

Received April 21, 2020, accepted May 23, 2020, date of publication May 27, 2020, date of current version June 8, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2997912

Hyperspectral Remote Sensing Image Classification With CNN Based on Quantum Genetic-Optimized Sparse Representation

HUAYUE CHEN^{ID}^{1,2,3}, FANG MIAO^{ID}^{1,4}, (Member, IEEE),
AND XU SHEN^{ID}², (Member, IEEE)

¹School of Geophysics, Chengdu University of Technology, Chengdu 610059, China

²School of Computer Science, China West Normal University, Nanchong 637002, China

³Internet of Things Perception and Big Data Analysis Key Laboratory of Nanchong, Nanchong 637002, China

⁴Key Laboratory of Pattern Recognition and Intelligent Information Processing of Sichuan, Chengdu University, Chengdu 610041, China

Corresponding author: Huayue Chen (sunnyxiaoyue@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771087, in part by the Key Projects of Sichuan Provincial Department of Education under Grant 16ZA0177, in part by the Talents Fund of China West Normal University under Grant 17YC147, and in part by the Scientific Research Fund of Key Laboratory of Pattern Recognition and Intelligent Information Processing of Chengdu University under Grant MSSB-2019-09.

ABSTRACT Due to the characteristics of the spectrum integration, information redundancy, spectrum mixing phenomenon and nonlinearity of the hyperspectral remote sensing images, it is a major challenging task to classify the hyperspectral remote sensing images. Therefore, a hyperspectral remote sensing image classification method, named QGASR-CNN is proposed in this paper. In the QGASR-CNN, a quantum genetic-optimized sparse representation method is designed to obtain the over-complete dictionary with sparsity, and achieve the feature sparse representation to construct the sparse feature matrix of hyperspectral remote sensing image pixel groups. Then the convolution neural network(CNN) directly convolutes with image pixels to build the feature mapping relation by using convolution operation. Finally, in order to testify the effectiveness of the QGASR-CNN, the actual hyperspectral remote sensing image datasets are selected in here. The comparison results show that the QGASR-CNN sparsely represents the features of hyperspectral remote sensing images and improves the classification accuracy. It can effectively alleviate the problems of the small samples and ‘salt and pepper misclassification’.

INDEX TERMS Hyperspectral remote sensing, image classification, sparse representation, convolutional neural network, quantum genetic.

I. INTRODUCTION

A hyperspectral sensor obtains approximately continuous spectral curve of the ground object in ultraviolet, visible, near infrared, mid infrared, and other electromagnetic wave bands. It combines the spectral information reflecting the reflection characteristics of ground object with the image information reflecting the spatial position relationship of ground object. It is composed of dozens or even hundreds of continuous wave band images into a three-dimensional data cube, which has the characteristics of combination of image and spectrum, and it has been widely used in resource

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

exploration, environmental monitoring, precision agriculture, disaster assessment, target recognition, and other fields [1]–[3]. Hyperspectral remote sensing image classification is based on the recognition information with specific meaning learned from the original spectral information, and each pixel is accurately divided into its own feature categories. Compared with general remote sensing images, hyperspectral remote sensing images have many characteristics, such as a large number of bands, fine and continuous spectrum, large data volume and information redundancy, spectral mixing phenomenon and data nonlinearity [4]–[6]. The data and the spatial structure is relatively complex and the dimension is very high. Traditional image processing and classification methods are difficult to achieve classification results. As a

typical ultra-high dimensional data, how to mine the diagnostic information of interested objects from the massive high dimensional data has become a bottleneck in the applications of hyperspectral remote sensing image.

In recent years, a lot of research on feature extraction and classification of hyperspectral images have been conducted by using sparse representation and convolution neural network(CNN) theory. Each test sample to be classified is taken as a linear combination of training samples [7]. Then the sparse coefficient is solved, and the classification is determined by minimizing residual between the test sample and all kinds of training samples, but only the spectral information is used in the feature extraction, resulting in the non-smoothness of the classification results. A joint sparse representation model based on the maximum likelihood estimation is proposed [8]. The traditional quadratic loss function is replaced by a coding residual function used to measure the joint approximation error class. The coding residual estimation transforms the traditional maximum likelihood estimation into an iterative weighted joint sparse representation. This method can reduce the number of non-uniform neighborhood pixels. However the robustness of outliers is poor. Besides, the iteration is slow due to the high dimension and large amounts of data, which reduces the efficiency. The group sparse principal component transformation is studied [9]. The signal is represented by the eigenvector of the covariance matrix of the training sample. The eigenvector is regarded as the base atom of the signal, and the atoms are orthogonal to each other to maintain the main components of the original spectrum. But the sparse prior is not used explicitly. Aiming at the problem that the independence of sparse reconstruction process will result in the loss of similarity information of coefficients between similar samples [10], a regularized sparse representation algorithm is proposed, which combines the centralized quadratic constraint as a regular term into the objective function of sparse representation to retain the similarity information. However, the solution of sparse representation ignores the Euclidean distance relationship between samples. The hyperspectral remote sensing image is classified by using the deep CNN network model [11]. It obtains the spectral vectors of each pixel, extracts the local spectral information on the spectral vector by using CNN convolution layers, and takes the feature map generated by convolution operation as the input of the full connection layer, and finally completes the hyperspectral remote sensing image classification. This method makes use of CNN's local connection, weight sharing and other characteristics, greatly reduces the model parameters and training difficulty, further improves the classification performance, but fails to make full use of the rich space spectrum information of hyperspectral remote sensing image, reducing the ability of feature learning. In reference [12], the spectral information, spatial structure information and semantic context awareness information of each pixel are represented jointly by different regions, and the features of multiple scales and different layers are combined from different layers for depth feature extraction.

However, this method divides the original image into six patches, which are global, top, bottom, left, right and center. Each patch corresponds to a CNN structure. After the feature extraction of the six CNN branches, the whole connection operation is carried out. The input information of CNN is not positively related to the classification effect. Complex inputs enlong the training time and classification time, which easily leads to over-fitting. After extracting spatial features from principal components analysis (PCA) dimensionally reduced hyperspectral image data by CNN [13], the spatial features extracted by CNN are further processed by sparse coding. Finally, the classification task is completed based on the sparse coding features. But it only considers the spatial features of hyperspectral remote sensing image, not its rich spectral features, and the accuracy of classification needs to be improved. A new pixel pair method is proposed to increase the number of CNN training samples [14]. For a test pixel, the trained CNN classifies the central pixel and each surrounding pixel, and then determines the final label through avoidance strategy. In this method, CNN is used to learn the features of pixel pairs, which has better classification performance. However, the generalization ability of CNN model is insufficient, and the accuracy of the classification needs to be improved because of the unclear quantification of pixel pairs.

To sum up, at present, only sparse representation is used to reduce dimension and extract features of hyperspectral remote sensing images, and then traditional SVM, random forest [15]–[17] and other classifiers are used to complete the classification of hyperspectral images and hyperspectral image classification by using CNN alone. Although certain achievements have been achieved, most of them do not use sparse prior explicitly, resulting in ‘salt pepper misclassification’. The extraction of joint features of space spectrum and the utilization of spatial information still need to be diversified, and there are deficiencies in the mining of hyperspectral data structure information. On the other hand, the traditional dimensionality-reduction methods are used to participate in the convolution operation of CNN models [18]–[20].

Therefore, in view of the characteristics of hyperspectral remote sensing image, such as the number of bands, the combination of atlas, information redundancy, spectral mixing and data nonlinearity, sparse representation with quantum genetic algorithm and convolutional neural network are integrated in order to propose a novel QGASR-CNN method, which is used to effectively describe the features and achieve The classification for hyperspectral remote sensing image. The actual hyperspectral remote sensing data are used to verify the effectiveness of the QGASR-CNN.

The presented research is organized as follows: Section 2 introduced basic methods of quantum genetic algorithm, sparse representation and convolutional neural network. In the Section 3, the proposed QGASR-CNN is described and the contribution is highlighted. In the Section 4, the experimental verification and analysis are provided.

The Section 5 is given to conclude the QGASR-CNN and suggest some works in the future.

II. BASIC METHODS

A. QUANTUM GENETIC ALGORITHM

Quantum genetic algorithm (QGA) is a probabilistic search optimization method that combines quantum computing theory with genetic algorithm. It has better population diversity and computational parallelism, faster convergence speed, higher search efficiency and stronger global optimization ability [21]–[23]. The QGA introduces the concept and principle of quantum computation into the genetic algorithm, and adopts chromosome coding based on Q-bit. The Q-bit in QGA is a unit vector defined in a two-dimensional complex vector space. A Q-bit can be expressed as $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$, where α and β are probability amplitudes of the corresponding states of Q-bit. $|\alpha|^2$ and $|\beta|^2$ represents the probability that the quantum state is observed as “0” state and “1” state, respectively. The normalization of states needs to meet $|\alpha|^2 + |\beta|^2 = 1$. In order to find the optimal solution, we apply the representation of the Q-bit probability amplitude to chromosome coding, quantum rotation gate and other operations to chromosome update.

B. SPARSE REPRESENTATION THEORY

As a state-of-the-art data mining technology, sparse representation can effectively extract the hyperspectral terrain information by using the high redundancy of massive high-dimensional data and the sparsity of interested signals [24]–[26]. According to the sparse representation theory, for a given dictionary, each signal can be expressed by the linear combination of a few primitives in the dictionary. Given the image data set $\{x_1, x_2, \dots, x_m\}$, its sparse representation mathematical model can be described as follows.

$$\min_{D, a_i} \sum_{i=1}^m \|x_i - Da_i\|_2^2 + \lambda \sum_{i=1}^m \|a_i\|_1 \quad (1)$$

where $D \in G^{d \times k}$ is the dictionary matrix, k is the sparsity, λ is the regularization parameter, and $a_i \in G^k$ is the sparse representation of the sample $x_i \in G^d$. Sparse representation learning of hyperspectral remote sensing image uses optimization algorithm to obtain sparse representation coefficient a_i , which can well reconstruct x_i when a_i is as sparse as possible.

C. CONVOLUTIONAL NEURAL NETWORK

In hyperspectral remote sensing image processing, a convolutional neural network (CNN) makes full use of the information of the adjacent areas in the image, greatly reduces the scale of parameters, reduces the complexity of calculation and improves the convergence speed of the network by sparse expression and weight sharing. It is usually composed of convolution layers, pooling layers, full connection layers and an output layer [27]–[29]. Assuming that the pixel value of the sub image of the input image at (i, j) position is p_{ij} the element value of the convolution kernel matrix at (x, y) position is k_{xy} ,

and the convolution kernel acts on a certain position of the image, the output is given by

$$W_{ij} = f \left(\sum_{x=1}^s \sum_{y=1}^s p_{i+x-1, j+y-1} \cdot k_{xy} + b \right) \quad (2)$$

where f is the activation function and b is the offset term. The CNN can be regarded as a kind of neural network with weight sharing and local connection. According to the previous definition, the forward propagation formula of convolution layer is described as follow, i.e.

$$p_{ij}^{(l)} = f \left(u_{ij}^{(l)} \right) = f \left(\sum_{x=1}^s \sum_{y=1}^s p_{i+x-1, j+y-1}^{(l-1)} \times k_{xy}^{(l)} + b^{(l)} \right) \quad (3)$$

By using the same convolution kernel, i.e., weight sharing and local connection, the CNN greatly reduces the trainable parameters in the network, the complexity of the model, and the risk of over fitting, so as to obtain better generalization ability.

III. A NEW HYPERSPECTRAL REMOTE SENSING IMAGE CLASSIFICATION

A. THE BASIC THOUGHT

Hyperspectral remote sensing images have the characteristics of numerous bands, fine and continuous spectrum, integrated spectrum, large data volume and redundant information, spectral mixing phenomenon and non-linear data. The spatial structure of images is complex and the dimension is higher. Therefore, traditional image processing and classification methods are difficult to effectively obtain better classification results. Sparse representation can represent the image features with as few atoms as possible and describe the image nature with less characteristic coefficients in a given over-complete dictionary. Quantum genetic algorithm combines quantum computing with genetic algorithm, which has better population diversity and computing parallelism, and takes on faster convergence speed, higher search efficiency and stronger global optimization ability. Convolution neural network is a kind of feed forward neural network, which contains convolution calculation and has deep structure. It has stronger classification ability according to the hierarchical structure. Therefore, in order to improve the classification accuracy and efficiency, based on the manifold structure, polymorphism and low rank of hyperspectral remote sensing images, the quantum genetic algorithm is used to sparsely decompose the over-complete dictionary of hyperspectral image according to the space-spectrum characteristics and sparse priori of pixels. Then the feature sparse representation coefficient of the central pixel is obtained by using the spatial continuity constraints, and the spectral information in the image is combined. The spatial structure and sparse coefficient are used to obtain the sparse representation vector of each pixel in order to construct the sparse feature matrix of hyperspectral image pixel group. Finally, the sparse feature matrix is regarded as the inputs of convolution kernel to establish the feature mapping relation of pixels. The novel

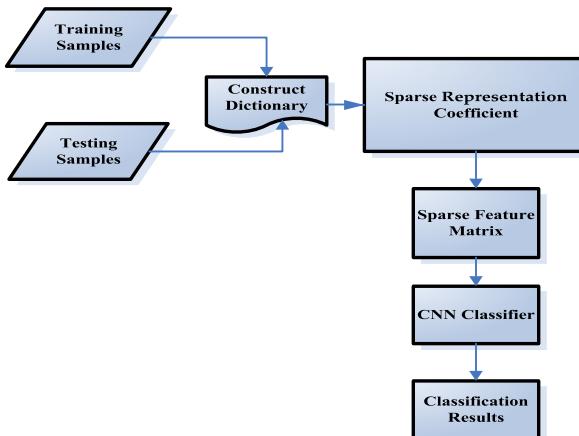


FIGURE 1. The flow of the QGASR-CNN method.

QGASR-CNN method is proposed to effectively achieve The classification of hyperspectral image.

B. OVERVIEW OF QGASR-CNN

The proposed QGASR-CNN method includes dictionary construction, sparse coefficient solution and CNN classifier. Firstly, the data is selected, and the over-complete dictionary is constructed by learning. Then, the sparse representation with quantum genetic algorithm is used to extract the features and construct sparse feature matrix. Finally, the sparse feature matrix is considered as the inputs of the CNN classifier, and the scene classification results is obtained by Softmax function. The flow of the QGASR-CNN is shown in Figure 1.

C. THE REALIZATION OF QGASR-CNN

1) CONSTRUCT DICTIONARY

The sparse representation of hyperspectral image is to decompose the hyperspectral image under the over-complete dictionary, and obtain the feature sparse representation. It mainly includes over-complete dictionary design and sparse decomposition algorithm. K-SVD algorithm is usually used to design over-complete dictionary [31]. It uses the original image to construct training samples, which can obtain redundant dictionary to be suitable for the original signal through self-adaptive learning. To a large extent, it guarantees the sparse representation effect of the over-complete dictionary. In order to make the pixel image objects are as close to sparse representation as possible, and fully reflect the local structural features among the pixels, it is necessary to completely reconstruct the original image information. In hyperspectral image processing, any given test pixel can be represented by a given set of labeled pixels with sparse linearity by using l_0 - norm or l_1 - norm regularization operator. For a given test pixel x , the sparse representation aims to obtain the weight vector α of the linear combination. At the same time, it meets the minimum reconstruction error $\|x - Da\|_2^2$ and the minimum sparse constraint $\|\alpha\|_1$. It can be expressed as $\text{argmin} \|x - Da\|_2^2 + \lambda_1 \|\alpha\|_1$, where λ_1 is the regularization parameter.

2) SPARSE REPRESENTATION

The parameter selection of sparse decomposition greatly affects the performance of sparse representation and the effect of feature extraction. At present, the parameter of sparse decomposition mainly depends on the experiences, which is complex and diverse, and has a wide range of values. Therefore, the parameter value obtained by experiences cannot be the optimal value. Quantum intelligent optimization algorithm combines quantum theory with intelligent computing, makes up for the shortcomings of traditional intelligent optimization algorithm. It has the characteristics of quantum parallel computing, accelerates the convergence speed, and avoids premature phenomenon. Compared with the traditional optimization algorithm, quantum evolution algorithm has good population dispersion, strong global search ability and fast search speed. Therefore, the QGA is applied to solve the problem of sparse representation parameter optimization. In the QGA, the representation of Q-bit probability amplitude is applied to chromosome coding, quantum revolving gate and other operations are used to update chromosomes, so as to find the optimal solution, optimize the parameter values of sparse representation, improve the performance of sparse representation and the sparse representation effect of hyperspectral images. The parameters of the sparse representation mainly include the number of atoms M in the over-complete dictionary, the maximum allowable error ε , the sparsity K , and the number of iterations n . With the increase of the number of dictionary atoms, the adaptability of the dictionary is stronger, but it also increases the capacity for storing the dictionary, reduces the sparsity of pixel data, and increases the complexity of operation. When the number of dictionary atoms is too large, the sparse representation of image loses meaning. Therefore, the QGA adopts the coding method based on Q-bit. Q-bit is a unit vector to define a two-dimensional complex vector space. A Q-bit can be expressed as $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$, α and β are probability amplitudes of corresponding states of Q-bit. $|\alpha|^2$ and $|\beta|^2$ represent the probability that the quantum state is “0” state and “1” state, respectively. The normalization of states must meet the requirements $|\alpha|^2 + |\beta|^2 = 1$. In this case, an m -quantum state can be expressed as follow.

$$\begin{cases} q_j = \left[\begin{array}{c|c|c|c} \alpha_1 & \alpha_2 & \cdots & \alpha_m \\ \beta_1 & \beta_2 & \cdots & \beta_m \end{array} \right] \\ |\alpha|^2 + |\beta|^2 = 1, \quad i = 1, 2 \dots m \end{cases} \quad (4)$$

Quantum population based on Q-bit coding can be defined as $Q(t) = (q_1^t, q_2^t \cdots q_m^t)$, where n represents the population size, t represents genetic algebra, and q_j^t represents a quantum chromosome, which is defined as follow.

$$q_j^t = \begin{bmatrix} \alpha_1^t & \alpha_2^t & \cdots & \alpha_m^t \\ \beta_1^t & \beta_2^t & \cdots & \beta_m^t \end{bmatrix} \quad (5)$$

where m is the number of Q-bit, i.e. the length of the quantum chromosome, which can be represented 2^n states. That is to

say, a chromosome represents the superposition of multiple states. In addition, because the norm square of probability amplitude tends to 0 or 1, the experience tends to a single state, which ensures the good convergence of QGA. The basic steps of parameter optimization for sparsity coefficients are described as follows.

Step 1: Initialize test sample X , training dictionary $A = [\alpha_1 \alpha_2 \cdots \alpha_N]$, sparsity K , initial parent chromosome $\gamma_0 = X$, initial population $Q(t) = (a_1^t, a_2^t \cdots a_m^t)$, optimal individual $A_0 = \emptyset$, initial iteration value T .

Step 2: The binary solution set of $Q(t)$ state $(\alpha_t \beta_t)$ is obtained by the spatial and spectral information of the pixel. A quantum chromosome q_i^t is defined according to the formula (5).

Step 3: Save the current optimal solution to measure each chromosome gene position (Q-bit), and obtain a state according to the formula(4).

Step 4: Calculate the fitness value for each state. There is $\gamma_k = \arg \max_{i=1 \dots k} |(\gamma_{k-1} q_i)|$, the fitness function is used to evaluate the fitness value of each individual and record the best individual γ_k .

Step 5: Add the best individuals of genetic evolution into the training dictionary.

$$A_k = A_{k-1} \cup \{\gamma_k\} \quad (6)$$

Step 6: The population renewal of each generation of chromosomes is carried out by quantum revolving gate $U(0.025\pi)$.

$$U(0.025\pi) = \begin{bmatrix} \cos 0.025\pi & -\sin 0.025\pi \\ \sin 0.025\pi & \cos 0.025\pi \end{bmatrix} \quad (7)$$

$$\gamma_k = X - A_k(UA_k)^{-1}UX \quad (8)$$

Step 7: Determine whether the end conditions are met. If the end condition is met, execute Step 8. Otherwise, there is $k = k + 1$, return to Step 4.

Step 8: Output sparse vectors.

$$\hat{a}(x) = A_k(UA_k)^{-1}UX \quad (9)$$

Because the spectral curve structure of the same pixel is very similar, the obtained coefficients from the over-complete dictionary are also similar. $\hat{a}(x)$ can be regarded as a new representation of pixel x ($x \in X$). That's to say, the sparse representation of pixel is x in the over-complete dictionary A .

3) A NOVEL SCENCE CLASSIFICATION METHOD

When the CNN is used to classify the scenes of hyperspectral images, its depth is the main factor for affecting the classification accuracy. Experimental results show that when a certain number of layers are built, the number of layers in the CNN is further increased. The classification accuracy does not improve, and it will take more time [32]. The CNN model is shown in Figure 2, which consists of a convolution layer, a pooling layer, a full connection layer, and a Softmax layer.

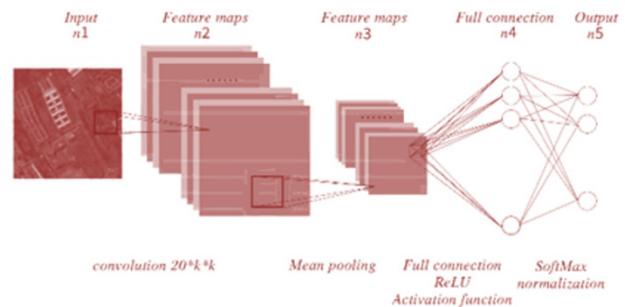


FIGURE 2. The structure CNN model.

The sparse vector of each pixel can be regarded as a two-dimensional image with one height, so the size of the input layer is $(n_1, 1)$ and n_1 is the total number of wavebands. The first convolution layer is obtained by filtering $20 k_1 \times 1$ convolution check (k_1 is the sparsity) for input images $(n_1, 1)$. It contains $20 \times n_2 \times 1$ nodes, where $n_2 = n_1 - k_1 + 1$. In order to reduce the amount of calculation and avoid the over-fitting issue, we pool the average value in the pool layer, which contains $20 \times n_3 \times 1$ nodes, where there are n and $n_3 = n_2/k_1$.

The first full connection layer has n_4 nodes, and the number of training samples from the pool layer to this layer is $(20 \times n_3 \times 1) \times n_4$. The neurons generated by the Relu activation function are used as the input of the second full connection layer. The second full connection layer has n_5 nodes, and the number of samples is $(n_4 + 1) \times n_5$ to be trained. The neurons generated by the Relu activation function are also used as the input of the Softmax layer. In the CNN model, if x_i is the input of the i -th layer, then there is $x_{i+1} = f_i(S_i^T x_i)$, where S_i^T is the sparse vector of the input pixel data of the layer, and $f_i(\cdot)$ is the Relu activation function. The second full connection layer will generate n_5 type labels, which will be put into the Softmax function for normalization operation, and then output all class probabilities with the dimension of $y = x_{L+1}$. The Softmax function is defined as

$$y = \frac{1}{\sum_{k=1}^{n_5} \exp(S_{L,k}^T x_L)} \begin{bmatrix} \exp(S_{L,1}^T x_L) \\ \exp(S_{L,2}^T x_L) \\ \vdots \\ \exp(S_{L,n_5}^T x_L) \end{bmatrix} \quad (10)$$

IV. EXPERIMENTAL VERIFICATION AND ANALYSIS

A. EXPERIMENTAL DATA AND ENVIRONMENT

In order to verify the effectiveness of the QGASR-CNN, hyperspectral remote sensing image datasets from Pavia University and Indian pines were selected in here [32]. The PCA, sparse representation (SR), and the sparse representation with QGA(QGASR) are used to extract the dimensionality-reduction features of the hyperspectral images. Then classification is trained by using the CNN model. The experimental results are compared with other algorithms in order to verify the effectiveness of the proposed QGASR-CNN.

TABLE 1. Basic data information of Pavia University.

Data	Pavia University
Collection location	Northern Italy
Acquisition equipment	ROSIS
Spectrum-coverage (um)	0.43-0.86
Data size (pixels)	610x340
Spatial resolution (m)	1.3
Band number	115
Band number after denoising	103
Sample size	42776
Number of categories	9

TABLE 2. Statistics of 9 types of features in Pavia University.

Type	Class name	Number of samples
1	Asphalt	6631
2	Meadows	18649
3	Gravel	2099
4	Trees	3064
5	Painted metal sheets	1345
6	Bare Soil	5029
7	Bitumen	1330
8	Self-Blocking Bricks	3682
9	Shadows	947

The experiment executed on a PC, and the basic configuration is Intel(R)Core(TM)i5-6300HQ CPU @ 2.30GHz.

Pavia University dataset is a hyperspectral remote sensing image dataset collected from Pavia University in the northern Italy by German airborne reflectance optical spectral imager [32]. The spectral imager continuously imaged 115 wavebands in the wavelength range of 0.43-0.86 μm , and the spatial resolution of the image was 1.3 m. Among them, 12 wavebands were eliminated due to the noise. The remaining 103 images are generally used. The size of the image is 610×340 , including 207400 pixels in total. A large number of background pixels are removed, and the remaining pixels are 42776 in total. The basic information of images is shown in Table 1, and the images are categorized into 9 types in Table 2.

Indian pines were imaged by AVIRIS, an airborne visible infrared imaging spectrometer, for an Indian pine tree in Indiana, USA. The size of 145×145 was intercepted and labeled as a hyperspectral image classification test dataset. The imaging wavelength range is 0.4-2.5 μm , and the ground objects are continuously imaged in 220 consecutive bands. However, since the 104-108, 150-163 and 220 bands cannot be reflected by water, we use the remaining 200 bands as the research object. The data contains 21025 pixels in total, a large number of background pixels are eliminated, and 10249 pixels including images are left. The basic information of the 16 types is shown in Table 3. The images are categorized into 16 types in Table 4.

B. DATA CONFIGURATION AND PARAMETER SETTING

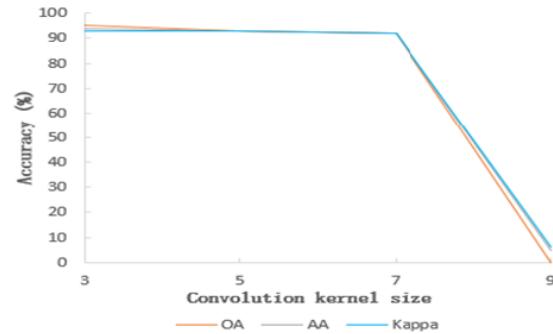
In terms of the selection of the size of convolution kernels, we used three-dimensional, five-dimensional, seven-dimensional, and nine-dimensional convolution kernels to conduct experiments on the datasets of Indian pines and Pavia

TABLE 3. Basic data information of Indian pines.

Data	Indian Pines
Collection location	Indiana USA
Acquisition equipment	AVIRIS
Spectrum-coverage (um)	0.4-2.5
Data size (pixels)	145×145
Spatial resolution (m)	20
Band number	220
Band number after denoising	200
Sample size	10249
Number of categories	16

TABLE 4. Statistics of 16 types of Indian pines.

Type	Class name	Number of samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-min	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-min	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Bldg-Grass-Tree-Drives	386
16	Stone-Steel-Towers	93

**FIGURE 3.** Influence of convolution check of different sizes on classification accuracy in Pavia University dataset.

University, respectively. By comparing the overall accuracy (OA), average accuracy(AA) and kappa coefficients(kappa), we can evaluate the influence of the size of convolution kernels on the classification performance of the model. The convolution influences with different sizes on classification accuracy are shown in Figure 3 and Figure 4.

From Figure 3 and Figure 4, it can be seen that two datasets of 5×5 convolution check obtain better classification performance. Therefore, the size of convolution core is set as 5×5 in this experiment. At the same time, we have carried out experiments on the influence of the proportion of training samples and test samples in the Indian pines dataset on the classification accuracy. The number of selected training samples are 10, 50, 100, 150, and 200, respectively.

As can be seen from Figure 5, the performance of the classification with growing training samples. As can be seen from the back part of the line graph, the addition of datasets

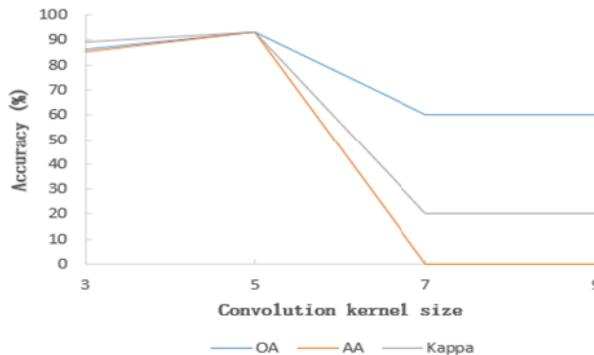


FIGURE 4. Influence of convolution check of different sizes on classification accuracy in Indian pines dataset.

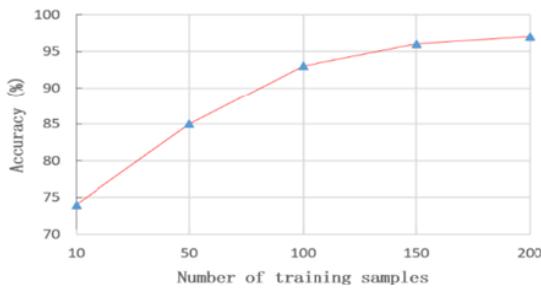


FIGURE 5. The influence of training sample number on classification accuracy.

TABLE 5. Pavia University data configuration.

No.	Category name	Training set / piece	Testing set / piece
1	Asphalt	200	6431
2	Meadows	200	18449
3	Gravel	200	1899
4	Trees	200	2864
5	Painted metal sheets	200	1145
6	Bare Soil	200	4829
7	Bitumen	200	1130
8	Self-Blocking Bricks	200	3482
9	Shadows	200	747
	Total	1800	40976

does not lead to over-fitting appearance. Therefore, 200 training samples from Pavia University and Indian pines datasets are selected in here. The remaining samples are used as test sets. The data configuration details of Pavia University dataset are shown Table 5.

Due to the less number of the samples for some surface feature types in the Indian pines dataset, we only reserve 8 types of surface feature with a large number of samples in this experiment. Similarly, 200 samples are selected for 8 types of surface feature as the training set and the rest samples regarded as the test set. The data configuration of Indian pines is shown in Table 6.

C. EXPERIMENTAL RESULTS AND ANALYSIS

In order to evaluate The classification effect of the proposed algorithm, The classification results are compared with the PCA, SR and QGASR. In the experiment, the commonly used classification and comparison evaluation indexes are overall

TABLE 6. Indian pines data configuration.

No.	Category name	Training set / piece	Testing set / piece
2	Corn-notill	200	1228
3	Corn-mintill	200	630
5	Grass-pasture	200	283
8	Hay-windrowed	200	278
10	Soybean-notill	200	772
11	Soybean-mintill	200	2255
12	Soybean-clean	200	393
14	Woods	200	1065
	Total	8180	2069

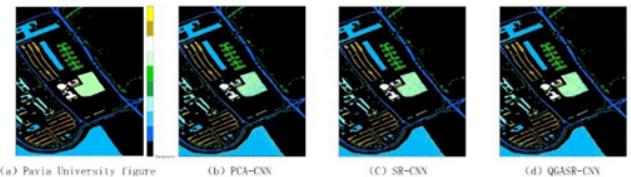


FIGURE 6. Classification diagram of three methods.

TABLE 7. Comparison of evaluation indexes of three algorithms.

Method	Training time(s)	Testing time(s)	Total accuracy(%)	Kappa coefficient
PCA-CNN	78.54	4.56	86.4	0.841
SR-CNN	68.34	4.78	88.7	0.863
QGASR-CNN	82.96	5.97	91.6	0.882

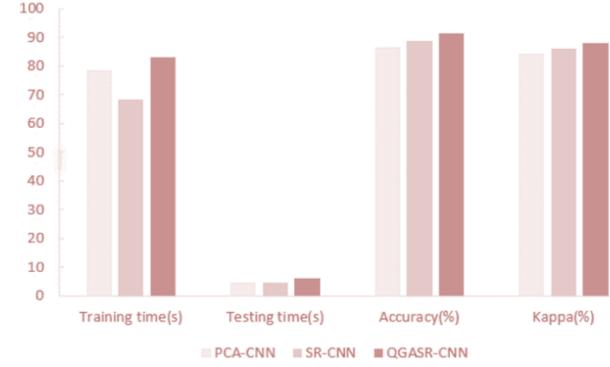


FIGURE 7. Comparison of evaluation indexes of three methods.

accuracy and kappa coefficient. All algorithms and experiments are repeated for 20 times, the maximum number of iterations is 80. The overall accuracy and kappa coefficients are the mean values for 20 tests.

Experiment 1: After the experiment is executed for hyperspectral remote sensing image dataset from Pavia University, the results comparison between the original hyperspectral remote sensing image and the classification results is shown in Figure 6.

The training time, testing time, total accuracy and kappa coefficient among three methods are obtained in Table 7 and Figure 7.

As can be seen from the Table 7 and Figure 7, the total accuracies of the PCA-CNN, SR-CNN and QGASR-CNN are 86.4%, 88.7% and 91.6%, respectively. The total accuracy

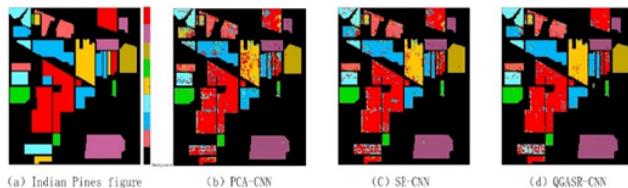


FIGURE 8. Classification chart of three methods of Indian pines data set.

TABLE 8. Comparison of evaluation indexes of three algorithms.

Method	Training time	Testing time	Total accuracy(%)	Kappa coefficient
PCA-CNN	23.52	1.572	88.2	0.861
SR-CNN	24.31	2.021	90.3	0.887
QGASR-CNN	25.09	2.219	94.1	0.921

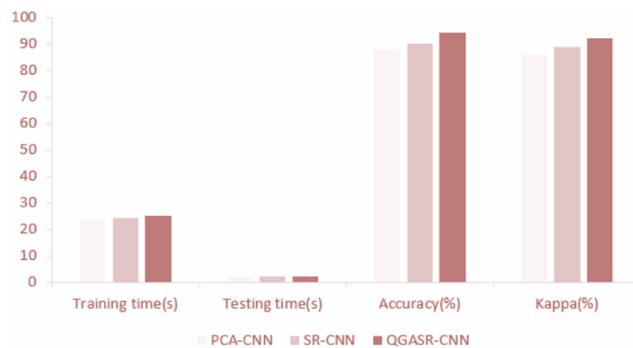


FIGURE 9. Comparison of evaluation indexes of three methods.

of the QGASR-CNN is best, which is 5.2% higher than that of the PCA-CNN. The kappa coefficients of PCA-CNN, SR-CNN and QGASR-CNN are 0.841, 0.863 and 0.882, respectively. The kappa coefficient of the QGASR-CNN is best among three methods. From the comparison results of the evaluation indexes among three methods for the hyperspectral remote sensing image, it can be seen that the QGASR-CNN is superior to the PCA-CNN and SR-CNN in classification accuracy and kappa coefficient. The proposed QGASR-CNN improves the classification accuracy, maintains the smoothness of classification results, and effectively reduces the classification error, but the operation efficiency of the QGASR-CNN is slightly lower than those of PCA-CNN and SR-CNN.

Experiment 2: After the experiment is executed for hyperspectral remote sensing image dataset from Indian pines, the results comparison between the original hyperspectral remote sensing image and the classification results is shown in Figure 8.

The operation time, total accuracy and kappa coefficient of the three methods are shown in Table 8 and Figure 9.

As can be seen from the Table 8 and Figure 9, the total accuracies of the PCA-CNN, SR-CNN and QGASR-CNN are 88.2%, 90.3% and 94.1%, respectively. The total accuracy of the QGASR-CNN is best, which is 5.9% higher than that of the PCA-CNN. The kappa coefficients of PCA-CNN, SR-CNN and QGASR-CNN are 0.861, 0.887 and 0.921, respectively. The kappa coefficient of the QGASR-CNN is

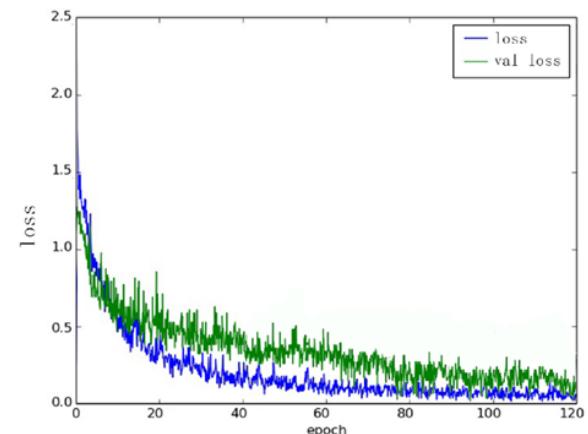


FIGURE 10. The loss curve for Pavia University dataset.

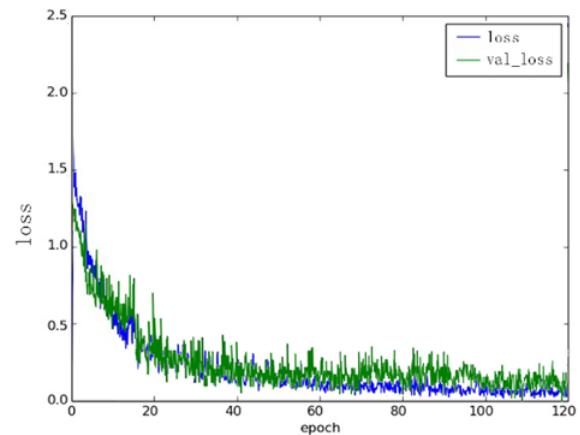


FIGURE 11. The loss curve for Indian pines dataset.

best among three methods. From the comparison results of the evaluation indexes among three methods for the hyperspectral remote sensing image, it can be seen that the QGASR-CNN is superior to the PCA-CNN and SR-CNN in classification accuracy and kappa coefficient. The proposed QGASR-CNN improves the classification accuracy, maintains the smoothness of classification results, and effectively reduces the classification error. The operation efficiency of the QGASR-CNN is close to those of PCA-CNN and SR-CNN.

Finally, in order to test whether the QGASR-CNN has over-fitting phenomenon, we have made a comparison between Pavia University dataset and Indian pines dataset. The maximum number of iterations is 120. The loss curve curves of the two datasets are shown in Figure 10 and Figure 11. In here, the loss indicates error value of training set, and the val_loss indicates error value of testing set.

As can be seen from the Figure 10 and Figure 11, that the convergence effect of the QGASR-CNN is more better. The error values of training set and testing set are minimum values, which show that the QGASR-CNN takes on better classification ability, generalization performance, and stability and robustness. When the QGASR-CNN runs

at the number of iterations (80), it has achieved a good classification accuracy. The experiment results show that the QGASR-CNN has faster convergence speed. At the same time, the QGASR-CNN does not exist any over-fitting phenomenon for classifying hyperspectral remote sensing image s from Pavia University dataset and Indian pines dataset.

As can be seen from the Figure 10 and Figure 11, that the convergence effect of the QGASR-CNN is more better. The error values of training set and testing set are minimum values, which show that the QGASR-CNN takes on better classification ability, generalization performance, and stability and robustness. When the QGASR-CNN runs at the number of iterations (80), it has achieved a good classification accuracy. The experiment results show that the QGASR-CNN has faster convergence speed. At the same time, the QGASR-CNN does not exist any over-fitting phenomenon for classifying hyperspectral remote sensing image s from Pavia University dataset and Indian pines dataset.

V. CONCLUSION

The traditional image processing and classification methods are difficult to achieve better classification results for hyperspectral remote sensing image with the same spectrum and different spectrum, the serious phenomenon of the same spectrum of foreign objects, the complex distribution of ground objects, the large difference of spatial scale, the small number of labeled samples, the complex and diverse noise types. A new hyperspectral remote sensing image classification method based on sparse representation with quantum genetic algorithm and convolutional neural network, namely QGASR-CNN is proposed in this paper. The sparse representation used to extract the characteristics of image and represent the image as a linear combination of base atoms in the dictionary. The quantum genetic algorithm is used to sparsely decompose the image to generate the sparse representation of the image. The sparse feature matrix of pixel group is constructed, and the matrix is used as convolution kernel to obtain the feature map relation of pixel in the CNN model, so as to achieve better classification effect. The hyperspectral remote sensing images from Pavia University and Indian Pines are used to prove the effectiveness of the proposed method. The classification accuracy can reach 94.1%. The experiment results show that the QGASR-CNN has a certain improvement in classification accuracy compared with the traditional method, which effectively solves the problem of “salt and pepper misclassification” in hyperspectral remote sensing image classification, and there is no any over-fitting phenomenon.

In the future work, we will further study the relationship between parameter selection and feature mapping in quantum optimization sparse decomposition, and further improve classification accuracy and operation efficiency.

REFERENCES

- [1] D. Landgrebe, “Hyperspectral image data analysis,” *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 17–28, Jan. 2002.
- [2] L. Zhang, L. Zhang, and B. Du, “Deep learning for remote sensing data: A technical tutorial on the state of the art,” *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.
- [3] H. Zhao, J. Zheng, W. Deng, and Y. Song, “Semi-supervised broad learning system based on manifold regularization and broad network,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 3, pp. 983–994, Mar. 2020.
- [4] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, “Deep learning in remote sensing: A comprehensive review and list of resources,” *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.
- [5] W. Deng, J. Xu, Y. Song, and H. Zhao, “An effective improved co-evolution ant colony optimization algorithm with multi-strategies and its application,” *Int. J. Bio-Inspired Comput.*, pp. 1–10, Mar. 2020.
- [6] L. Fang, S. Li, X. Kang, and J. A. Benediktsson, “Spectral-spatial hyperspectral image classification via multiscale adaptive sparse representation,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 12, pp. 7738–7749, Dec. 2014.
- [7] J. Peng, L. Li, and Y. Y. Tang, “Maximum likelihood estimation-based joint sparse representation for the classification of hyperspectral remote sensing images,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 6, pp. 1790–1802, Jun. 2019.
- [8] Z. Khan, F. Shafait, and A. Mian, “Joint group sparse PCA for compressed hyperspectral imaging,” *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4934–4942, Dec. 2015.
- [9] H. Yuan, Y. Yan Tang, Y. Lu, L. Yang, and H. Luo, “Hyperspectral image classification based on regularized sparse representation,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2174–2182, Jun. 2014.
- [10] W. Deng, H. Liu, J. Xu, H. Zhao, and Y. Song, “An improved quantum-inspired differential evolution algorithm for deep belief network,” *IEEE Trans. Instrum. Meas.*, early access, Mar. 25, 2020, doi: 10.1109/TIM.2020.2983233.
- [11] M. Zhang, W. Li, and Q. Du, “Diverse region-based CNN for hyperspectral image classification,” *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2623–2634, Jun. 2018.
- [12] H. Liang and Q. Li, “Hyperspectral imagery classification using sparse representations of convolutional neural network features,” *Remote Sens.*, vol. 8, no. 2, p. 99, Jan. 2016.
- [13] W. Li, G. Wu, F. Zhang, and Q. Du, “Hyperspectral image classification using deep pixel-pair features,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 844–853, Feb. 2017.
- [14] X. Huang and L. Zhang, “An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 257–272, Jan. 2013.
- [15] W. Li and Q. Du, “Gabor-filtering-based nearest regularized subspace for hyperspectral image classification,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1012–1022, Apr. 2014.
- [16] N. He, M. E. Paoletti, J. M. Haut, L. Fang, S. Li, A. Plaza, and J. Plaza, “Feature extraction with multiscale covariance maps for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 755–769, Feb. 2019.
- [17] G. Cheng, Z. Li, J. Han, X. Yao, and L. Guo, “Exploring hierarchical convolutional features for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6712–6722, Nov. 2018.
- [18] Y. Li, W. Xie, and H. Li, “Hyperspectral image reconstruction by deep convolutional neural network for classification,” *Pattern Recognit.*, vol. 63, pp. 371–383, Mar. 2017.
- [19] W. Song, S. Li, L. Fang, and T. Lu, “Hyperspectral image classification with deep feature fusion network,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3173–3184, Jun. 2018.
- [20] W. Deng, J. Xu, and H. Zhao, “An improved ant colony optimization algorithm based on hybrid strategies for scheduling problem,” *IEEE Access*, vol. 7, pp. 20281–20292, Feb. 2019.
- [21] Y. Zhong, Q. Cao, J. Zhao, A. Ma, B. Zhao, and L. Zhang, “Optimal decision fusion for urban land-use/land-cover classification based on adaptive differential evolution using hyperspectral and LiDAR data,” *Remote Sens.*, vol. 9, no. 8, p. 868, Aug. 2017.

- [22] R. Liu, L. Zhang, and B. Du, "A novel endmember extraction method for hyperspectral imagery based on quantum-behaved particle swarm optimization," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 4, pp. 1610–1631, Apr. 2017.
- [23] J. Li, H. Zhang, and L. Zhang, "Efficient superpixel-level multi-task joint sparse representation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5338–5351, Oct. 2015.
- [24] L. Gan, J. Xia, P. Du, and J. Chanussot, "Multiple feature kernel sparse representation classifier for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5343–5356, Sep. 2018.
- [25] C. Zhao and W. Li, "Collaborative representation-based binary hypothesis model with multi-features learning for target detection in hyperspectral imagery," *J. Indian Soc. Remote Sens.*, vol. 46, no. 5, pp. 847–862, May 2018.
- [26] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.
- [27] H. Lee and H. Kwon, "Going deeper with contextual CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4843–4855, Oct. 2017.
- [28] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "A new deep convolutional neural network for fast hyperspectral image classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 145, pp. 120–147, Nov. 2018.
- [29] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [30] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
- [31] S. Yu, S. Jia, and C. Xu, "Convolutional neural networks for hyperspectral image classification," *Neurocomputing*, vol. 219, pp. 88–98, Jan. 2017.
- [32] M. Ahmad, S. Protasov, A. M. Khan, R. Hussain, A. M. Khatta, and W. A. Khan, "Fuzziness-based active learning framework to enhance hyperspectral image classification performance for discriminative and generative classifiers," *PLoS ONE*, vol. 13, no. 1, Art. no. e0188996, doi: [10.1371/journal.pone.0188996](https://doi.org/10.1371/journal.pone.0188996).



HUAYUE CHEN received the B.S. degree in computer science and technology from the Sichuan Normal College, Nanchong, China, in 2002, and the M.S. degree in computer software and theory from Chongqing University, Chongqing, China, in 2005. She is currently pursuing the Ph.D. degree in geodetection and information technology with the Chengdu University of Technology, Chengdu, China. Since 2012, she has been an Associate Professor with China West Normal University, Nanchong. Her research interests include artificial intelligence, optimization method, and hyperspectral remote sensing image processing.



FANG MIAO (Member, IEEE) received the B.S. degree in radioactive geophysical prospecting, the M.S. degree in radioactive geology and exploration, and the Ph.D. degree in radioactive geology and exploration from the Chengdu Institute of Geology, Chengdu, China, in 1982, 1985, and 1990, respectively. Since 1995, he has been a Professor with the Chengdu University of Technology, Chengdu. His research interests include artificial intelligence, optimization method, and hyperspectral remote sensing image processing.



XU SHEN (Member, IEEE) received the B.S. degree in software engineering from China West Normal University, Nanchong, China, in 2020. His research interests include computer vision and natural language processing.

• • •