The Journal of Supercomputing (2021) 77:2829–2843
https://doi.org/10.1007/s11227-020-03377-w

Multi-spectral remote sensing land-cover classification based on deep learning methods

Tongdi He[1] **· Shengxin Wang**[2]

Published online: 6 July 2020

Abstract
It is of great significance and practical application value to extract land-cover type accurately. However, the input data usually used in classification such as reflectance data or vegetation index are very simple and quantitative remote sensing products are rarely used. In this paper, a multi-spectral land-cover classification method based on deep learning is proposed. Using the excellent detail capture ability of contourlet transform to obtain the potential information to supplement the spectral feature space, combined with deep learning for feature selection and feature extraction, a spectral–texture classification model is constructed. The multi-spectral sensing remote data and field measurement data in Dadukou District of Chongqing, northern Negev, and Changping region of Beijing were used for evaluation. Experiment results show the proposed method can outperform principal component analysis, linear discriminant analysis and neural network, and effectively improve the classification accuracy of multi-spectral images; this method provides a new perspective for land-use classification.

Keywords Multi-spectral images · Land-cover · Classification · Deep learning · Contourlet transform · Feature extraction

1 Introduction

Classification of different land-cover regions of remote sensing images is essential for efficient interpretation of them. As a mirror of nature, vegetation directly reflects the external characteristics of the complex [1]. It is recognized as an important and

✉ Tongdi He
hetongdi@hxu.edu.cn

Shengxin Wang
wsx870101364@email.swu.edu.cn

[1] College of Physical and Electrical Engineering, Hexi University, Zhangye 734000, Gansu, China

[2] College of Engineering and Technology, Southwest University, Chongqing 400715, China

sensitive indicator of changes in the ecological environment. Basically, the regions like vegetation, soil, water bodies, concrete structure, etc., of a natural scene are often not well separated leading to overlapping regions. Vegetation remote sensing classification is limited by the difficulty of remote sensing image acquisition and the immature automatic classification algorithm [2]. It is often interpreted by manual visual interpretation. With the development of remote sensing technology and the improvement of computer automatic classification methods, the accuracy and efficiency of vegetation remote sensing classification have been greatly improved [3]. However, traditional multi-spectral remote sensing is affected by a small number of sensor bands, low spectral resolution, complex vegetation types and spectral similarity, which makes a large number of "homologous" or "homologous foreign matter" phenomena widely exist, making it impossible to obtain higher vegetation [4]. Classification accuracy cannot meet the needs of precision agriculture, forest tree species classification and grass species identification.

Multi-spectral remotely sensed images comprise information over a large range of variation on frequencies (information), and these frequencies change over different regions, which needs to be estimated properly for improved classification [5, 6]. The multi-spectral remote sensing image data have both spectral features with correlated bands and spatial features correlated in the same band (also known as spatial correlation). An efficient utilization of the spectral and spatial (contextual) information can improve the classification performance significantly compared to the conventional non-contextual information-based classification methods (*k*-nearest neighbors, maximum likelihood, neural networks, etc.) [7]. Such conventional (non-contextual) approaches may be reasonable if spatial resolution is high or when the spectral intensities are well separated for different classes, which is rarely found in any real-life data [8]. For example, in the urban areas, the densities of the spectral intensities are seldom well separated. Thus, it is important to decide whether the arrangements of spatial data or a transformation of the pixels to a different space that uncorrelated the information to separate the data into desired classes can be used as features [9].

In recent years, with the deep learning method gradually being explored by people, its motivation is based on simulating the human brain for analytical learning, imitating the mechanism of the human brain to interpret the data, cognizing the data layer by layer and abstracting the shallow classification [10]. The data description and abstraction have limitation. Although the deep learning model can extract the features of the image well, it cannot directly extract the features of the high-dimensional data image [11]. Contourlet transform is a new extension of wavelet transform. It is a multi-resolution, localized, multi-directional image representation method. Its image base is distributed in multi-scale and multi-direction, and it also has full translation invariance [12]. With excellent detail capture, it can effectively reflect the texture and detail changes of the image.

In this paper, a multi-spectral land-cover classification method combining air-spectrum information and deep learning is proposed. Using excellent detail capture ability of the second-generation wavelet transform to obtain the potential information to supplement the spectral feature space, combined with deep learning for feature selection and feature extraction, a spectral–texture classification method for expressing

multi-information of features is constructed. The experiment was verified by multi-spectral image in Dadukou area of Chongqing, northern Negev, and Changping region of Beijing. The experimental results show the advantage of proposed methods.

2 Multi-spectral remote sensing land-cover classification based on deep learning

2.1 Correlation analysis

The spectral response characteristic is the most direct and important interpretive element of the ground object recognition of multi-spectral remote sensing images. Due to different material compositions and structures, various ground objects on the earth surface have their own unique spectral reflection and radiation characteristics, which are expressed in gray value differences at each wave band in the image [13]. Multi-spectral remote sensing data have provided a large number of ground object spectral characteristics, but during the process of data analysis and thematic information extraction, the visual interpretation of images is still needed. In addition, as ground objects have correlation among the radiation characteristics at various wave bands, a lot of redundancies may be present in the information contained in wave bands with large correlation [14]. In view of this, wave band portfolio selection needs to consider that the variance of the wave band radiant quantity should be large as much as possible, the wave band correlation should be small, and the spectral difference of ground object types should be large (which is related to the selection of the optimum temporal). The correlation coefficient method is adopted for the selection of the best wave band, so as to achieve the goals of keeping the main information, compressing the data size, enhancing and extracting the wave band data with better visual interpretation effects [15].

Presently, there are many available methods for correlation analysis, such as product moment correlation and rank correlation. In this paper, the method adopted is product moment correlation, a method to calculate the linear correlation between two variables, proposed in the twentieth century by British statistician Pearson. It is also called Pearson's correlation, a frequently used correlation analysis method at present. The correlation coefficients matrices for seven wave bands extracted from multi-spectral remote sensing images of Dadukou District, Chongqing City, and the correlation coefficient between the bands of remote sensing data is shown in Table 1.

By analyzing the results in Table 1, it can be seen that the correlation among Band 3, Band 4 and Band 7 of visible light bands is rather good, so these bands form the best wave band portfolio. Remote sensing image fusion was carried out on these three bands, and the ground objects were classified by the obtained fused images.

2.2 Spectral and texture feature extraction

Because the classification results only using the spectral features of images are often not ideal, it is necessary to use its diversified features to obtain potential information through image transformation to supplement the spectral feature space to enhance the

Table 1 The correlation coefficient between the bands of remote sensing data

Band	B1	B2	B3	B4	B5	B6	B7
Band 1	1						
Band 2	0.157	1					
Band 3	0.235	0.284	1				
Band 4	0.126	0.235	0.374	1			
Band 5	0.358	0.490	0.472	0.534	1		
Band 6	0.3422	0.374	0.368	0.472	0.538	1	
Band 7	0.752	0.783	0.795	0.034	0.639	0.723	1

image, combining spectral and texture information making classification results more effective. For each pixel to be classified, the spectral features are described by the rectangular area composed of the pixel itself and its surrounding neighborhood pixels. If the size of the neighborhood window is L, then for each pixel, the $L * L$ spectral feature vector can be constructed. The use of a vector composed of regions to represent a single-pixel spectral feature can weaken the spectral difference of similar features to a certain extent, while eliminating the influence of image speckle noise.

The contourlet transform, also known as the pyramidal direction filter bank (PDFB), is composed of Laplacian pyramid (LP) and direction filter bank (DFB), image base of the PDFB is distributed in multi-scale and multi-direction, and it also has full translation invariance. With excellent detail capture, it can effectively reflect the texture and detail changes of the image [22]. Laplacian pyramid (LP) decomposition is an effective way to achieve multi-resolution analysis of images. Each level of Laplacian pyramid decomposition will produce an approximate part and the remaining part, in which the approximate part is extracted to become the original 1/2, the decomposition of the next layer is only performed on the approximate part, and finally, the nth layer is formed. A pyramidal image decomposition consists of an approximate part and N detail parts. The specific implementation is shown in Fig. 1.

The directional filter bank (DFB) is a good way to decompose images in the direction and has good reconfiguration [22]. The implementation of DFB is through the decomposition of the l-tree of the l-layer Fig. (2). By dividing the frequency domain, 2^i sub-bands as shown in Fig. 3 are generated, which can be represented by 2^i channel structures.

The contourlet transform is implemented by combining the pyramid decomposition and the directional filter, and the image is not well described by any part alone [23]. The pyramid decomposition does not have directionality, and the directional filter can decompose well for the high-frequency part, but not fit for low-frequency part. The

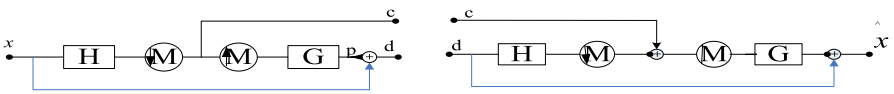

Fig. 1 Pyramid decomposition and reconstruction structure

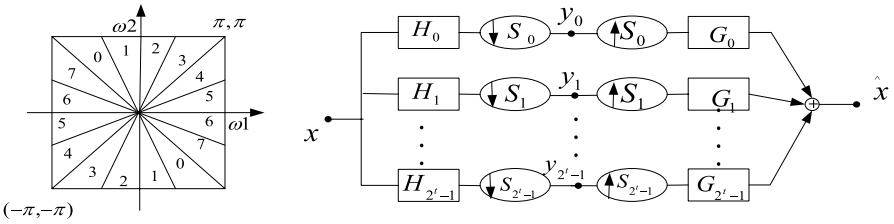

Fig. 2 The structure of DFB

combination of the two can just make up for the shortcomings of the other party and thus get a good image description as shown in Fig. 3.

The calculation method is as follows: A k-level NSCT transform is performed on the source image I, the generated band-pass sub-images are, respectively, $I_1, I_2, \ldots, I_K$, and the number of directional decompositions of a sub-image $(1 \le m \le k)$ is d_m. Then, the directional sub-images generated by $I_m (1 < m < k)$ are expressed as $I_m^1, I_m^2, \ldots, I_m^{2^{d_m}} (1 \le m \le k)$. For source image I, the texture feature vector is defined as:

$$F_{\mathrm{ij}} = \left[I_1^1(i,j) I_1^2(i,j), \ldots, I_1^{2^{d_1}}(i,j), I_2^1(i,j), I_2^2(i,j), \ldots, I_2^{2^{d_2}}(i,j), \ldots, I_k^1(i,j), I_k^2(i,j), \ldots, I_k^{2^{d_k}}(i,j) \right] \tag{1}$$

where $I_m^n(i,j)$ is the coefficients at which the directional sub-images are located and are represented, and $\sum_{m=1}^{k} 2^{d_m}$ is the dimensions of vector F_{ij}. The spectral feature vector and the texture feature vector of each pixel to be separated are connected end to end, the resulting eigenvector with dimension is $L * L + \sum_{m=1}^{k} 2^{d_m}$, and it is the spectral–texture feature vector of the pixel.

2.3 Deep belief network (DBN)

Typical deep learning models have three types: DBN, SAE and CNN [16]. The DBN was proposed by Hinton et al. in 2006 [17]. It has received extensive attention as a deep learning model and has been successfully applied in the fields of speech and object

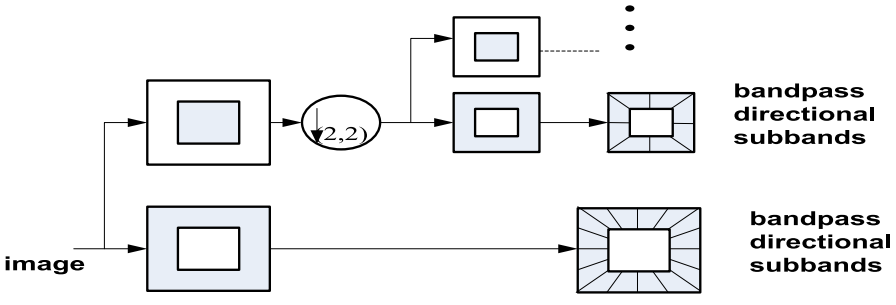

Fig. 3 The structure of PDFB

recognition, etc. The DBN is composed of Boltzmann machine (RBM) network and supervised back-propagation (BP) network. The DBN is a probability generation model that is composed of multiple RBM. A typical DBN model with three hidden layers is shown in Fig. 1. The underlying data training set is DBN, the first RBM model is Gauss–Bernoulli RBM, and the remaining RBM models are Bernoulli–Bernoulli RBM. The learning process of DBN mainly includes pre-training and fine-tuning, pre-training uses layer-by-layer training to train RBM in each layer separately, and weights and deviations between layers are fixed and saved for further analysis. In the fine-tuning phase, the weights and deviations of the DBN model will be updated by the back-propagation algorithm using the labeled input data. For details, refer to [18].

In Fig. 4, W is the weight between the two layers, h is the hidden layer,v is the visible layer, I is the number of neurons in the visible layer, J is the number of neurons in the hidden layer, h_j is the jth neuron in the hidden layer, v_i is the ith neuron in the visible layer [19]; the energy of the RBM is defined as:

$$E(v, h|\theta) = -\sum_{i=1}^{I} a_i v_i - \sum_{j=1}^{J} b_j h_j - \sum_{i=1}^{I}\sum_{j=1}^{J} w_{ij} v_i h_j \tag{2}$$

where a_i and b_j are the offset values of v_i and h_j, $\theta = (W_{ij}, a_i, b_j)$ is the parameter of the RBM, and RBM is an energy-based model that yields a joint overview of (v, h):

$$P(v, h|\theta) = \frac{1}{z_\theta} \exp[-E(v, h|\theta] \tag{3}$$

where $Z_\theta = \sum_v \sum_h \exp[-E(v, h|\theta]$ is the normalized term (also known as the fit distribution function). In the practical problem, the distribution $P(v|\theta)$ of the observed

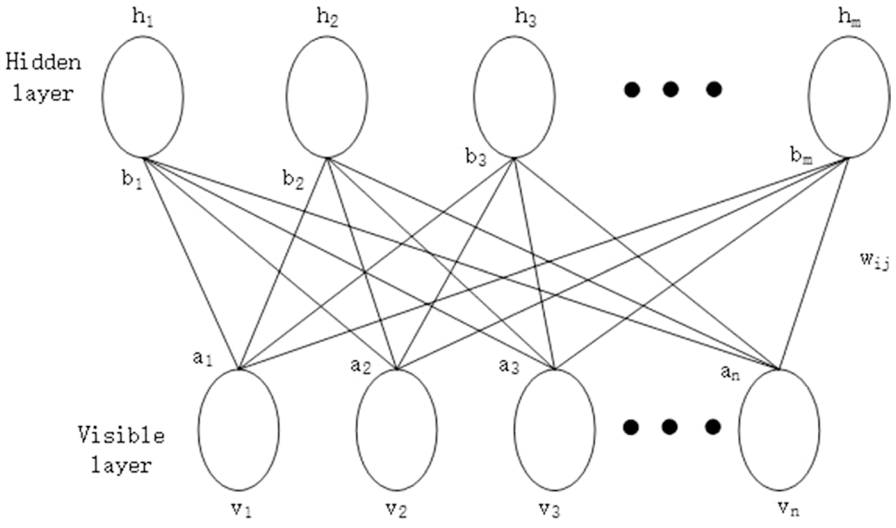

Fig. 4 The structure of RBM

variable is defined by RBM, which is the edge distribution of the joint probability $P(v, h|\theta)$ [20, 21], and the probability of the jth hidden layer node is:

$$P(h_k = 1|v, \theta) = \sigma\left(b_j + \sum_{i=1}^{I} v_i w_{ji}\right) \tag{4}$$

Among them, $\sigma(x) = 1/(1 + e^{-x})$ is a sigmoid function. It is given the state of the hidden layer node, and the ith visible layer node is:

$$P(v_k = 1|h, \theta) = \sigma\left(a_i + \sum_{i=1}^{J} h_j w_{ji}\right) \tag{5}$$

The RBM is trained in an iterative manner, and the goal of the training is to learn the value of the $\theta = (W_{ij}, a_i, b_j)$; parameter θ is obtained by finding maximum log likelihood function on training set (assuming T samples are included):

$$\theta^* = f_{\arg\max_\theta} \sum_{i=1}^{T} \ln P\left(v^{(t)}|\theta\right) \tag{6}$$

where the maximum value of $v^{(t)}$ is usually obtained by the gradient ascending method, and the key is to calculate the gradient $\frac{\partial \ln(\theta)}{\partial \theta}$, which is available through derivation:

$$\frac{\partial \ln(\theta)}{\partial \theta} = \sum_{i=1}^{T} \left[-\sum_{h} P(h|v^{(t)}, \theta) \frac{\partial E(v^{(t)}, h|\theta)}{\partial \theta} + \sum_{v.h} P(v|h, \theta) \frac{\partial E(v, h|\theta)}{\partial \theta} \right] \oplus \tag{7}$$

where $P(h|v^{(t)}, \theta)$ is the probability that visible layer unit is defined as hidden layer of known training sample $v^{(t)}$, and the update rules for each parameter are as follows:

$$\Delta W_{ij} = \varepsilon\left(\left\langle v_i h_j \right\rangle_{\text{data}} - \left\langle v_i h_j \right\rangle_{\text{recon}}\right), \tag{8}$$

$$\Delta a_i = \varepsilon\left(\langle v_i \rangle_{\text{data}} - \langle v_i \rangle_{\text{recon}}\right), \tag{9}$$

$$\Delta b_j = \varepsilon\left(\left\langle h_j \right\rangle_{\text{data}} - \left\langle h_j \right\rangle_{\text{recon}}\right), \tag{10}$$

where $\langle \cdot \rangle_{\text{data}}$ is the mathematical expectation of training dataset definition distribution, $\langle \cdot \rangle_{\text{recon}}$ is the distributional expectation defined by the reconstructed model, and ε is pre-training learning rate. The classification process of DBN consists of two parts:

1. Training process: the vector composed of training samples is sent to the visible layer of RBM in the first layer as input. Then, through the layered training, the correlation of data was mapped to the hidden layer h_1, h_2, … h_m successively. In order to speed up the training process, batch learning is generally adopted, a fixed

number of training samples are randomly sampled each time and input the neural network as a mini-batch, and each mini-batch carries out a weight update.
2. Prediction process: the data of test samples are obtained by similar data organization of training samples, so they have the same data dimension. Based on the trained network structure, classification results by forward propagation are obtained.

2.4 Land-cover classification

The steps to construct a spectral–texture classification model that expresses the multi-information of the ground objects are as follows:

1. Performing correlation analysis to analyze the relationship between features and spectra to ensure the feasibility of classification;
2. Using contourlet to calculate the texture features and construct the spectral–texture features of the classified images;
3. Selecting training samples and setting of DBN network parameters; using the set DBN network combined with texture feature extraction to classify features of multi-spectral remote sensing images;
4. Evaluating classification accuracy for various algorithms.

The procedure of remote sensing land-cover classification uses deep learning method as shown in Fig. 5.

3 Experimental results and analysis

3.1 Parameter settings

In experiment, the window day for spectral feature calculation is set to $L = 3$, the scale of the contourlet transform is 3, and the number of each level is 4. The obtained texture feature vector dimension is $16 + 16 + 8 = 40$. The experiment calculates the spectral–texture eigenvectors of the three bands separately and combines the three eigenvectors into one eigenvector as the input of the classification. The input vector dimension is $3 * (3 * 3 + 40) = 1470$ (Figs. 6, 7, 8).

The initial value of the learning rate in the experiment was set to 0.01, and the training was dynamically adjusted by the reconstruction error. Divided by 20 elements from a random number of normal distribution $N(0, 0.01)$, the elements of the hidden layer offset h are initialized to 0. The number of nodes of the hidden layer is selected by calculating the reconstruction error of the RBM, the reconstruction error is based on the training data as the initial state, and the Gibbs, the sample obtained after sampling, and the original data are different according to the distribution of the RBM (generally evaluated by a norm or a two norm).

The III group experimental samples were taken from the remote sensing data, and 100 samples from each type of ground and 700 randomly selected points were

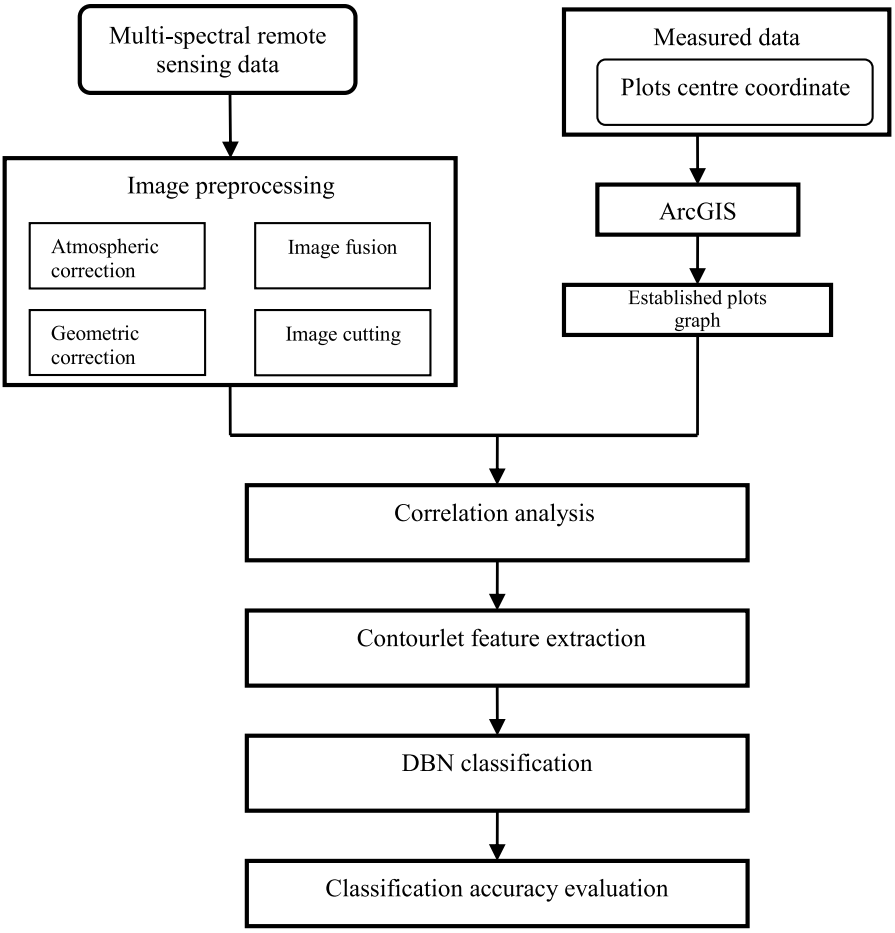

Fig. 5 The procedure of remote sensing land-cover classification uses deep learning method

randomly selected to form experimental sample data, and multiple small sample sets were trained in batches. For the RBM hidden layer node parameters, since the number of output layer nodes is 5 and input layer nodes is 147, respectively, the reconstruction error corresponding to the number of different nodes is calculated. When the number of nodes is 56, the reconstruction error is the smallest, thereby determining the first hidden layer node.

3.2 Classification results

The proposed method was compared with PCA, LDA and RBF neural network. Principal component analysis (PCA) is a typical unsupervised linear feature extraction method, in which the original m-dimensional data are embedded into the d-dimensional subspace, which is spanned by the leading eigenvectors in the data

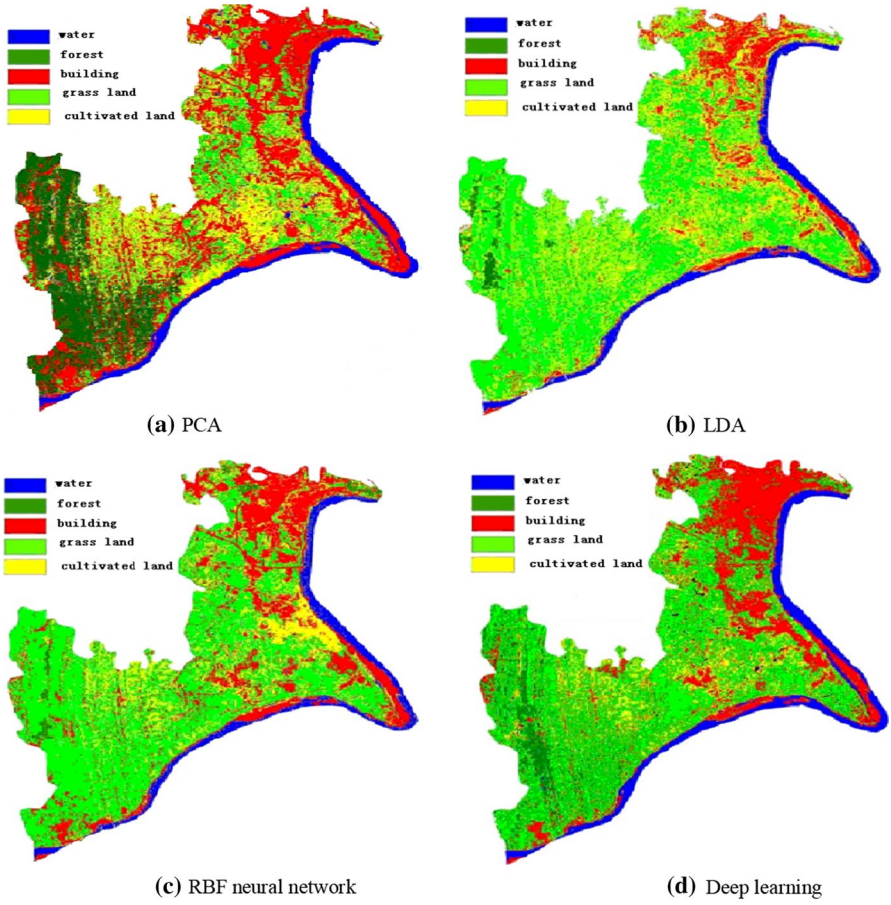

Fig. 6 The classification results of Dadukou area of Chongqing

covariance matrix. The goal of PCA is to find a set of orthogonal basis functions to capture the direction of the maximum variance in data, so as to maintain the Euclidean distance of the point to the data. The main steps are as follows: Firstly, the remote sensing dataset is represented by covariance matrix; then, the covariance matrix is decomposed by feature value. Finally, the order of feature value is adjusted so that the feature value is arranged according to the corresponding feature value after decomposition, and the corresponding vectors of the previous feature value are taken as the basis of the subspace, and the representation of the dataset is obtained under the basis mapping.

The basic idea of LDA is the supervised feature extraction algorithm first proposed by Fisher. Its purpose is to select the vector that makes Fisher's criterion function reach the extreme value as the best projection direction, so that after the data are projected to the low-dimensional space, all kinds of samples are as "compactly" as possible, and different kinds of samples are as "scattered" as possible. The main

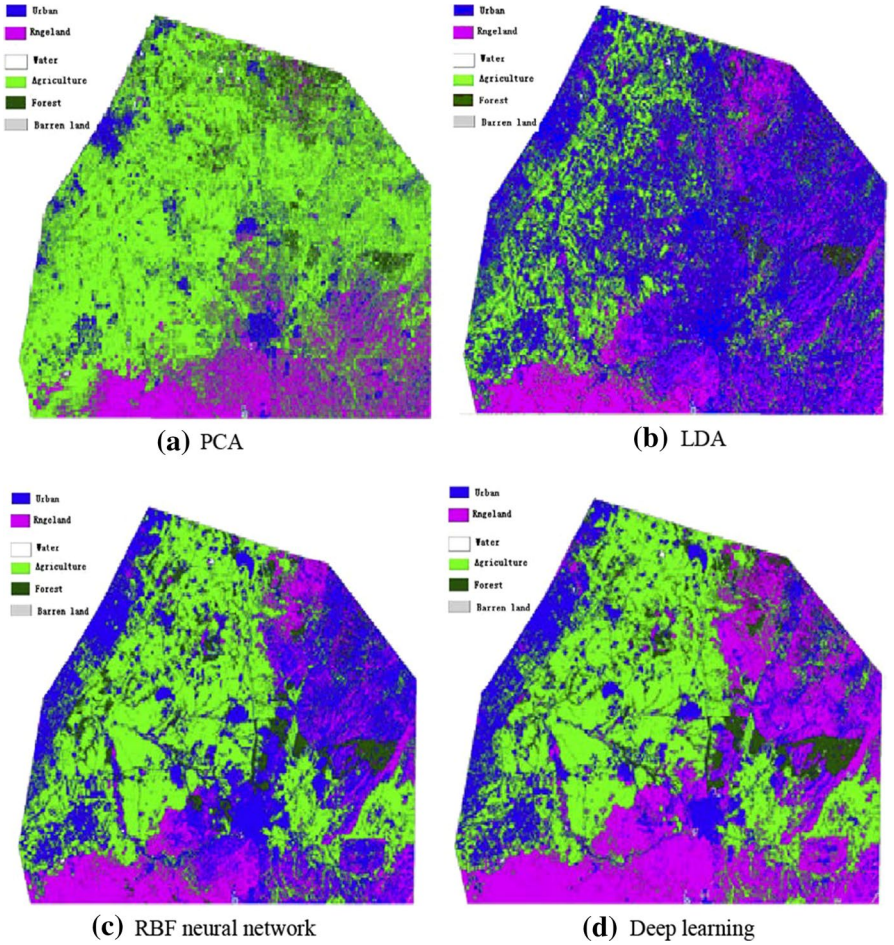

Fig. 7 The classification results of northern Negev

steps are as follows: Firstly, the overall dispersion matrix is determined by using remote sensing datasets; then, the objective function is determined and the feature value corresponding to the nonzero feature value is obtained. Finally, the regular subspace is obtained and the sample data are mapped from the high-dimensional space to the low-dimensional space.

The specific method of multi-spectral remote sensing image classification model use RBF neural network: Firstly, the self-organizing learning algorithm is used for unsupervised clustering to determine the RBF center; secondly, the RBF width is determined according to the average distance of the sample belonging to the center; finally, gradient descent method is used to train the connection weight between the RBF layer and the output layer.

The I group experimental data are the ETM+ imagery captured by Landsat-7 at Dadukou District of Chongqing in 2000. The remote sensing image is preprocessed

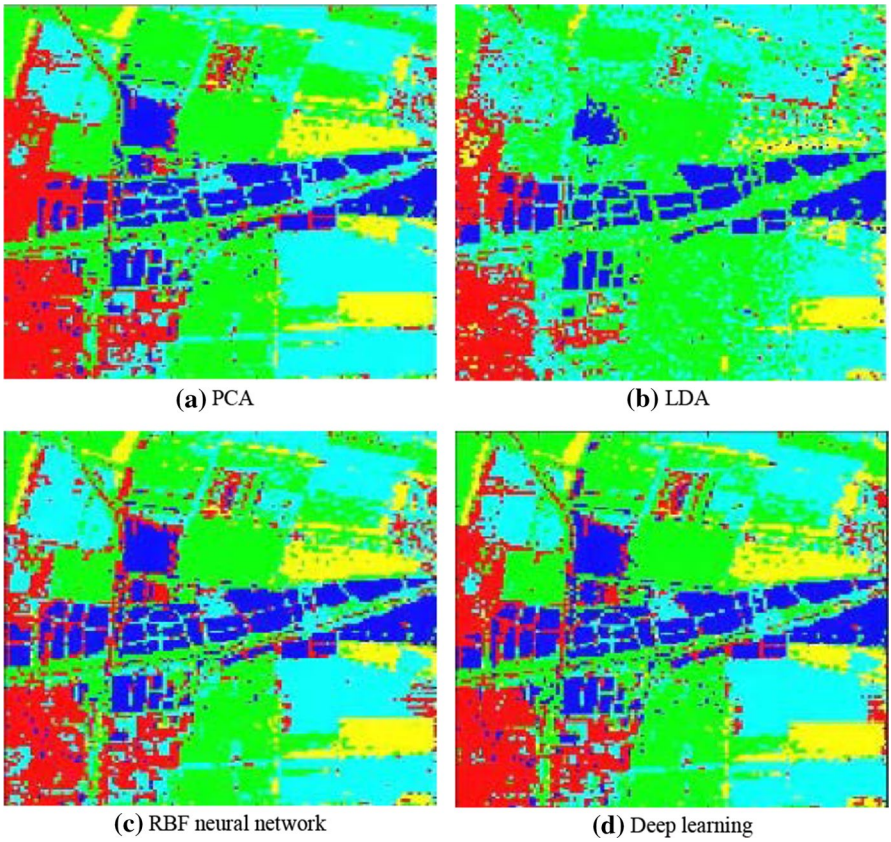

Fig. 8 The classification results of Changping region of Beijing city

by geometric correction, radiation correction and image enhancement according to the ground control points, and the research area is extracted from the remote sensing image through the combination of administrative division vector data and GIS technology.

The II group experimental data is a Landsat-5 TM image acquired on January 30, 2009. The research area over the Israeli Negev desert fringe is about 4000 km^2 in size.

The III group experimental data are the OMIS imagery captured at Changping region of Beijing city, and the experimental image consists of 512 rows and 512 pixels.

3.3 Evaluation of classification accuracy

In this paper, the overall classification accuracy is used as the basis for the classification result evaluation. The overall classification accuracy (OA) refers to the sum of pixels that are correctly classified divided by the total number of pixels. The real

image of the surface or the real region of interest of the surface limits the pixels true classification. Because the multi-spectral image is divided into five categories in this paper, the overall accuracy of each class is different. We take the highest classification accuracy as the overall classification accuracy of this algorithm to evaluate its classification ability. The classification accuracy comparison on the actual datasets using different algorithms is shown in Table 2.

It can be seen from Table 2 that with the increase in the number of training samples, the overall classification accuracy of all algorithms is increasing, because the more the training samples, the more effective the features of images obtained through learning will be, thus improving the classification accuracy. The recognition performance of RBF is significantly better than that of PCA and LDA, and it can capture the various structural characteristics of data effectively, and the classification effect is better than of the first four algorithms with a classification accuracy improved by 6%. Because the proposed method has excellent detail capture ability by contourlet transform to obtain the potential information to supplement the spectral feature space, and combined with deep learning for feature selection and feature extraction, the classification accuracy is greatly improved. Its classification accuracy is improved by 17% than PCA, which is the highest among the various classification methods.

4 Conclusions

Land-cover change is a frontier and hot spot of global change research at presently, and it will affect the function of the earth's biochemical sphere and the cycle of material energy. Land-cover change is closely related to global ecological evolution, biodiversity, climate change and survival and development of human. The commonly used approaches of land-cover classification mainly use spectral information and neglecting spatial information. A multi-spectral land-cover classification method combining air-spectrum information and deep learning is proposed. Using the excellent detail capture ability of the second-generation wavelet transform to obtain the potential information to supplement the spectral feature space, combined with deep learning for feature selection and feature extraction, a spectral–texture classification model for expressing multi-information of features is constructed. The results show deep learning combining spectral–texture features of images can effectively improve the classification accuracy, and more accurately mine the spatial distribution law of multi-spectral remote sensing images. Also, it is observed that different structures of the classified regions obtained with the proposed scheme are well shaped. But those

Table 2 Classification accuracy of six methods on the actual dataset (unit: %)

	PCA	LDA	RBF	Proposed method
Experiment **I**	73.0	78.3	86.2	90.1
Experiment **II**	68.7	72.8	76	81
Experiment **III**	71.2	75.6	80.1	87

methods cannot describe the relationship between land-cover change and ecosystem function improvement or decline well. At the same time, the evaluation index cannot describe the magnitude and speed of change well. In future research, we will further study a variety of deep learning models, in order to improve the classification accuracy and reliability of deep learning.

Acknowledgements This research work has been partially supported by National Science Foundation of China under Grant No. 41371338.

References

1. Alvarez-Mieles G, Lrvine K, Griensven A (2013) Relationships between aquatic biotic communities and water quality in a tropical river-wetland system. Environ Sci Policy 5(3):1–13
2. Gao XG, Wang L, Qi DH (2013) Eco-Security evaluation of wetland based on PSR model-taking Dashanbao wetland as an example. J Nat Sci Human Normal Univ 36(2):86–90
3. He TD, Zhao KL (2018) Multispectral remote sensing land use classification based on RBF neural network with parameters optimized by genetic algorithm. Int Conf Sensor Netw Signal Process (SNSP) 8:118–123
4. Kaya GT, Ersoy OK, Kamasak ME (2011) Support vector selection and adaptation for remote sensing classification. IEEE Trans Geosci Remote Sens 49(6):2071–2079
5. Diao W, Sun X, Zheng X et al (2016) Efficient saliency-based object detection in remote sensing images using deep belief networks. IEEE Geosci Remote Sens Lett 13(2):137–141
6. Lv Q, Dou Y, Niu X, et al (2014) Classification of land cover based on deep belief networks using polarimetric RADARSAT-2 data. In: Proceedings of Geoscience and Remote Sensing Symposium, pp 4679–4682
7. Do MN, Vetterli M (2002) Contourlets: a directional multi resolution image representation. In: IEEE ICIP, pp 357–360
8. Tian B, Xiong WZ (2018) A side information generation method using deep learning for distributed video coding. J Phys Conf Ser 6(5):210–223
9. Shankar B, Meher Saroj K, Ghosh Ashish (2011) Wavelet-fuzzy hybridization: feature extraction and land-cover classification of remote sensing images. Appl Soft Comput 11:2999–3011
10. Lecun Y, Bottou L, Bengio Y et al (2014) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
11. Qu F, Zhang JT, Shao ZT, Qi SZ (2017) An intrusion detection model based on deep belief network. In: International Conference on Network, Communication and Computing, vol 77, pp 15–27
12. Huang Hai B, Huang Xiao R, Li Ren X, Teik C (2016) Sound quality prediction of vehicle interior noise using deep belief networks. Appl Acoust 35(12):1–12
13. Bengio Y, Lamblin P, Dan P et al (2006) Greedy layer-wise training of deep networks. In: Proceedings of International Conference on Neural Information Processing Systems. MIT Press, Cambridge, pp 153–160
14. Liu L, Sun WW, Ding B (2016) Offline handwritten Chinese character recognition based on DBN fusion model. In: IEEE International Conference on Information and Automation, pp 1911–1918
15. Fan X, Peter WT (2018) Combined deep belief network in deep learning with affinity propagation clustering algorithm for roller bearings fault diagnosis without data label. J Vib Control 34(1):1–26
16. Niu X, Ban YF (2014) A novel contextual classification algorithm for multitemporal polarimetric SAR data. IEEE Geo-sci Remote Sens Lett 11(3):681–685
17. Burt PJ, Adelson EH (2013) The Laplacian pyramid as a compact image code. IEEE Trans Commun 31(4):532–540
18. Bamberger RH, Smith MT (2002) A filter bank for the directional decomposition of images: theory and design. IEEE Trans Signal Proc 40(4):882–893
19. Mahdianpari M, Salehi B, Rezaee M, Mohammadimanesh F, Zhang Y (2018) Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. Remote Sens 10:1119. https://doi.org/10.3390/rs10071119

20. Zhang P, Ke Y, Zhang Z, Wang M, Li P, Zhang S (2018) Urban land use and land cover classification using novel deep learning models based on high spatial resolution satellite imagery. Sensors 18:3717. https://doi.org/10.3390/s18113717
21. Li C, Wang Y, Zhang X, Gao H, Yang Y, Wang J (2019) Deep belief network for spectral-spatial classification of hyperspectral remote sensor data. Sensors 19:204. https://doi.org/10.3390/s19010204
22. Li J, Xi B, Li Y, Du Q, Wang K (2018) Hyperspectral classification based on texture feature enhancement and deep belief networks. Remote Sens 10:396. https://doi.org/10.3390/rs10030396
23. Gaetano R, Ienco D, Ose K, Cresson R (2018) A two-branch CNN architecture for land cover classification of PAN and MS imagery. Remote Sens 10:1746. https://doi.org/10.3390/rs10111746