it, which are Hue (H), Viscosity (V) and Saturation (s). We express this conversation with another formal denotation:

\[ T: \text{RGB} \rightarrow \text{HSV} \]

At last, we quantize HSV space into 256 colors. The quantization for is 16 levels, S and V are the same 4 levels. For the convenient calculation, we use the one-dimensional feature space to represent the HSV three-dimensional space. [7] The process is

\[ Q: \text{HSV} \rightarrow C \]

In this formula, \( C = \{ C_i | i = 0,1,2,\ldots, 255 \} \), \( C_i \) is the i-th quantized color.

(2) Extraction of graphic texture features. Texture is principle standard feature used in the graphical evaluation. It means pixel gray level or the color variation. This article researches how to better obtain texture feature for better evaluation, understanding and classification. There are three methods for texture feature extraction: texture feature extraction based on structuration, and the texture feature extraction based on statistics. The 6components given by texture feature extraction correspond to the 6features from psychology. They are regularity, orientation degree, roughness, and contrast and alignment degree.

The texture feature extraction in this paper is based on the texture feature of the gray co-occurrence matrix. In the gray co-occurrence, we investigate the two pixel combinations about the gray configuration. It is the representative calculation of second order statistics about the texture feature. It decides the distance d and the direction \( \theta \). The direction is on the line of \( \theta \), e.g., 0, 45, 90, 135degrees. If one pixel is \( i \), the other pixel has distance \( d \) from the pixel \( i \), and the gray level of pixel is the \( j \) emerging frequency. Then we normalize them to express the probability \( p_{ij} \). The resulted value is the \( (i,j) \) array element value of gray co-occurrence matrix. [5] We set \( L \) as the graphic gray level. \( (p_{ij}) \) denotes the gray co-occurrence matrix. Actually, it is the tea-color histogram of the two pixels. We use the obtained gray co-occurrence matrix can elicit a series of textural features. This article applies gray co-occurrence matrix from Haralick [6] to calculate the statistics.

B. Adaptive Neural Network Using Particle Swarm

The nonlinear regression of independent variable X relative to the dependent variable \( \hat{Y} \) is the maximum estimated value of random variable \( y \) which is relative to the random variable \( x \). We set the function \( f(x, y) \) is the joint probability of the random variable \( x, y \). \( X \) is the measured value of \( x \) as well as the known quantity, and the nonlinear regression of random variable \( y \) is:

\[ \hat{Y} = E(y|X) = \frac{\int_{-\infty}^{\infty} y f(x,y) dy}{\int_{-\infty}^{\infty} f(x,y) dy} \] (1)

where \( \hat{Y} \) is the \( Y \) estimated output which under the input variable \( X \). The sample data set \( \{x_i, y_i\}_{i=1,2,\ldots,n} \) applies non-parametric estimation and we can obtain the estimated density function \( \hat{f}(X, y) \).

\[ \hat{f}(X, y) = \frac{1}{n(2\pi)^{p/2} \sigma^{p+1}} \exp \left[ \frac{-(X - Y)^2}{2\sigma^2} \right] \] (2)

\[ \sum_{i=1}^{n} \exp \left[ -\frac{(X - X_i)^2}{2\sigma^2} \right] \] (3)

GRNN network is similar to the RBF network in the network structure. The GRNN structure has input layer, pattern layer, summation layer, and output layer. Figure 3 has the details. Input layer, GRNN input layer has the same number of neurons in comparison with the input vector dimension of training sample. Each neuron directly transmits input vector to the next layer neuron without function calculation. Pattern layer. The neural number of GRNN pattern layer is same with the training sample quantity. The pattern layer transfer function is in the following:

\[ p_i = \exp \left[ -\frac{(X - X_i)^2}{2\sigma^2} \right], \quad i = 1,2,\ldots,n \] (3)
where \( X_j \) is the corresponding sample observation of the \( i \)-th neuron. Summation layer. GRNN summation layer uses two types of functions for the summation. The first type of function means the arithmetic summation of all the neuron output in the pattern layer. The neuron connection weight between pattern layer and summation layer is 1, and the transfer function is \( S_D = \sum_{n=1}^{n} P_i \). The second type of function is the weighting summation of all the neuron output in the pattern layer. The \( j \)-th neuron and the \( i \)-th neuron in the pattern layer have the link weight. The transfer function is \( S_{\phi/j} = \sum_{i=1}^{n} y_i P_i \), Output layer. The neuron number of GRNN output layer equal to the vector dimension of training sample output, which means:

\[
y_j = \frac{S_{\phi/j}}{S_D}, \quad j = 1, 2, \ldots, k
\]

(4)

Here we used Particle swarm algorithm. The algorithm process is the following: (1) randomly initialize \( N \) particles; (2) calculate the particle adaptation; (3) track the individual optimal position and group optimal position of \( N \) particles; (4) modify the particle speed; (5) particle new speed is used to modify the location; (6) If the entire adaption has found or the iterations have reached the upper limit, it will stop calculation. Otherwise, return to step 2, re-perform the calculation. The optimizing process is shown in Figure 1.

**C. Least Square SVMs**

Suykens (1999) proposed least square support vector machine (LS-SVM). Choosing different optimization function SVM turns into inequality restrict. Suppose there are \( N \) samples \((x_i, y_i) \in \mathbb{R}^n \times \mathbb{R}\). Consider a nonlinear function \( \phi(x) \), with which the sample was mapped from the original space to feature space. In the high dimensional feature space setting up \( y(x) = w \cdot \phi(x) + b \).

Therefore, we can map the nonlinear function in the origin space into the linear function in high dimensional feature space. Using structural risk minimum principle to determine the coefficient \( w \) and \( b \), building the linear programming model is as follows:

\[
\min_{w, b, \xi} J_p (w, \xi) = \frac{1}{2} w^T R w + c \sum_{i=1}^{N} \xi_i^2
\]

(5)

where \( c \) is the penalty coefficient. Among them, the distance of any point \( x_i \) to hyperplane can meet certain conditions. The corresponding Lagrange function is:

\[
L(w, b, e, a) = J_p (w, e) - \sum_{i=1}^{N} a_i \left\{ w^T \phi(x) + b + e_i - y_i \right\}
\]

where \( a_i \) is the Lagrange multiplier, it can be used to solve the following problem. Kernel function is the key of support vector machine. Note that \( K(x_i, x_j) = K_{ij} = \phi(x_i) \cdot \phi(x_j) \) is a real value symmetric function defined in \( \mathbb{R} \times \mathbb{R} \), which makes the integral operator positive. This paper chooses the kernel function. It is the symmetric function that can meet the Mercer conditions. The radial basis kernel function,

\[
K(x, x) = \exp\left(\frac{||x-x||^2}{2\sigma^2}\right)
\]

III. EXPERIMENTAL RESULTS

**A. Data Sources**

The dataset used in this experiment comes from Corel graphics library which is collected from Internet. It is the popular and standard database for the pattern recognition filed. This database collects 9 types of patterns with 2000 graphs. They are buildings, birds, flowers, people, trees, elephants, clouds, mountains and cars, which have about 200 samples in each category. This database covers a wide range of objects. The buildings and cars belong to man-made objects or scene. Birds, flowers, people, trees, and elephants belong to the natural objects. Clouds and mountains are part of the typical natural scene. They are collected from different locations, points of view, and the illumination. Therefore, this database has the typicality and strong challenging and can provide comprehensive and credible evaluation of the suggested pattern recognition methods.

**B. Algorithm Application**

The solution of neural network algorithm is based on the gradient descent. It will suffer from local minimum, slow convergence rate, and long-training time. Moreover, the optimization algorithm of particle swarm can deal with the above problems. We do not need the differentiable function in the gradient descent, and this algorithm can shorten the training time. This paper combines the particle swarm and neural network algorithm. Using the particle swarm optimizing to training of neural network has the following algorithm process in Table 1.

Explanation: (1) In the process 2, each parameter is: population size is \( m \), learning factor is \( c_1, c_2 \), the inertia weight is \( 2 \). Moreover, the maximum speed, the maximum iterations, accuracy requirement, particle initial position, and the initial velocity are \( v_{max}, m, max, \varepsilon, Z, V \); (2)
In step 3, if the iterations \( k \) is lesser than the maximum iterations, it will go on the next step [7].

C. Main Experimental Result

We extract color feature and textural features of each image. The color feature extraction method is in the Section 2.1.1, and the textural feature extraction is in the Section 2.2.2. We can use the features to represent the samples. Moreover, we can put the features into the neural network of particle swarm and SVM classifiers for the experiment of complex pattern recognition. Each experiment will repeat for 20 times, and the final result is the average of error rate and the training time of each round. In the experiment, the first 10 groups of samples are the training samples, and the remainder 5 groups of data are the testing samples. Moreover, we treat the input sample classification as the corresponding classification of the maximal node of the neural network during judging the actual output of neural network. At last, we compare the recognition result between the classification of neural network of particle swarm and that of SVM. The neural network classification of particle swarm algorithm has the parameter setting of: population size \( n \) is 200 and the maximal iteration time is 300. We set \( c_1 = c_2 = 4, \ w_{\text{max}} = 0.8, \ w_{\text{min}} = 0.3 \). This paper made the experiments respectively by SVM algorithm and the method proposed in this paper.

In the first experiment, it used the SVM method to do the experiment. From whose experiment results we can give a comparison with the method proposed in this paper. Then by means of the comparison result, we can see the effectiveness of the method in this paper. During the experiment, it uses the \( O_i \) to express the \( i \)-th round experiment, and each line corresponds to a recognition error rate of every category in each round experiment. We setting the experiment parameters are \( c_1 = c_2 = 4, \ w_{\text{max}} = 0.8, \ w_{\text{min}} = 0.3 \). And the error rate and the training time are as evaluation standards of the experiment. From the experiment results in table 2 we could find that the error rate is concentrated on 8% or 9%, and the highest can be 9.87%. So we can see it is much higher than that of other methods. The reasons accounting for these are as follows: Firstly, the Support vector machine (SVM) can use the hyperplane directly to perform classification. The method shows great advantages when dealing the linear problems. But in fact, we often meet the nonlinear problems, as the problem in this paper. Secondly, when the SVM algorithm deals with the nonlinear problem, the classification result greatly depends on the selection of the kernel function. But now it has not a fixed method to select the kernel function. So for the different kernel functions we may get the different classify results. Last, due to the SVM algorithm is based on the gradient descent method, so it is easy to suffer from local minimum, slow convergence speed, long training time and other issues in the classification process, which will have a big impact on the classification results. All the above reasons can lead to the high error rate. The experimental result is in the following Table 2, [1]

<table>
<thead>
<tr>
<th>Algorithm 1</th>
<th>Algorithm process of particle swarm optimization to the neural network</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Based on the network structure, we map the network threshold and weight into the particle dimensionality ( D ).</td>
</tr>
<tr>
<td>2</td>
<td>The parameter initial setting of particle swarm algorithm;</td>
</tr>
<tr>
<td>3</td>
<td>Judge whether the iterative frequency ( k ) reach the maximum ( n_{\text{max}} );</td>
</tr>
<tr>
<td>4</td>
<td>Calculate particle adaption value of the population ( f(z) );</td>
</tr>
<tr>
<td>5</td>
<td>Based on the particle adaptation value to update the best position (pbest) of each particle and the current optimizing value;</td>
</tr>
<tr>
<td>6</td>
<td>Judge whether the optimal position MSE (mean square error) can satisfy the set accuracy requirement. If MSE is larger than the requirement, it will go on the next step. Otherwise, it will turn to 9;</td>
</tr>
<tr>
<td>7</td>
<td>From the above (2) and (3) to update particle position ( V_i ), and speed ( X_i ) of each population.</td>
</tr>
<tr>
<td>8</td>
<td>Add 1 to the iterative times, and adjust inertial weight ( w ) at the same time, then return to step 3;</td>
</tr>
<tr>
<td>9</td>
<td>The global optimum particle mapping of the obtained population is network weight and threshold;</td>
</tr>
<tr>
<td>10</td>
<td>Stop the learning of algorithm.</td>
</tr>
</tbody>
</table>

TABLE II. CLASSIFICATION RESULT OF THE PROPOSED ALGORITHM

<table>
<thead>
<tr>
<th>Experimental output result</th>
<th>0.0845</th>
<th>0.0858</th>
<th>0.0795</th>
<th>0.0735</th>
<th>0.0790</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0767</td>
<td>0.0748</td>
<td>0.0795</td>
<td>0.0705</td>
<td>0.0780</td>
<td></td>
</tr>
<tr>
<td>0.0867</td>
<td>0.0870</td>
<td>0.0890</td>
<td>0.0905</td>
<td>0.0887</td>
<td></td>
</tr>
<tr>
<td>0.0865</td>
<td>0.0887</td>
<td>0.0869</td>
<td>0.0840</td>
<td>0.0749</td>
<td></td>
</tr>
</tbody>
</table>

TABLE III. NEURAL NETWORK CLASSIFICATION OF PARTICLE SWARM

<table>
<thead>
<tr>
<th>Experimental output result</th>
<th>0.0743</th>
<th>0.0856</th>
<th>0.0891</th>
<th>0.0846</th>
<th>0.0882</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0864</td>
<td>0.0842</td>
<td>0.0890</td>
<td>0.0837</td>
<td>0.0870</td>
<td></td>
</tr>
<tr>
<td>0.0869</td>
<td>0.0907</td>
<td>0.0904</td>
<td>0.0893</td>
<td>0.0873</td>
<td></td>
</tr>
<tr>
<td>0.0962</td>
<td>0.0884</td>
<td>0.0919</td>
<td>0.0941</td>
<td>0.0924</td>
<td></td>
</tr>
</tbody>
</table>

The second experiment, it uses the Neural Network of Particle Swarm Algorithm which is proposed in this paper. From the experiment we can test the effectiveness of the method in the paper. The experiment parameters are \( c_1, c_2, w_{\text{max}} \) and \( w_{\text{min}} \). And we also set them \( c_1 = c_2 = 4, \ w_{\text{max}} = 0.8, \ w_{\text{min}} = 0.3 \). The evaluation standards of the experiment are error rate and the training time. From the following experimental results in Table 3 we can see the error rates of the method proposed in this paper are lesser than 1%, and the highest is only 0.95% which is significantly lesser than that of the Support Vector Machine method. The method proposed in this paper shows great advantage over the Support Vector Machine method. The reasons are form the two aspects. Firstly, the particle swarm optimization algorithm doesn’t need the differentiable function in the search process of the gradient descent method, which makes the algorithm greatly flexible and adaptable when it applied to the classification problem. Finally, by combining the particle swarm optimization (pso) algorithm and the neural network algorithm, it can compensate for the local
minimum, slow convergence speed, long training time and over fitting shortcomings. Besides, they can better play to the potential of each algorithm and improve local research ability and optimize the global searching ability, so it can shows great classification accuracy.

In the third experiment, it gave a comparison about the training time of the above two method. From the comparison results it can better show the effectiveness of the neural network particle swarm classify method. The experiment parameters are \( c_1, c_2, w_{\text{max}} \) and \( w_{\text{min}} \). And we also set as follows \( c_1 = c_2 = 4, w_{\text{max}} = 0.8, w_{\text{min}} = 0.3 \). The evaluation standards of the experiment are error rate and the training time. From the results in Table 4 we can see the average training time (15.453 s) of SVM method in the five round experiments is significantly longer than the proposed training time (8.749 s). And the average training error of the former one is about 8.8% and the latter is only 0.39%. From table4, we could find the classification accuracy, training time and training error of particle swarm neural network are better than the SVM method. The reasons counting for the experiments results are threefold. Firstly, when the SVM algorithm deals with the nonlinear problem, the classify result greatly depend on the selection of the kernel function. But now it has not a fixed method to select the kernel function. So for the different kernel functions we may get the different classification results. Second, the method proposed in this paper combines the particle swarm optimization (pso) algorithm and the neural network algorithm, which can compensate for the local minimum, slow convergence speed, long training time and overfitting shortcomings. Besides, they can improve the local research ability and optimize the global searching ability, so it can shows great classification accuracy. Thirdly, the SVM method is under the gradient descent, which will has the local minimum, slow convergence rate, and long term of training time. Secondly, the particle swarm optimization can fix the above problems. Because it doesn’t need the differentiable function in the gradient descent.

<table>
<thead>
<tr>
<th>Round</th>
<th>Algorithm</th>
<th>Training time</th>
<th>Training error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SVM</td>
<td>15.45s</td>
<td>0.0879</td>
</tr>
<tr>
<td></td>
<td>Neural network of particle swarm</td>
<td>8.74s</td>
<td>0.0039</td>
</tr>
<tr>
<td>2</td>
<td>SVM</td>
<td>16.82s</td>
<td>0.0834</td>
</tr>
<tr>
<td></td>
<td>Neural network of particle swarm</td>
<td>8.90</td>
<td>0.0037</td>
</tr>
<tr>
<td>3</td>
<td>SVM</td>
<td>15.39s</td>
<td>0.0892</td>
</tr>
<tr>
<td></td>
<td>Neural network of particle swarm</td>
<td>9.16</td>
<td>0.0041</td>
</tr>
</tbody>
</table>

Experimental conclusions are two folds. (1) FromTable4, we can find out the classification veracity, training time, and training error of particle swarm neural network is better than the BP neural network. (2) The particle swarm neural network places great advantages of complex image pattern recognition. The algorithm of BP neural networks under the gradient descent, which will have the local minimum, slow convergence rate, and long term of training time. The particle swarm optimization can fix the above problems. It is not the differentiable function in the gradient descent. At the same time, this algorithm can shorten the training time.

### IV. SUMMARY

This paper presents pattern recognition research based on the neural network of particle swarm. The essence of artificial neural network algorithm is upon the gradient descent, it will suffer from local minimum, slow convergence rate, and long-term training time. Moreover, the optimization algorithm of particle swarm can make up the above problems. We do not need the differentiable function in the gradient descent, and this algorithm can shorten the training time. Therefore, the combination of particle swarm and neural network can improve local research ability and optimize the global search ability. Moreover, we treat the input sample classification as the corresponding classification of the maximal node during judging the actual output of neural network. At last, we compare the recognition algorithm results between the classification of neural network of particle swarm and SVM. The result shows that, the accuracy, training time, training error of particle swarm neural network classification is better than that of SVM. The neural network of particle swarm places great advantages during the pattern recognition application. However, the experiment greatly depends on the initial parameter setting. How to effective set the parameters the research hot spot in the future.

### REFERENCES


Zemin Liu is currently an associate professor in the College of Mathematics and Computer Science at the Panzhihua University of P. R. China. He received a bachelor degree of Engineering from Xi'an Jiao Tong University in 1987 and a master's degree from University of Electronic Science and Technology of china in 2007. He is in charge of teaching compiler principles, operating system. He hosts and participates in a number of projects and research topics. At present his research interests include: machine intelligence, computer graphic, digital image processing.

Wei Zong received the B.S. degree in Electrical Engineering and Computer Science, the M.S. and the Ph.D. degrees in Biomedical Engineering, from Xi’an Jiaotong University (XJTU), Xi’an, P. R. China, in 1983, 1986, and 1993, respectively. From 1993 to 1997, he was an associate professor at the Institute of Biomedical Engineering, XJTU. Since 1997, he has been affiliated with the Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology (MIT). His research interests include signal and image processing, pattern recognition, machine learning, data mining, and medical informatics.