



Research article

Multi-spectral remote sensing images feature coverage classification based on improved convolutional neural network

Yufeng Li*, Chengcheng Liu, Weiping Zhao and Yufeng Huang

Key Laboratory of Aerospace Information Perception and Intelligent Processing, College of Electronic and Information Engineering, Shenyang Aerospace University, Shenyang 110122, China

* **Correspondence:** Email: liyufeng@sau.edu.cn; Tel: +008618040038866.

Abstract: With the continuous development of the earth observation technology, the spatial resolution of remote sensing images is also continuously improved. As one of the key problems in remote sensing images interpretation, the classification of high-resolution remote sensing images has been widely concerned by scholars at home and abroad. With the improvement of science and technology, deep learning has provided new ideas for the development of image classification, but it has not been widely used in remote sensing images processing. In the background of remote sensing huge data, the remote sensing images classification based on deep learning proposed in the study has more research significance and application value. The study proposes a high-resolution remote sensing images classification method based on an improved convolutional neural network. The traditional convolutional neural network framework is optimized and the initial structure is added. The actual classification results of radial basis functions and support vector machine are compared horizontally. The classification results of hyperspectral images were presented that the improved method can perform better in overall accuracy and Kappa coefficient. The commission errors of support vector machine classification method are more than 6 times of that of the improved convolutional neural network classification method and the overall accuracy of the improved convolutional neural network classification method has reached 97% above.

Keywords: convolutional neural network; support vector machine; radial basis function; remote sensing images classification; inception

1. Introduction

The processing of satellite remote sensing images plays a vital role in the protection of natural resources, the technical support of civilian equipment and military reconnaissance [1]. Therefore, the image processing has always been the focus of research remote sensing images classification, which is to classify each pixel in the image into one of several categories [2]. The effective features are extracted by analyzing the spectral and spatial information of various types of objects [3]. Then the suitable feature parameters are analyzed and selected. The feature space is divided into several non-intersecting subspaces, and each pixel in the image is divided into these subspaces [4].

The traditional classification methods of satellite remote sensing images are generally supervised classification, unsupervised classification and ancient remote sensing visual images interpretation [5]. The common supervised classification methods are support vector machine, minimum distance, maximum likelihood and so on [6–7]. Unsupervised classification methods include iterative self-organizing data analysis techniques algorithm, K-means clustering algorithm and other unsupervised classification methods [8]. The results of the visual image interpretation method are more accurate, but it is subjectively influenced by the interpreters. The above classification algorithms are shallow learning algorithms. Although they utilize the spectral information of the pixels in the image, these shallow learning algorithms are limited due to the limited computational units of these algorithms and the large amount of sample size and the complex and diverse features of the satellite remote sensing images [9]. When comes to complex classification problems, its generalization ability is restricted, which makes it impossible to express complex features effectively [10]. Such shallow models will eventually be replaced by some emerging methods.

Since the concept of deep learning is put forward by Hinton and other professors at the University of Toronto in the top academic journal *Science* in 2006, it has attracted great attention all over the world. Hinton uses a multi-layer mechanism model similar to the human brain to reduce dimensionality and classify information [11]. This deep learning method has made great achievements in image, speech recognition and other fields [12]. The convolutional neural network proposed by Razavian is a multi-layer neural network structure with excellent training performance [13]. It has good application in remote sensing images processing [14]. Haobo Lyu proposed a new change detection algorithm named REFEREE in 2016. The REFEREE method can detect multi-class changes for multi-temporal images [15]. Nataliia Kussul proposed deep learning classification of land cover and crop types using remote sensing data in 2017. This is the first attempt to apply CNNs to multisource multitemporal satellite imagery for crop classification [16]. R. Marc, K. Marco proposed an automated end-to-end approach for multi-temporal classification in 2018, which achieved state-of-art accuracy in crop classification tasks with a large number of crop classes [17]. However, there are relatively few works on the application of deep learning to the interpretation of high-resolution satellite remote sensing images. Therefore, the study proposes a new method to optimize the structure of traditional convolutional neural network for high-resolution remote sensing images classification. This research based on convolutional neural network classifies high-resolution multi-spectral remote sensing images automatically, optimizes the traditional convolutional neural network framework and adds Inception structure, compares its classification effect with support vector machine algorithm and radial basis functions horizontally, and it has a better improvement.

2. Materials and methods

2.1. Traditional classification principles

2.1.1. Support vector machine classification method

Support vector machine was proposed by Cortes and Vapnik in 1995. Support vector machine is a general learning method developed based on statistical learning theory [18]. Its basic idea is to transform the original classification space into high-dimensional space by inner product kernel function. In the transformed high-dimensional space, the decision plane of maximum edge interval is also constructed, which is also called optimal decision hyper-plane. It achieves the goal of seeking the best compromise between learning accuracy and learning ability in a small sample space and achieving optimal promotion ability. The classification performance of support vector machine depends on the selection of support vector machine classification model. However, there is still no good general solution for the selection of parameter model.

Support vector machine can learn, classify and predict the sample data. The classification flow chart of support vector machine is shown in Figure 1.

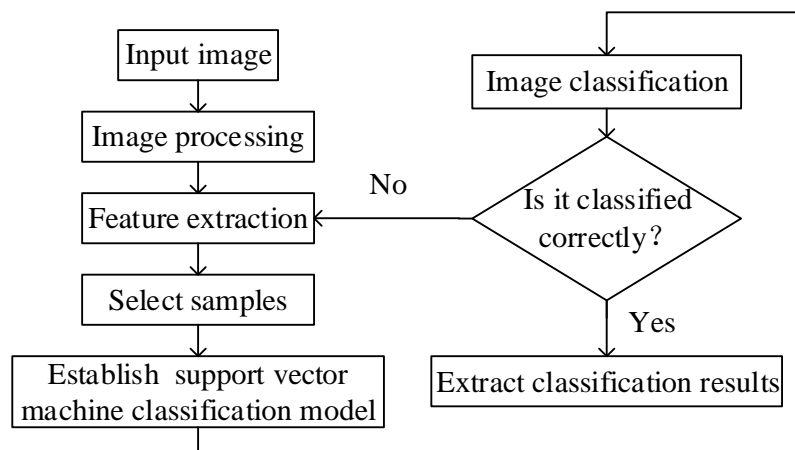


Figure 1. Support vector machine classification flow chart.

2.1.2. Radial basis function classification method

The radial basis function neural network is divided into three layers [19]. The first layer is the input layer, which completes the introduction of feature vectors into the network. The second layer is the hidden layer, which transforms the low-dimensional input mode into the high-dimensional space to facilitate the classification and recognition of the output layer. The number of nodes in the hidden layer depends on the need to solve the problem. The selection of hidden layer nodes generally uses Gaussian function as the transfer function:

$$\Phi_i(x) = \exp[-\|x - c_i\|^2/2\delta^2] \quad i = 1, 2, \dots, m \quad (1)$$

In the formula x is the n -dimensional input vector, c_i is the center of the i -th basis function,

which has the same dimension as x , and δ is the i -th perceived variable (also can be a freely selected parameter), δ determines the width of the basis function around the center point and the scope of the corresponding basis function of the center point, m is the number of perceptual units.

The third layer is the output layer. If the number of hidden layer nodes is m , the output is:

$$y_i = \sum_{i=1}^m W_i \Phi_i (\|x - c_i\|) \quad (2)$$

Where: W is the weights the center of the basis function and $\| \cdot \|$ is a 2-norm.

The structure of radial basis function neural network classifier is shown in Figure 2.

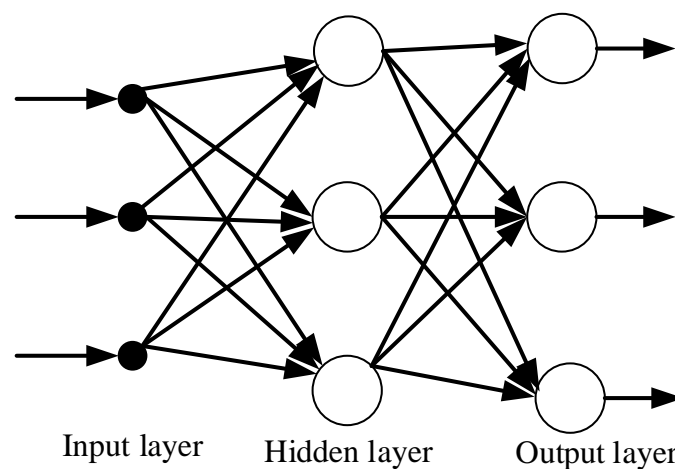


Figure 2. Radial basis function neural network classifier structure.

Radial basis function neural network has good generalization ability and fast learning convergence. It has been successfully applied to data classification, pattern recognition, information processing, image processing and so on. The network has faster computing speed and stronger nonlinear mapping capability [20].

2.1.3. Convolutional neural network classification method

Convolutional neural network is a typical model of deep learning [21]. The schematic diagram of traditional convolutional neural network is shown in Figure 3. Generally, the convolutional neural network model is represented by different structural layers, including convolutional layer, pooling layer (subsampling layer), one or more fully connected layers and output layer. The convolution layer is used to convolute the input image using the specified filter, and usually occurs alternately with the pooling layer [22].

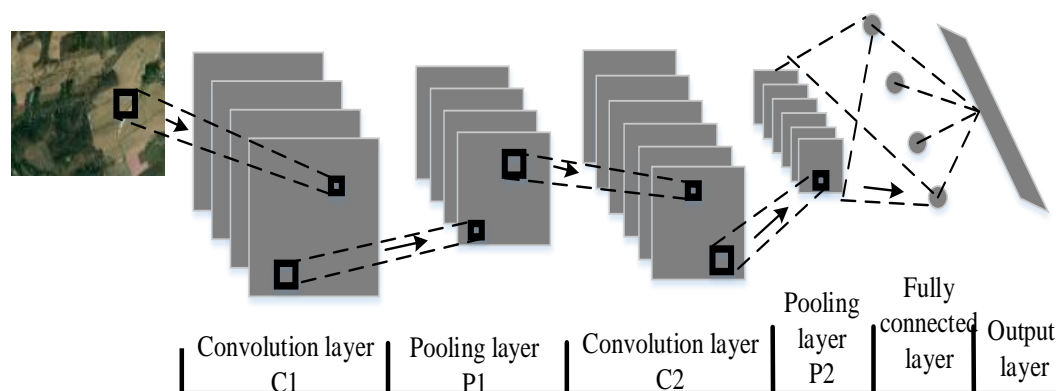


Figure 3. Schematic diagram of traditional convolutional neural network.

In ordinary neural networks, one neuron connects to all the neurons in the next layer. In convolutional neural network, neurons are sparsely connected, usually within the self-defined sensory range of each designated neuron. In addition, some interconnected neurons in a layer have the same weights and deviations. To a large extent, they can help to reduce parameters. The pooling layer is the feature extraction layer. The continuous range of the feature map obtained by the convolution of the previous layer is the action area, and only the features generated by repetitive hidden units are pooled. These pooling units have translation invariant, and the whole convolutional neural network has translation invariant. Even after a small translation, the input image still produces the same features.

2.2. Improved convolutional neural network classification method

The performance improvement of a convolutional neural network is usually to increase the depth or width of the network, which is to increase the number of layers or the number of neurons in each layer. However, this design method is not only easy to overfitting, but also increases the computational complexity. The solution to these two problems is to reduce the parameters while increasing the depth and width of the network. In order to reduce the parameters, the natural full connection needs to become a sparse connection. There will not be a qualitative improvement, because most of the hardware is optimized for dense matrix calculations. Although the sparse matrix has a small amount of data, the time consumed is difficult to reduce.

The method used in the study is to add the Inception structure to the traditional convolutional neural network classification model. The method used in the study is to add the Inception structure to the traditional convolutional neural network classification model. In deep learning, large-scale convolution kernels can bring about larger receptive fields, but they also generate more parameters. For example: 5×5 convolution kernels have 25 parameters, and 3×3 convolution kernels have 9 parameters. The former is 2.78 times of the latter. If a small size filter is used to replace a large size filter, a small network composed of two 3×3 convolutional layers connected in series is used to replace a single 5×5 convolutional layer, which reduces the number of parameters while maintaining the receptive field range. Optimize the traditional convolutional neural network classification effect [23]. The Inception structure is shown in Figure 4.

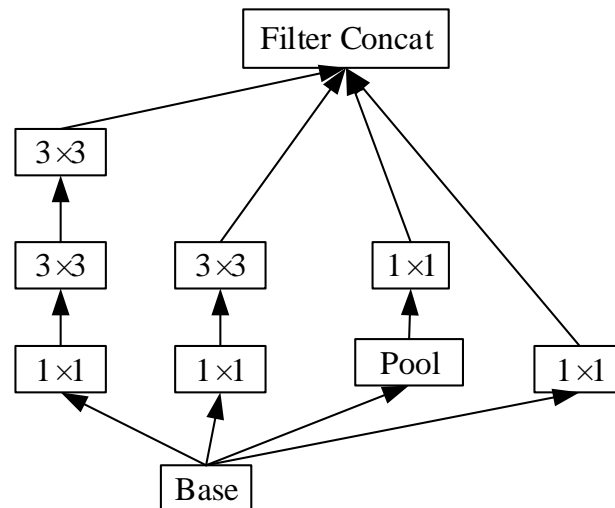


Figure 4. Inception structure.

This research adopts AlexNet model, which is a convolutional neural network model published by Alex Krizhevsky in 2012. Alex presented this network structure model at the 2012 ImageNet Image Classification Challenge and wins the championship of Beyond ImageNet Large Scale Visual Recognition. Due to AlexNet model is not too deep and has good classification ability, this study uses AlexNet model as the basic framework and optimizes remote sensing images classification. And the pooling layer of this study uses the maximum pooling method and non-overlapping sampling method. Its principle is to select the maximum value of the image region as the value of the region after sampling [24]. The improved convolutional neural network for classification is shown in Figure 5.

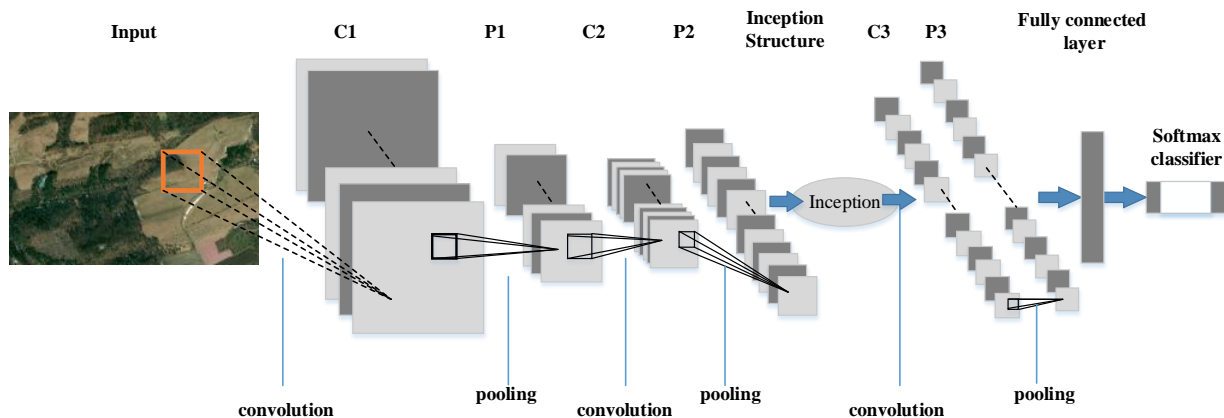


Figure 5. Improved convolutional neural network for classification.

The input layer of convolutional neural network is used to receive images, and the convolution layer is used to extract various features of images and reduce the impact of noise on classification [25].

Suppose the original input image is X , Y_i represents the feature map of the i -th layer, $Y_0 = X$, then

$$Y_i = f(K_i * Y_{i-1} + b_i) \quad (3)$$

In formula (3): K_i represents the weight of kernel of convolution layer i ; the operator $*$ represents convolution of the K_i with the feature map of layer $i - 1$; b_i represents the bias vector of layer i ; f is activation function. Compared with sigmoid activation function and tanh function, ReLU activation function can overcome the problem of vanishing gradient and accelerate the training speed [26]. Therefore, the method of this study uses ReLU function as the activation function. The expression of ReLU is:

$$\text{ReLU}(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases} \quad (4)$$

Generally, the pooling layer follows the convolution layer closely, and the feature map output from the previous pooling layer is sampled based on the local correlation of the image, and its scale remains unchanged. Generally, there are two kinds of action modes of pooling layer: Max pooling and mean pooling.

Convolution layer and pooling layer are connected alternately. Complete connection layer synthesizes the previously extracted features and reduces the image feature information from two-dimensional to one-dimensional. The final output layer generates a label corresponding to the sample based on the feature vector obtains by the fully connected layer.

The core of the classification process based on convolutional neural network lies in the training of the whole network, which is similar to the learning process of human brain. The process is divided into two stages. The first stage is forward propagation, so that the feature of the sample image is learned from the input layer to the output layer. The second stage is back propagation, which calculates the error between the actual output value and the expected output value according to the loss function, also known as "residual", and adjusts the network parameters according to the gradient descent method. Cross entropy loss function is the most widely used in convolutional neural network. Cross entropy is used to evaluate the difference between the probability distribution obtained from the actual output and the expected output of model training. Reducing the cross entropy is to improve the prediction accuracy of model. Its discrete function form is:

$$H(p, q) = - \sum_x p(x) \log q(x) \quad (5)$$

Here, $p(x)$ is the real distribution of data, $q(x)$ is the distribution of training. The larger the value of cross entropy, the greater the difference between the training sample and the distribution of the model. The goal of training convolutional neural network is to reduce the loss function of network through the gradient descent method.

When training a deep neural network, if the model has too many parameters and too few training samples, the trained model is prone to overfitting. The specific performance of overfitting is that the model has a small loss function on the training data and a high prediction accuracy but the test data has a large loss function and a low prediction accuracy. In order to solve the problem of overfitting, a model integration method is generally adopted to train multiple models for combination. However, this method will cause the problem that it takes too long to train the model and test multiple models. In 2012 and 2014, Hinton proposed dropout in his study [27]. When a complex

feedforward neural network is trained on a small data set, it is easy to cause overfitting. In each training batch, the over fitting phenomenon can be significantly reduced by ignoring half of the feature detectors. In forward propagation, the activation value of a neuron stops working with a certain probability p . This can make the model more general, because it does not rely too much on some local features.

In this way, the deep neural network can avoid from the time-consuming problem. The structure diagram of standard neural net is shown in Figure 6. The structure diagram of after applying dropout is shown in Figure 7.

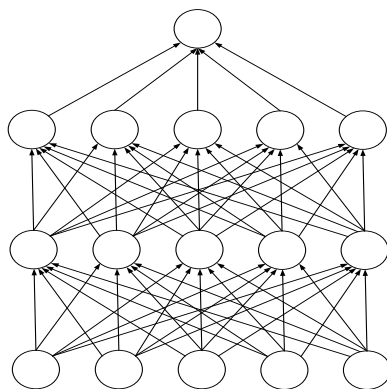


Figure 6. Standard neural net.

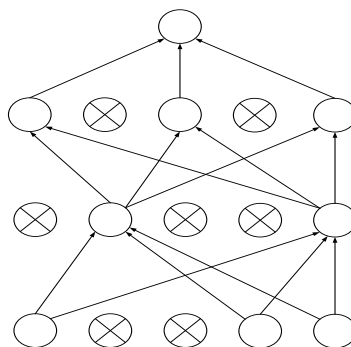


Figure 7. After applying dropout.

The Inception structure uses a small convolution kernel to replace a large convolution kernel, uses a non-linear saturation activation function to perform non-linear transformation. He obtained features are processed to achieve the application of multi-scale features. The addition of the Inception structure can reduce the parameters while increasing the depth and width of the network, thus optimizing the convolutional neural network classification effect.

The output layer of convolutional neural network usually uses a classifier, and the number of neuron nodes in the output layer depends on different classification tasks. Softmax classifier is based on the multinomial distribution model, and different classification probabilities can be obtained through the software classifier. Therefore, the classification performance of the softmax classifier is

better for a variety of non-overlapping categories.

For the given test input x , a probability value $p(y = j|x)$ is estimated for each category j , that is, the probability of each classification result for x is estimated. Suppose the function will output k -dimensional vector to represent the k estimated probability values. The system equation of softmax classifier for k -class classification is as follows:

$$h_{\theta}(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1|x^i; \theta) \\ p(y^{(i)} = 2|x^i; \theta) \\ \vdots \\ p(y^{(i)} = 3|x^i; \theta) \end{bmatrix} \quad (6)$$

3. Experimental results and analysis

To improve the classification effect of ground cover in multi-spectral remote sensing images, the study proposes a classification method that optimizes the traditional convolutional neural network framework and adds Inception structure. In order to prove the superiority of the improved convolutional neural network classification method, the study compares two traditional classification algorithms to classify the public satellite data of the National Oceanic and Atmospheric Administration (NOAA).

NOAA is the third generation of practical meteorological observation satellites from the National Oceanic and Atmospheric Administration. The first generation is called "TIROS" (1960–1965), the second generation is called "ITOS/NOAA" (1970–1976), and the third generation is called "TIROS-N/NOAA".

The purpose of NOAA satellite application is daily weather services. There are two satellites in operation. AVHRR is the main detection instrument of NOAA series satellites. Details of AVHRR data are shown in Table 1. There are two aspects in the application of AVHRR data. On the one hand, it is a large-scale regional (including national, continental, and global) survey. Which has advantages that other remote sensing cannot compare. The work that has been carried out includes the land cover surveys in the United States (Loveland et al. 1991), the land cover surveys in Africa (Tucker et al. 1985), the land cover surveys in South America (Townshend et al. 1987), the global land cover surveys (Defries 1994) and other surveys. On the other hand, it is a survey of small and medium-scale areas. The application of this aspect is mainly due to the difficulty of obtaining high-resolution remote sensing data now and the remote sensing surveys have poor live performance. Using the AVHRR data to obtain the macroscopic, good temporal resolution and accurate ground information.

The use of multiple bands or the selection of appropriate band combinations for classification is helpful to overcome the homology of foreign objects, which can improve the classification accuracy. Using the AVHRR data received in 1998, and performing projection transformation and geometric correction on it, three band data sets were generated (AVHRR has five bands, only three bands are used here). The test image selected in this study is a remote sensing image of a farm in the United States in the public satellite data set of NOAA as shown in Figure 8. The classification results based on support vector machine and radial basis function are compared with the improved convolutional neural network classification method proposed in the study, as shown in Figures 9–11.

Table 1. Details of AVHRR data.

Channel	Wavelength (μm)	Waveband	Ground resolution (km)	Application
AVHRR-1	0.58–0.68	Visible light	1.10	Daytime clouds, ice, snow, vegetation
AVHRR-2	0.725–1.10	Near-infrared	1.10	Daytime clouds, vegetation, water, agricultural estimation, land usage survey
AVHRR-3A	1.58–1.64	Middle-infrared	1.10	Daytime clouds, ice, snow, soil moisture, drought monitoring
AVHRR-3B	3.55–3.93	Middle-infrared	1.10	Night clouds, forest fire, volcanic activity
AVHRR-4	10.30–11.30	Far-infrared	1.10	Day and night image, land surface temperature, sea surface temperature
AVHRR-5	11.50–12.50	Far- infrared	1.10	Day and night image, land surface temperature, sea surface temperature

In this study, 10% samples are randomly selected as the training set. In order to verify the effectiveness of the improved convolutional neural network classification method in the classification task, the above experiments use three different algorithms (support vector machine, radial basis function, improved convolutional neural network classification method) on the same data set (NOAA) for verification. The criteria for measuring the effectiveness of classification are: User's accuracy, commission error, overall accuracy, Kappa coefficient and other factors. Before the experiment, all classification algorithms are set up under the same environment configuration. TensorFlow 1.1.0 open source framework is adopted. The built environment is PC, the operating system is Ubuntu 16.04, the processor is Intel (R) Xeon (R) CPU E5-1603 v3 @ 2.80 GHz, the graphics card is NVIDIA Quadro K2200 version, the running memory is 16 G, and the CUDA version is 8.0. The improved convolutional neural network classification method adopts AlexNet model, the pooling layer adopts max-pooling, the gradient descent method is used to adjust the selected cross-entropy loss function. Inception structure is added to the traditional convolutional neural network classification model, which makes use of its non-linear change ability and increases the network width by parallel convolution layers of different scales, thus improves the feature extraction ability.



Figure 8. Remote sensing image of a farmland in the United States published by NOAA.

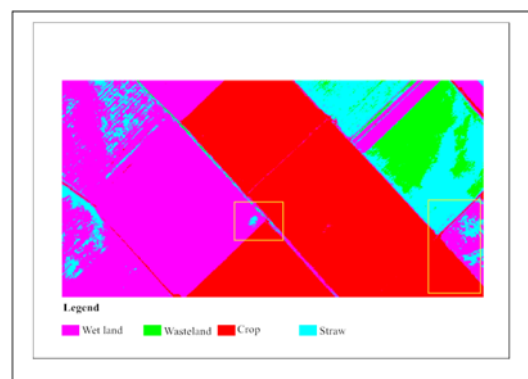


Figure 9. Classification results for Figure 8 based on support vector machine.

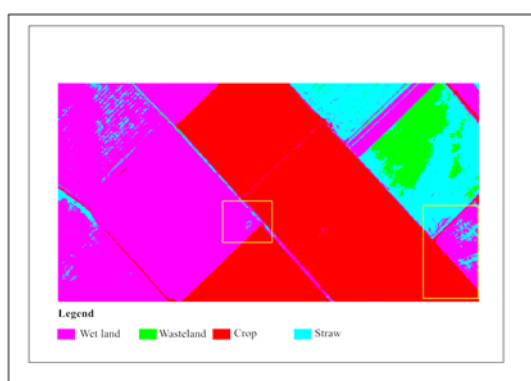


Figure 10. Classification results for Figure 8 based on radial basis function .

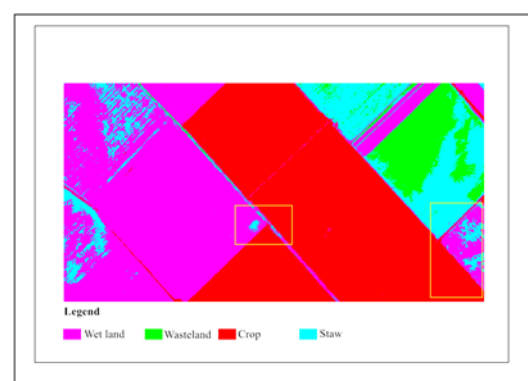


Figure 11. Classification results for Figure 8 based on Inception in convolutional neural network.

According to the characteristics of the image, the target types are divided into wetland, wasteland, crop and straw. In order to test the accuracy of image classification, 200 samples (800 samples in total) were randomly selected for each target type for analysis, and the confusion matrix of the classification results shown in the table was obtained. The evaluation results are shown in Tables 2–4. According to the experimental classification results, the commission errors of support vector machine classification method is more than 6 times that of the improved convolutional neural network classification method. Especially for the terrain with complex features such as straw, the commission errors of support vector machine classification method are much higher than that of the improved convolutional neural network classification method. The accuracy of radial basis function classification method is relatively high when it is used to classify large areas of ground objects, but it is not enough for the local confused ground objects. However, the overall accuracy of support vector machine classification method and radial basis function classification method is far less than that of the improved convolutional neural network classification method. The advantages of using the proposed classification method are highlighted through the comparative experiments. Therefore, it is feasible to use the proposed classification method based on the improved convolutional neural network.

Table 2. Evaluation results of support vector machine.

Remote sensing image	Support vector machine classification accuracy evaluation			
	User's accuracy	Commission error	Production accuracy	Omission error
Wet land	87.62%	12.38%	99.28%	0.72%
Wasteland	93.25%	6.75%	73.43%	26.57%
Crop	81.55%	18.5%	84.85%	15.15%
Straw	84.48%	15.52%	86.73%	13.27%
Overall accuracy				87.52%
Kappa coefficient				0.8223

Table 3. Evaluation results of radial basis function.

Remote sensing image	Radial basis function classification accuracy evaluation			
	User's accuracy	Commission error	Production accuracy	Omission error
Wet land	83.45%	16.55%	83.16%	16.84%
Wasteland	78.89%	21.11%	78.02%	21.98%
Crop	83.43%	16.57%	82.82%	17.18%
Straw	76.87%	23.13%	80.71%	19.29%
Overall accuracy				82.01%
Kappa coefficient				0.7457

Table 4. Evaluation results of improved convolutional neural network.

Remote sensing image	Improved convolutional neural network classification accuracy evaluation			
	User's accuracy	Commission error	Production accuracy	Omission error
Wet land	97.86%	2.14%	98.92%	1.08%
Wasteland	98.07%	1.93%	98.07%	1.93%
Crop	97.94%	2.06%	95.96%	4.04%
Straw	98.21%	1.79%	97.35%	2.65%
Overall accuracy				97.99%
Kappa coefficient				0.9715

4. Conclusions

The study improves the classification method based on convolutional neural networks, and proposes the idea of adding an Inception structure. The Inception structure can perform non-linear transformation. It can process the obtained features to achieve the application of multi-scale features. Inception uses a parallel convolution kernel to increase the network width, and it is located at a higher number of layers, so it improves the ability of network feature extraction. This is the key to further improve the classification effect of high-resolution multi-spectral remote sensing images. The improved convolutional neural network model adopts AlexNet model and adds Inception structure to improve the classification effect of the network. Softmax classifier also plays a great role in improving the classification accuracy of the network. The experimental results show that the improved convolutional neural network classification method used in this research improves the

overall accuracy of high-resolution multi-spectral remote sensing images classification by about 10%. The commission errors of the improved convolutional neural network model are much smaller than that of the classification method based on support vector machine and radial basis function. The improved convolutional neural network model improves the overall accuracy and classification effect of multi-spectral remote sensing images classification.

Acknowledgments

This work was supported by National Science and Technology Major Project of High-Resolution Earth Observation (70-Y40-G09-9001-18/20), Liaoning Provincial Natural Science Foundation of China (20180550334), Key Project of Ministry of Education of China (2017A02002), Liaoning education department science and technology research project (L201701, L201704 and L201735). The authors deeply appreciate the supports.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. W. J. Yan, D. H. Chen, L. Liu, Research progress of hyperspectral image classification, *Opt. Precis. Eng.*, **27** (2019), 680–693.
2. S. Y. Chen, H. Z. Lin, X. Zhao, G. Wang, Deep learning-based classification of hyperspectral data, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **7** (2014), 2094–2107.
3. F. W. Fu, B. W. Zou, Review of remote sensing image classification based on deep learning, *Appl. Res. Comput.*, **35** (2018), 3521–3525.
4. J. Zhao, Y. Zhong, H. Shu, L. P. Zhang, High-resolution image classification integrating spectral-spatial-location cues by conditional random fields, *IEEE. Trans. Image Process. Publ. IEEE Signal Process. Soc.*, **25** (2016), 4033–4045.
5. A. B. Salberg, Detection of seals in remote sensing images using features extracted from deep convolutional neural network, *Geosci. Remote Sens. Symp. IEEE*, (2015), 1893–1896.
6. C. Cortes, V. Vapnik, Support-vector networks, *Mach. Learn.*, **3** (1995), 273–297.
7. D. C. Feng, G. Chen, W. Y. Du, X. Y. Wu, K. L. Xiao, Remote sensing image classification based on minimum distance method, *J. Beihua. Inst. Aerosp. Technol.*, **22** (2012), 1–3.
8. A. Ahmad, S. Hashmi, K-Harmonic means type clustering algorithm for mixed datasets, *Appl. Soft Comput.*, (2016), 39–49.
9. Y. Zhang, H. T. Yang, C. H. Yuan, A survey of remote sensing image classification methods, *J. Weapon Equip. Eng.*, **39** (2018), 108–112.
10. M. Volpi, D. Tuia, Dense semantic labeling of subdecimeter resolution images with convolutional neural networks, *IEEE Trans. Geosci. Remote Sens.*, (2017), 1–13.
11. G. E. Hinton, R. R. Salakhutdinov, Reducing the dimensionality of data with neural networks, *Sci.*, **313** (2006), 504–507.
12. W. Y. Pang, L. M. Sun, H. X. Jiang, L. X. Li, Convolution in convolution for network in network, *IEEE Trans. Neural Networks. Learn. Syst.*, **29** (2018), 1587–1597.

13. A. S. Razavian, H. Azizpour, J. Sullivan, S. Carlsson, CNN features off-the-shelf: An astounding baseline for recognition, *Comput. Vision Pattern Recognit. Workshops. IEEE*, (2014), 512–519.
14. J. P. Zhao, W. W. Guo, S. Y. Cui, Z. H. Zhang, Convolutional neural network for SAR image classification at patch level, *Geosci. Remote Sens. Symp. IEEE*, (2016), 945–948.
15. H. B. Lyu, H. Lu, L. C. Mou, Learning a transferable change rule from a recurrent neural network for land cover change detection, *Remote Sens.*, **8** (2016).
16. N. Kussul, M. Lavreniuk, S. V. Skakun, A. Y. Shelestov, Deep learning classification of land cover and crop types using remote sensing data, *IEEE Geosci. Remote Sens. Lett.*, **14**(2017), 778–782.
17. R. Marc, K. Marco, Multi-temporal land cover classification with sequential recurrent encoders, *ISPRS Int. J. Geo-Inf.*, **7** (2018).
18. Y. Liu, Z. S. Liao, Multiple kernel learning with the generalized error bound of support vector machine, *J. Wuhan Univ. (Nat. Sci. Ed.)*, **58** (2012), 149–156.
19. Q. J. Zhang, X. J. Zhang, Z. Zhao, Y. J. Wang, Classification of polarimetric SAR images based on multiscale segmentation and radial basis function neural network, *Geomat. Spat. Inf. Technol.*, (2019), 67–71.
20. M. Zhou, Dimension reduction and classification of hyperspectral remote sensing image based on RBF neural network, *Territ. Nat. Resour. Study*, (2016), 14–16.
21. P. V. Arun, K. M. Buddhiraju, A. Porwal, CNN based sub-pixel mapping for hyperspectral images, *Neurocomput.*, **311** (2018), 51–64.
22. A. Krizhevsky, I. Sutskever, G. Hinton, Image net classification with deep convolutional neural networks, *NIPS*, **60** (2017), 84–90.
23. C. Szegedy, V. Vanhoucke, S. Loffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, *IEEE Conf. Comput. Vision Pattern Recognit.*, 2016.
24. Krizhevsky, Alex, Sutskever, Ilya, Hinton, E. Geoffrey, Image net classification with deep convolutional neural networks, *Commun. ACM*, **60** (2012), 84–90.
25. Q. Z. Gong, P. Zhong, Y. Yu, D. W. Hu, Diversity-promoting deep structural metric learning for remote sensing scene classification, *IEEE Trans. Geosci. Remote Sens.*, (2017), 1–20.
26. X. Glorot, A. Bordes, Y. Bengio, Deep sparse rectifier neural networks, *J. Mach. Learn. Res.*, 2011.
27. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: A simple way to prevent neural networks from overfitting, *J. Mach. Learn. Res.*, (2014), 1929–1958.



AIMS Press

©2020 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)