Research Article

Classification of ETM+ Remote Sensing Image Based on Hybrid Algorithm of Genetic Algorithm and Back Propagation Neural Network

Haisheng Song,[1,2] Ruisong Xu,[1] Yueliang Ma,[1] and Gaofei Li[1]

[1] *Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, 510640 Guangzhou, China*
[2] *University of Chinese Academy of Sciences, 100049 Beijing, China*

Correspondence should be addressed to Haisheng Song; songhs@ustc.edu

The back propagation neural network (BPNN) algorithm can be used as a supervised classification in the processing of remote sensing image classification. But its defects are obvious: falling into the local minimum value easily, slow convergence speed, and being difficult to determine intermediate hidden layer nodes. Genetic algorithm (GA) has the advantages of global optimization and being not easy to fall into local minimum value, but it has the disadvantage of poor local searching capability. This paper uses GA to generate the initial structure of BPNN. Then, the stable, efficient, and fast BP classification network is gotten through making fine adjustments on the improved BP algorithm. Finally, we use the hybrid algorithm to execute classification on remote sensing image and compare it with the improved BP algorithm and traditional maximum likelihood classification (MLC) algorithm. Results of experiments show that the hybrid algorithm outperforms improved BP algorithm and MLC algorithm.

1. Introduction

Satellite remote sensing data is widely used in resource exploration, military reconnaissance, environmental disaster monitoring, land use, crop yield assessment, urban planning, and many other areas [1]. Classification of remote sensing image has been a focus and a difficulty in the field of remote sensing and a critical step of transformation from remote sensing technology to practical application [2]. The most commonly used classification processing method of remote sensing image is statistical pattern recognition method which is based on the spectral characteristics of the image data [3]. But sometimes the classification effect of this method is not very satisfactory. It often makes people see the confusing results because the method is only based on the statistical characteristics of gray-scale data of each band [4].

Jia [5] used BP algorithm to classify three-band RGB color images for automatic visual inspection of seed maize and compared it with minimum distance and maximum likelihood classifications. Xiong et al. [6] combined BPNN and ground truth database and attained good classification effect. Mousavi Rad et al. [7] used BPNN classifier to classify rice varieties. In order to classify, they extracted twenty-two features from sixty color and texture features. Experiments showed that classification result was satisfactory. Yang et al. [8] and Liu et al. [9] applied improved BPNN algorithm to classify TM image and achieved good classification result.

Using BPNN algorithm for classification of remote sensing image can eliminate fuzziness and uncertainty to some extent. But simple BPNN algorithm used in remote sensing image classification has great limitations mainly in five ways [10, 11]. Firstly, BP algorithm learning process has the possibility of falling into local minimum easily, so it cannot guarantee that the network's learning process always tends to be in globally stable state. Secondly, BP algorithm has some defects such as slow convergence velocity and requiring a large amount of training samples. Thirdly, the determination of learning rate of network depends on one's experience. Furthermore, the choice method of learning rate can directly influence the stability and efficiency of the learning process. Fourthly, determination of the number of nodes in the hidden layer is based not on some theory but on one's experience.

Fifthly, there exists overfitting phenomenon in the process of learning. Classification errors generate as thus [12–14].

On the contrary, GA has advantages of strong ability of global optimization and being not easy to fall into local minima. But it has poor ability of local optimization. Combined with BP algorithm's strong ability of partial optimization, Hybrid algorithm is constructed. Cao and Jin [15] fused Landsat ETM+ image and ERS-2 SAR image with principal component analysis (PCA) and constructed a hybrid algorithm of BP-ANN/GA. Then they applied the hybrid algorithm to classify urban terrain surfaces in Pudong New Area of Shanghai, China. They got the satisfactory result especially in building and road classification. But in his hybrid algorithm, traditional BPNN with one hidden layer and twelve nodes of hidden layer was adopted. It has weak points of slow adjustment of learning rate, being fit for urban area with less ground features.

In this paper, we constructed a hybrid algorithm of GABPNN with GA and improved BPNN instead of traditional BPNN. Initial structure of BPNN is generated by using GA. It makes the best of both algorithms. By using this hybrid algorithm, we got the excellent classification results on ETM+ remote sensing image.

2. Data Source

The data in this study comes from American Landsat-7 satellite Enhanced Thematic Mapper plus (ETM+) image. The image acquisition time is January 10, 2003. The image has a spatial resolution of 30 meters; and the image coverage of the area is Shunde District, Foshan City, Guangdong Province, China. In order to fulfill research needs, 500 × 500-pixel subimages were tailored from the original image as a study area. Because the sixth band of ETM+ images is thermal infrared band and it has a resolution of 60 meters, meanwhile, the eighth band is panchromatic wave band and it has a resolution of 15 meters, the two bands have been removed [16, 17]. According to the information entropy of the remaining six bands and the correlation between each original image band, this paper chooses the third, fourth, and fifth bands to composite pseudocolor image because of their rich information and small correlation [18]. Figure 1 shows the pseudo-color composition image with band 5, 4, and 3.

3. Methodology

By means of visual interpretation and geographical knowledge of land use in Shunde, the study area can be classified as seven kinds of ground objects (Table 1) [2], that is, farmland (C1), woodland (C2), grass (C3), water (C4), residential area (C5), traffic land (C6), and other construction lands (C7).

A total of 1400 samples were gotten from the land use classification map of the same period. Each category has 200 samples. Among these 200 samples, 80 were used for training and 120 were used to validate the classification accuracy. In order to carry out a comparative study, three classification methods were implemented in this paper. The first method is the traditional maximum likelihood classification (MLC).

TABLE 1: Land cover types in classification.

Class name	Land-cover class	Description
C1	Farmland	Crop fields and vegetable fields
C2	Woodland	Deciduous or evergreen forest land, orchards and tree groves
C3	Grass	Lawn, grassland or rangeland
C4	Water	Reservoir, pond, lake, river and wetland
C5	Residential area	Rural and urban residences, commercial land
C6	Traffic land	Expressway, road, railway, urban road, country road
C7	Other construction lands	Industrial and other developed lands

Figure 1: Pseudo-color composition image in Shunde district of Guangdong, China.

The method is realized by software of the Environment for Visualizing Images (ENVI), and the version of ENVI is 4.8. The second classification method is improved BP neural network algorithm (BPNN). A 6-9-7 three-tier network architecture was adopted. That is to say, the input layer has six nodes, namely, band numbers of the remote sensing image. The hidden layer number is one and the node number of this layer is 9 by experience. The output layer node number is the same as classification category number, namely, 7. The classification method is realized by the software of Matlab 2011b. The third classification method is GABPNN hybrid algorithm. GA is used here for getting initial structure of BPNN, and classification is realized by modified BPNN. The input nodes of the BP network are band numbers, namely, 6. The output nodes are the category numbers, namely, 7. Layers number of hidden layers and nodes number of each hidden layer, thresholds, and connecting weights of each node in hidden or output layer are unknown. These unknown parameters are

determined through hybrid algorithm. The hybrid algorithm is realized by the software of Microsoft Visual C++ 2008.

4. Algorithm Principle

4.1. Improved BP Algorithm. The dynamic adaptive adjustment for learning rate is adopted in this paper to improve efficiency of common BP algorithm. Learning or training of BPNN is process of modification on weight vector of network repeatedly according to least mean square error principle, where the squared error is between the desired output and the actual output of the network. The learning process will not stop until the output value of the network is close to the desired output value. In other words, the total error is less than a given number or the iteration program ceases at the given maximum cycle times. For a training sample p, the square error E_p is defined as follows:

$$E_p = \frac{1}{2} \sum_k \left(t_{pk} - o_{pk}\right)^2, \tag{1}$$

where t_{pk} represents the desired output and o_{pk} represents the actual output, and the total mean square error of the whole BPNN is expressed as

$$E = \frac{1}{n} \sum E_n, \tag{2}$$

where n is the total number of training samples. Adjustment for weight vector of the output layer and the hidden layer is carried out by

$$w(k+1) = w(k) + \eta(k)D(k) + \alpha(w(k) - w(k-1)), \tag{3}$$

where $w(k)$ and $w(k-1)$ represent weight vectors of k moment and before, respectively. Momentum factor α is used to prevent the learning process from falling into a local minimum. Dynamic adjustment formula for learning rate η is as follows:

$$\eta(k) = 4^\lambda \eta(k-1). \tag{4}$$

$\eta(k)$ is learning rate for network of k moment. In formula (4), λ refers to adjustment coefficient, where

$$\lambda = \text{sign}(D(k)D(k-1)). \tag{5}$$

$D(k)$ is the negative gradient of k moment, which is defined as

$$D(k) = \left(-\frac{\partial E(w)}{\partial w}\right)\bigg|_{w=w(k)}. \tag{6}$$

So there are two situations of λ value. If two consecutive iterations have the same gradient direction, it means that the decline speed is too slow; then λ equals to +1. The learning rate increased to four times. While if two consecutive iterations have contrary gradient direction, it means that the decline speed is too fast, then λ equals to −1. The learning rate decreased to 25 percent. Such adjustment of the learning rate not only can avoid falling into local minima and

misconvergence, but also can help to shorten the learning or training time.

Sigmoid function is chosen as activation function of nodes for hidden layer and output layer. Consider

$$f(x) = \frac{1}{1 + \exp(-X)}, \tag{7}$$

where X is the sum of all input value multiplied by the corresponding weight plus the threshold value of the node. Using sigmoid type excitation function as incentive function can ensure that the function is nonlinear and continuously differentiable.

However, net input value of the node j for hidden or output layer is

$$\text{net}_j = \sum_j w_{ji} x_i. \tag{8}$$

Therefore the output of the node j is

$$o_j = f\left(\text{net}_i + b_j\right) = f\left(\sum_j w_{ji} x_j + b_j\right). \tag{9}$$

Thus the threshold b_j of node j can be regarded as a special weight vector, and adjustment of the threshold can also use (3).

4.2. Genetic Algorithm. Genetic algorithm is used to obtain the initial structure of the improved BP network.

4.2.1. Encoding. In this study, floating-point encoding is adopted. Individual gene encoded string is composed of all input weights, thresholds, learning rate, momentum factor, layers number of hidden layers, and nodes number of each hidden layer. Initial population of individuals is generated randomly. Population size of 40 is more appropriate in this study because a larger size increases the training time and the global optimal solution cannot be gotten by a smaller size.

4.2.2. Fitness Function. According to principle of single value, continuity, nonnegative value, maximization, rationality, consistency, compatibility, and small amount of calculation for the fitness function [10] and because the network training goal is minimum mean square error (MSE), the minimum value problem can be converted into the maximum value problem, so the fitness function of individual can be expressed as the following:

$$\text{fitness}(x) = \frac{1}{1 + E(x)}, \tag{10}$$

where $E(x)$ represents the total MSE of the output according to BP network structure of decoding of the individual x. Hence, the fitness of an individual is determined by the total MSE. The bigger the error is, the smaller the fitness is.

4.3. Genetic Operators

(1) *Selection.* The best individual is saved preferentially and the remaining individuals are selected stochastically. Namely,

the individual with maximum fitness value is preserved firstly, and then the remaining individuals are chosen at random. The selection method provides a zero-deviation and minimum individual extension. Individuals of small fitness value may be selected under the premise of reservation of the best individual, so that population can keep diverse.

(2) *Crossover*. For this asexual reproduction operator, owing to floating-point encoding for individual, heterogeneous arithmetical crossover is adopted in this study. Two parent individuals X_A^t and X_B^t are selected from the population by using the stochastic selection operation firstly, and then two new offspring individuals are created by crossover operation as

$$
\begin{aligned}
X_A^{t+1} &= \gamma X_B^t + (1 - \gamma) X_A^t, \\
X_B^{t+1} &= \gamma X_A^t + (1 - \gamma) X_B^t,
\end{aligned} \tag{11}
$$

where parameter $\gamma \in (0, 1)$ is not a constant but a variable determined by the current evolutional generation. Value of γ can be expressed as

$$
\gamma = \frac{1}{1 + g}, \tag{12}
$$

where g is number of current evolutional generation. Arithmetical crossover probability p_c plays a dominant role in GA from the beginning to the end. It directly affects the convergence of the algorithm. The greater p_c value is, the faster new individuals generate. But if it is too large, then the genetic model will be destroyed easily. Therefore it generally ranges between 0.4 and 0.99. In this study p_c takes a value of 0.66.

(3) *Mutation*. Uniform mutation method is used in this study. Firstly, the parent individual $X(x_1, x_2, \ldots, x_n)$ is chosen by selection operator before mutation operation. Then a random single point component $x_i, i \in [1, n]$, from the coding of the parent individual X is selected randomly. Accordingly, x_i is replaced by $x_i', x_i' = x_i(1 - \gamma)$, where γ is the same of (12). Mutation probability p_m recommended generally ranges between 0.0001 and 0.1. In this study p_m is given value of 0.005.

The property of combining these crossover and mutation operators can make a uniform search in the initial space in early generations and very locally at a later stage, favoring local tuning. It also greatly reduces the risk of premature convergence.

4.4. Hybrid Algorithm for Classification. Before implementing the hybrid algorithm, remote sensing data must be preprocessed. Preprocessing of remote sensing imagery data refers to the normalized processing of remote sensing image data. The normalized transform uses the following formula [19, 20]:

$$
x_i' = \frac{(x_i - M_{\min})}{(M_{\max} - M_{\min})}, \tag{13}
$$

where $M_{\min}$ represents the minimum value of a band of ETM+ image. Correspondingly, $M_{\max}$ is the maximum value of the identical band. Accordingly, all data after normalizing are distributed in the region of 0 and 1. This reduces the solving difficulty to a certain extent.

All weights, thresholds, and other parameters of BPNN are initialized at random firstly. They are encoded as real numbers. Initial population of GA is generated.

Training samples are input, and fitness of GA is calculated. The individual is assessed by program. If the individual gets specified precision, then program exits GA and enters into modified BPNN, or else program executes selection, crossover and mutation operator to generate new species. GA will not exit until program reaches the maximum number of evolution generation or the individual gets specified precision.

In modified BPNN, the best individual is decoded by program, and then initial structure of BPNN is gotten. Weights, thresholds of all nodes, layers number of hidden layers and node number of hidden layers are obtained. Then program executes training. Training will not exit until program reaches the maximum number of iterations or overall error is less than a specified minimum value. Finish of training means that the best stable effective BPNN is constructed.

Finally, program executes classification with trained BPNN and outputs the classified results. Flow chart of GABPNN hybrid algorithm can be seen in Figure 2.

5. Classification Results

Figure 3(a) illustrates the classification result by using conventional MLC algorithm. This algorithm is implemented by ENVI 4.8. In this software, training samples are defined as region of interest (ROI).

Figure 3(b) indicates the classification result by using improved BPNN algorithm. In this classification method, a 6-9-7 three-tier neural network architecture was adopted. The input layer has six nodes, namely, band numbers of the image. There is only one hidden layer and the node number of this layer is 9 by experience. The output layer node number is the same as classification category number, namely, 7. The anticipated outputs for the 7 surface classes: C1, C2, C3, C4, C5, C6, and C7, are defined as (1 0 0 0 0 0 0), (0 1 0 0 0 0 0), (0 0 1 0 0 0 0), (0 0 0 1 0 0 0), (0 0 0 0 1 0 0), (0 0 0 0 0 1 0), and (0 0 0 0 0 0 1), respectively. The error between the real output and the desired output are evaluated. If the BPANN output is beyond the allowed error, then the error feedback is propagating in the opposite direction to adjust the weights until the satisfied results are acquired. We first choose the training data for this modified BPANN from the ETM+ image at six bands and 80 samples for each class. Then the maximum number of iterations is set to 15000, and the initial learning rate is set to 0.05. Momentum factor is set to 0.5. The minimum MSE is set to 0.01. After 14520 times of training, an excellent BPNN is constructed. The training performance convergence curve is shown in Figure 4. Input ETM+ remote sensing image data and the classification result are gotten, as shown in Figure 3(b).
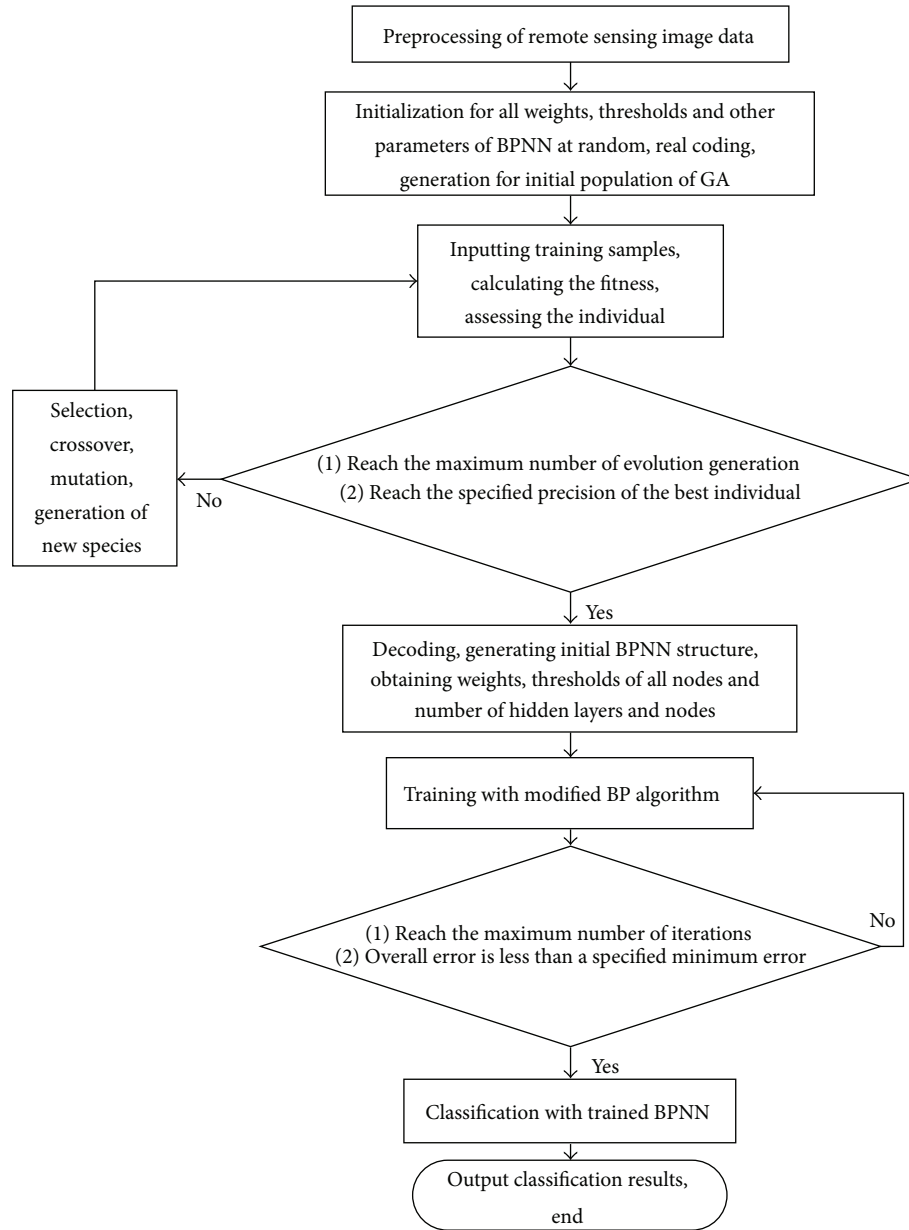
FIGURE 2: Flow chart of GABPNN hybrid algorithm.

Figure 3(c) shows the classification result by using GABPNN hybrid algorithm. First, we used GA to generate the initial structure of BPNN; the population size is set to 40; the crossover probability is set to 0.66; the mutation probability is set to 0.005; the maximum evolution generation is set to 500. The GA program exited after the 458th generation, and then the most suitable 4-layer structure BP network is obtained. In this BP network, there are two hidden layers; the nodes number of hidden layer near the input layer is 15, and the nodes number of hidden layers near the output layer is 18. Therefore a 6-15-18-7 structure of 4-layer BPNN is constructed. Then we utilized the previous improved BP algorithm for getting the exact solution of weights and thresholds. The maximum number of iterations is also set to 15000. The initial learning rate is set to 0.05. The momentum rate factor is set to 0.5. The minimum MSE is set to 0.01. After 11200 times of training, an efficient stable BPNN is constructed. The hybrid algorithm training performance convergence curve is shown in Figure 5. Using this fine-tuned BPNN, we obtained the classified result as shown in Figure 3(c).

6. Classification Assessment

In order to further verify the correctness of these three types of remote sensing image classification algorithm, we carried out the quantitative comparison for their classification accuracy [15]. The method is conventional, that is, confusion
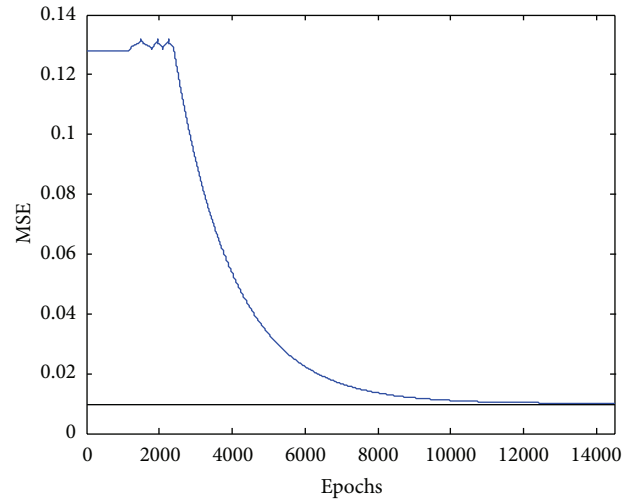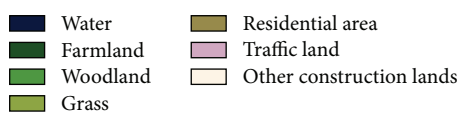
(a)

(b)

▮ Water	▮ Residential area
▮ Farmland	▮ Traffic land
▮ Woodland	▯ Other construction lands
▮ Grass	

(c)

FIGURE 3: Classified images derived from (a) MLC, (b) BPNN, and (c) hybrid algorithm of GABPNN.

FIGURE 4: Training performance of modified BPNN.

TABLE 2: Comparison of production and user accuracies, as a percentage.

Accuracy	C1	C2	C3	C4	C5	C6	C7
Production accuracy							
MLC	70.8	82.5	73.3	88.3	91.7	62.5	74.2
BPNN	75.0	83.3	77.5	91.7	91.7	66.7	75.0
GABPNN	79.2	91.7	87.5	95.8	92.5	75.0	81.7
User accuracy							
MLC	69.7	75.0	77.2	92.2	71.9	78.1	82.4
BPNN	72.6	78.1	81.6	94.8	72.9	81.6	82.6
GABPNN	84.1	77.5	83.3	99.1	84.7	89.1	88.3

TABLE 3: Comparison of the overall accuracy.

	MLC	BPNN	GABPNN
Overall accuracy	77.62%	80.12%	86.19%
Kappa coefficient	0.739	0.768	0.838

matrix, production and user accuracy, overall accuracy, and kappa coefficient [21].

Every 120 validation samples for each type are selected from existing LULC mapping or field investigation with global positioning system (GPS) receiver. The production and user accuracy, overall accuracy, and kappa coefficient are then computed [22]. They are shown in Tables 2 and 3.

From the previous classification images and accuracy tables, we find that farmland, woodland, and grass are prone to misclassification or leakage points because of the phenomena of "same object with different spectra" and "same spectrum with different objects," whereas water area has relatively high accuracy in three classification methods. The reason is that the spectral curves of water and others are different obviously, while farmland, woodland, and grass have similar or same spectrum. Residential area, traffic land, and

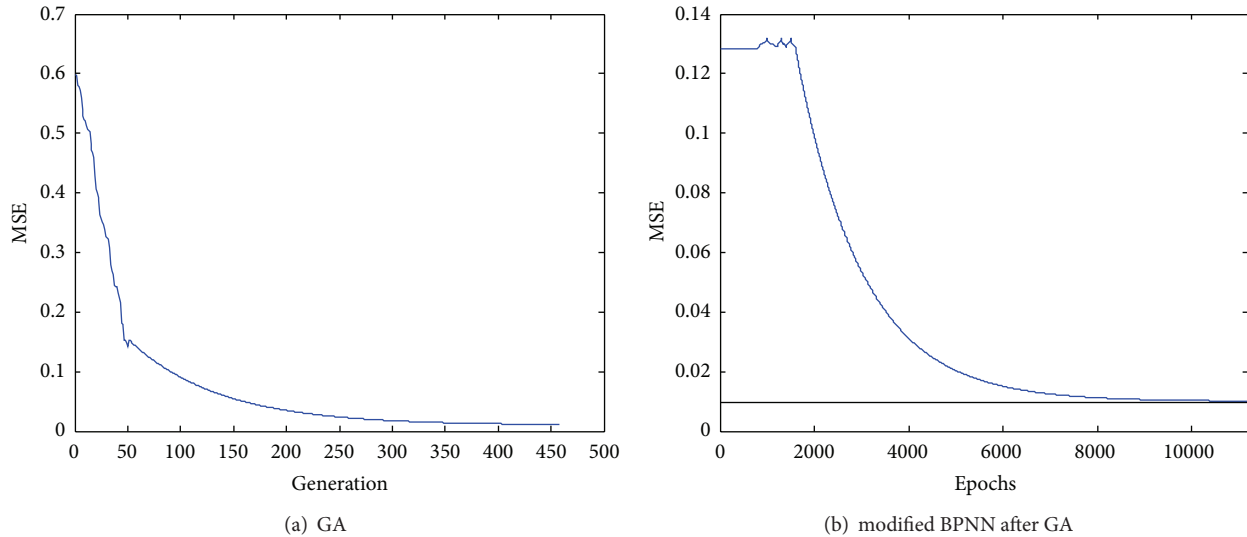(a) GA

(b) modified BPNN after GA

FIGURE 5: Hybrid algorithm training performance.

other construction land are easy to be misclassified because these classes contain large amounts of cement and other building materials which caused their spectra similar or alike.

As shown in Table 2, the user accuracy of farmland in MLC is 69.7%, and it is 72.6% and 84.1% in modified BPNN and GABPNN, respectively. The user accuracy of GABPNN is 14.4% higher than that of MLC, and it is nearly 12% higher than the modified BPNN algorithm, the same for other classification types.

The overall accuracy and kappa coefficient are shown in Table 3. From this table, we can find that the overall accuracy is from 77.62% of MLC to 80.12% of BPNN and to 86.12% of GABPNN hybrid algorithm. It increased significantly, as well as kappa coefficient.

7. Conclusion and Discussion

This paper developed a hybrid algorithm of GABPNN to optimize the network's structure and make fast convergence of BPANN. The algorithm is applied on ETM+ image in Shunde District, Guangdong, China.

(1) The training speed of modified BPNN after GA is speeded up and the classification accuracy for all surfaces is improved.

(2) Applying the GA algorithm to initialize the structure of the BPANN can take its advantage of optimization and overcome the shortcomings of the BPANN's slow convergence and falling into the local minima easily.

The probability distribution density of the training data of each type with a normal distribution characteristic is required in the conventional MLC, but the actual chosen sample data may be deviated from the normal distribution; therefore, classification precision of MLC is low relatively.

But on the contrary, there are no strict requirements, no strict limitations, and no need of normal distribution character on training data in BPNN. In addition, BPNN has the ability of complex nonlinear mapping. Hence, it is more suitable for the classification processing of massive high-dimensional remote sensing image data.

On the other hand, the ordinary BPNN classifier has defects of falling into local minima easily, convergence difficultly, and being time consuming, while the GABPNN developed in this study not only avoids shortcomings of BPNN but also improves the accuracy and efficiency of classification. Experience has proved that GABPNN hybrid algorithm is stable and effective. Results of experiments in this paper show that the hybrid algorithm outperforms improved BP algorithm and MLC algorithm. It combines the advantages of both and it can always get the global optimal solution. It is an excellent competent promising algorithm of classification processing for moderate resolution multispectral remote sensing data and it has a strong practical value.

References

[1] J. B. Campbell, *Introduction To Remote Sensing*, Taylor & Francis, London, UK, 2000.

[2] Y. M. Luo, M. H. Liao, J. Yan, C. Y. Zhang, and S. L. Shang, "Development and demonstration of an artificial immune algorithm for mangrove mapping using landsat TM," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 4, pp. 751–755, 2013.

[3] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, John Wiley & Sons, New York, NY, USA, 2nd edition, 2001.

[4] J. A. Richards and X. P. Jia, *Remote Sensing Digital Image Analysis, An Introduction*, Springer, Berlin, Germany, 3rd edition, 1999.

[5] J. Jia, "Pattern classification of RGB colour images using a BP neural network classifier," *Proceedings of the SPIE-The International Society for Optical Engineering*, vol. 1989, pp. 248–256, 1993.

[6] Z. Y. Xiong, K. Chen, and C. D. Gu, "An algorithm of image classification based on BP neural network," in *Proceedings of the IEEE International Conference on Computer Science and Automation Engineering (CSAE '12)*, pp. 523–526, 2012.

[7] S. J. Mousavi Rad, F. A. Tab, and K. Mollazade, "Classification of rice varieties using optimal color and texture features and BP neural networks," in *Proceedings of the 7th Iranian Conference on Machine Vision and Image Processing (MVIP '11)*, p. 5, November 2011.

[8] B. Yang, Z. Liu, Y. Xing, and C. Luo, "Remote sensing image classification based on improved BP neural network," in *Proceedings of the International Symposium on Image and Data Fusion (ISIDF '11)*, August 2011.

[9] Q. L. Liu, B. L. Jiao, and L. Liu, "Study On remote sensing image classification method based on improved BP neural network model," *Electronics Optics & Control*, vol. 16, no. 8, pp. 65–67, 2009.

[10] F. Yu, Y. S. Zhao, and H. T. Li, "Soil moisture retrieval based on GA-BP neural networks algorithm," *Journal of Infrared and Millimeter Waves*, vol. 31, no. 3, pp. 283–288, 2012.

[11] D.-Y. Ge, X.-F. Yao, and W.-J. Xiang, "Application of BP neural network and genetic algorithm in testing of micro-drill's rounded corner," *Sensor Letters*, vol. 9, no. 5, pp. 1943–1947, 2011.

[12] C. F. Luo, C. Y. Wang, and Y. H. Liu, "Classification of land use in Xinjiang with BP classifier using MODIS imagery," *Arid Land Geography*, vol. 28, no. 2, pp. 258–262, 2005.

[13] S. Yin, N. Shu, and X. Liu, "Classification of remote sensing image based on adaptive GA and improved BP algorithm," *Geomatics and Information Science of Wuhan University*, vol. 32, no. 3, pp. 201–204, 2007.

[14] J. F. Mas and J. J. Flores, "The application of artificial neural networks to the analysis of remotely sensed data," *International Journal of Remote Sensing*, vol. 29, no. 3, pp. 617–663, 2008.

[15] G. Cao and Y. Q. Jin, "A hybrid algorithm of the BP-ANN/GA for classification of urban terrain surfaces with fused data of Landsat ETM+ and ERS-2 SAR," *International Journal of Remote Sensing*, vol. 28, no. 2, pp. 293–305, 2007.

[16] B. Guo, S. R. Gunn, R. I. Damper, and J. D. B. Nelson, "Band selection for hyperspectral image classification using mutual information," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 4, pp. 522–526, 2006.

[17] L. Zhang, Y. Zhong, B. Huang, J. Gong, and P. Li, "Dimensionality reduction based on clonal selection for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 12, pp. 4172–4186, 2007.

[18] F. A. Kruse, A. B. Lefkoff, J. W. Boardman et al., "The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment*, vol. 44, no. 2-3, pp. 145–163, 1993.

[19] F. Ratle, G. Camps-Valls, and J. Weston, "Semisupervised neural networks for efficient hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 5, pp. 2271–2282, 2010.

[20] G. A. Carpenter, M. N. Gjaja, S. Gopal, and C. E. Woodcock, "Art neural networks for remote sensing: vegetation classification from landsat tm and terrain Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 2, pp. 308–325, 1997.

[21] H. Li, Y. Wang, Y. Li, and X. Wang, "Pixel-unmixing moderate-resolution remote sensing imagery using pairwise coupling support vector machines: a case study," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4298–4306, 2011.

[22] X. Liu, X. Li, L. Liu, J. He, and B. Ai, "An innovative method to classify remote-sensing images using ant colony optimization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 12, pp. 4198–4208, 2008.