Application of deep neural networks in classification of medium resolution remote sensing image

WANG Dian-lai¹, SU Ai-xia², LIU Wen-ping³

¹Department of Information Engineering, Shou Gang Institute of Technology, Beijing 100144, China;
²China Software Testing Center, Beijing 100048, China;
³College of Information, Beijing Forestry University, Beijing 100083, China.

achenhanjun@tiangong.edu.cn

Abstract. Aiming at the problem that high misclassification rate of remote sensing image occurs due to the phenomenon of "same spectrum different matter" and "same object different spectrum", the deep neural networks (DNN) is proposed to the classification of medium resolution remote sensing image. The deep neural networkstake advantage of the neural network with multiple hidden layers to learn the characteristics that can describe the essential attributes of data, and has achieved high accuracy in land cover classification in Miyun District, Beijing city. The experimental results indicate that the classification accuracy of this algorithm is the highest compared with SVM and BP neural network, the classification accuracy reaches 96.53%, and the classification accuracy is 10.39% higher than that of SVM. The kappa coefficient of DNN is 0.95, which is also the highest among the comparison algorithm. So, the DNN can apply to the classification of moderate resolution remote sensing image.

1. Introduction

There are a large number of phenomena of "different objects of the same spectrum" and "different spectra of the same objects" in remote sensing image, resulting in a high rate of misclassification of it. Using the spectral and texture features of remote sensing images, some methods based on statistical theories, such as minimum distance method, maximum likelihood method, KNN, etc., are applied to classification of remote sensing image. With the development of artificial intelligence algorithm, some new methods such as artificial neural network[1], support vector machine (SVM)[2], genetic algorithm[3], object-oriented classification[4] have also been applied to remote sensing image interpretation. Although the above methods have achieved good classification results, the degree of automation is not high[5], and the classification accuracy cannot satisfy some conditions. Therefore, classification algorithm with higher classification accuracy and higher automation degree is needed to solve a large number of remote sensing image classification problems.

In 2006, Hinton[6], et al. proposed the concept of deep learning for the first time. They pointed out that the training difficult problems of deep neural network can be solved through initialization layer by layer. Deep learning algorithm has been successfully applied to the following fields, such as image field, voice field, video field, text field and data analysis field[7]. Its core idea of deep learning is to let machines learn from data automatically by increasing the number of layers of the network[8].

The deep learning algorithm provides a new method for remote sensing image interpretation, and a large number of deep learning algorithm research in the field of remote sensing image classification...
has emerged. Deep belief network (DBN) model was used to study the classification of Toronto area in Canada by remote sensing image, and the classification accuracy was higher than that of SVM and NN algorithm (in 2014, LU Qi, et al.)[9]. CNN was applied to classify hyper spectral remote sensing images, and the classification accuracy was higher than that of SVM algorithm (in 2015, Wei Hu, et al.)[10]. DBN algorithm was employed to classify high-resolution remote sensing images using spectral and texture features, and the classification accuracy was higher than that of SVM and NN algorithms (in 2016, Liu Dawei)[11]. CNN was used to study the classification of Qinghai Lake Region by remote sensing images, and the classification accuracy was higher than that of the maximum likelihood method and SVM (In 2018, Ma Kai and Luo Ze)[12]. CNN was employed to verify the effectiveness of classification using improved feature extraction based remote sensing image (in 2018, Lu Mengxi)[13]. CNN and deeplav3 were applied to extract water bodies using high-resolution remote sensing images, and the classification accuracy reached 95.09% and 92.14% respectively (in 2019, Chen Qian, et al.)[14].

Based on the above analysis, it can be seen that deep learning algorithm is an effective method in the classification of high-resolution or hyper-spectral remote sensing image, and DBN and CNN are also widely used. In the classification of medium-resolution remote sensing image, the application of deep learning algorithm is limited by more mixed pixels and difficulty of extracting training samples. Aiming at these problems, deep learning algorithm is employed to research the classification of medium resolution (Landsat-5 TM) remote sensing image, and a deep neural network with seven hidden layers is designed to study the classification of land cover in Miyun District, Beijing city. The designed network takes advantage of neural network characteristics, such as initialization layer by layer and the reverse fine-tuning, and the classification accuracy is greatly improved. It is a useful attempt of the deep learning algorithm in the remote sensing image classification with medium resolution. The results show that the proposed algorithm is superior to the contrast algorithm in terms of overall classification accuracy and kappa coefficient.

2. Materials and Methods

2.1 data and data preprocessing
Landsat5-TM remote sensing image was used for research. The Landsat5-TM image contains seven bands, the spatial resolution is 30 meters, and only the sixth band’s spatial resolution is 120 meters. For more parameters of each band of landsat-5 TM, please refer to the relevant literature. Data was downloaded from website https://lta.cr.usgs.gov/TM, the data’s date is on July 12 2006, and data’s path/row is 124/32. The download data was processed using the software ENVI4.5, and only the area of Miyun District, Beijing city was left by masking and clipping operation. The image’s size is 2531*2376 pixels.

For analysis the land cover types in study area, the landsat-5 TM multispectral images were combined to a RGB image, red channel was replaced by band4, band3 took place of green channel, blue channel was substituted for band2. The sub-graph(a) in Fig.1 indicated that there are three types of land cover in study area: red color represents area of vegetation, blue color means water bodies area (Miyun reservoir), and the light green color shows a city or bare soil area.

According to the types of land cover in study area, three training samples, such as water bodies, vegetation and urban or bare soil were selected. The training samples were selected with the help of ROI tools of EVNI. Four thousands samples of each land cover type were selected, including band1 to band7. The four thousands samples of each land cover type were divided into two groups, the two thirds is for training, and onethird is for testing.

2.2 Methods
The important factors which effect the algorithm, such as network structure, parameter setting and data normalization are briefly described in the following sections.
2.2.1 The network structure of DNN and parameters setting

(1) Determine the structure of DNN neural network. Firstly, the number of input layer and output layer is determined. The training data in this paper includes 7 bands of Landsat5-TM B1~B7, and the input data is 7 dimensions, so the number of points of the input layer is 7. The surface in the study area is covered by water body, urban area or bare soil and vegetation, so the number of points of output layer is 3. Secondly, the number of hidden layers needs to be analyzed according to specific problems. It is generally believed that increasing the number of hidden layers can reduce the network error and improve the accuracy, but it also complicates the network, increases the training time of the network and increases the probability of "overfitting" phenomenon. Generally speaking, at least one hidden layer should be included. In order to make the model have better data characterization ability, the number of hidden layers in this paper is set as 7. Finally, the choice of the number of contacts in the hidden layer is equally important. Too small number will reduce the ability of feature extraction of the network and lead to the phenomenon of "underfitting". If the number is too high, the model will be too complex, which will lead to the phenomenon of "overfitting". Meanwhile, the excessively complex network structure will make the training time too long. After several rounds of test results, the empirical value is 64. In order to avoid the occurrence of "overfitting", the proportion of random dropout nodes in the training process is 30% to improve the generalization ability of the model.

(2) The input data is normalized and the label data of the sample is constructed. The values of the spectral data B1~B7 are normalized to the interval [-1,1]. Let Y1=[1,0,0]T, Y2=[0,1,0]T, Y3=[0,0,1]T, Y1 represents vegetation, Y2 represents water body, and Y3 represents urban area or bare soil. The above vectors are used to generate sample label data.

(3) DNN training parameter setting. The following parameters were set in the DNN model: learning rate, activation function, loss function, optimization function, batch size, and epoch.

The learning rate determines how much to learn and how much to update parameters in training procedure. In general, this value is too large or too small would prevent the DNN algorithm to achieve a good value. In this paper, the network learning rate is set as 0.009.

For deep learning algorithm, the choice of activation function is very important. Assume that there are \( n \) neurons \( x_1, \ldots, x_n \), with weights of \( w_1, \ldots, w_n \) and bias of \( b \), then the \( y \) is the output of neuron, and the formula of \( y \) value is shown in formula (1), where \( f \) is the activation function, that is, the output of the nerve needs to be output after the transformation of the activation function. The activation function \( f \) usually needs to satisfy the following characteristics: (1) the activation function must be a continuously differentiable nonlinear function. (2) The activation function and its derivative form must be simple to ensure network learning efficiency. (3) The derivative of the activation function should not be too large or too small, but should be stable around 1. If the value is too large, it will lead to gradient explosion; otherwise, it is easy to cause gradient disappearance. Based on the above principles, Swish activation function is selected instead of the traditional ReLU. The formula of Swish activation function is shown in formula (2). The range of Swish \((x)\) function is \((-1,1)\), and \(\beta\) is a super parameter, where \(e\) is Napier's constant, and its value is 2.7182. As \(\sigma(\bullet)\) approaches 1, \(y\) approaches \(x\). As \(\sigma(\bullet)\) approaches 0, \(y\) approaches 0. In Keres, Swish\((x)=x\) k.sigmoid \((x)\). K is a Backend for Keres.

\[
y = f(\sum_{i=1}^{n} w_ix_i + b) \quad (1)
\]

\[
Swish(x) = x\sigma(\beta x) = \frac{x}{1 + \exp(-\beta x)} \quad (2)
\]

The selection of parameter updating methods is particularly important. Common parameter updating methods include SGD, Momentum, Agrad and Adam, etc. The above methods have their own applicable conditions, and Adam method is selected as the parameter updating method in this paper. CrossEntropy is selected as the loss function, and its formula is shown in formula (3). Log
stands for natural Log base $e$, which $t_k$ is the correct label, $y_k$ is the output of the neural network. In keres the parameter is: Loss = 'categorical_crossentropy'.

\[ E = -\sum_k t_k \log y_k \quad (3) \]

For achieving a balance between efficiency and accuracy in DNN network training, batch size and epoch were set as 128 and 64 respectively in this paper.

2.3 The main steps of DNN algorithm
The main step of DNN algorithm is described as following:

1) The pre-processing of remote sensing image data mainly includes remote sensing image correction, registration, cutting and mask, and the remote sensing image containing band b1~b7 in Miyun District of Beijing is finally obtained, as detailed in 2.1 of this paper.

2) For training and testing, three types of sample selection was selected, such as water, vegetation and urban area or bare soil. This paper uses the ROI tool of ENVI software to pick the three types of the sample.

3) For the construction and training of DNN model, the above three types of samples were used. The model structure was shown in 2.2.1. After the training, the model structure and parameters were saved for subsequent image classification.

4) The trained DNN model in step (3) was used to classify the pending score data in the study area;

5) Judge whether all the classifications have been completed. If all the classifications have been completed, the classification results will be output and the algorithm will end. Otherwise wait until all classes have been executed;

3. Results and Discussion

3.1. Environment of experiments and tools
Experimental environment is as follows: The operating system is window7.0. Eclipse4.5.1 is used to process the data, such as, masking, picking up the training samples etc. The java development environment is JDK1.8. The remote sensing image extracting tools is VGTExtract. The algorithm of SVM and BP neural network was coded with Java, based on OpenCV API of java version. The deep neural network is implemented based on Keres API with python3, and keres’s Version is 2.3.0.

3.2. Results
Sub-graph(a) in Fig.1 is a false color image combined by landsat5-TM in Band B4, B3 and B2 using tools of ENVI4.5, in which the water body is blue, the vegetation is red, and the urban area or bare soil of Miyun is light green. Sub-graph (a) clearly displays the land cover types in Miyun District, Beijing city.

Sub-graph (b), sub-graph (c) and sub-graph (d) in Fig. 1 show the land cover classification results in Miyun district, Beijing city by SVM, BP neural network and DNN algorithm respectively. In sub-graph (b), (c) and (d) in fig.1, the land cover type vegetation represented by green, water body marked using blue, and urban area or bare soil represented by brown. Table 1 shows the fuzzy classification matrix of SVM, BP neural network and DNN algorithm. Table 2 shows the overall classification accuracy of the three algorithms and the Kappa coefficient.

| Table 1 classification in confusion matrix for each classification method |
|-----------------------------|-------|-----------------|-----------------|--------|-----------------|
| method         | vegetation | water | Urban area/bare soil | Sum of row | User's accuracy (%) |
| SVM            | Vegetation | 709   | 24   | 267   | 1000 | 70.90%       |
|                | Water      | 30    | 920  | 50    | 1000 | 92.00%       |
|                | Urban area/bare soil | 91 | 60 | 849 | 1000 | 84.90%       |
| Sum of column  | 830 | 1004 | 1166 | 3000 |     |               |
Producer’s accuracy (%) 85.42% 91.63% 72.81%

|          | vegetation | water | Urban area/bare soil | Sum of row | User’s accuracy (%) |
|----------|------------|-------|-----------------------|------------|---------------------|
| BP       | 683        | 5     | 312                   | 1000       | 68.30%              |
|          | 5          | 983   | 12                    | 1000       | 98.30%              |
|          | 81         | 33    | 886                   | 1000       | 88.60%              |
|          | 769        | 1021  | 1210                  | 3000       |                     |
| Producer’s accuracy (%) | 88.82% | 96%  | 73.22% |

DNN

|          | vegetation | water | Urban area/bare soil | Sum of row | User’s accuracy (%) |
|----------|------------|-------|-----------------------|------------|---------------------|
|          | 961        | 15    | 24                    | 1000       | 96.10%              |
|          | 18         | 972   | 10                    | 1000       | 97.20%              |
|          | 25         | 12    | 963                   | 1000       | 96.30%              |
|          | 1004       | 999   | 997                   | 3000       |                     |
| Producer’s accuracy (%) | 95.72% | 97.30% | 96.59% |

Table 2 Comparison of accuracy and Kappa coefficient to each classification method

|                  | SVM       | BP       | DNN      |
|------------------|-----------|----------|----------|
| Overall accuracy | 82.60%    | 85.07%   | 96.53%   |
| Kappa coefficient| 0.74      | 0.78     | 0.95     |

Fig.1 Classification results of land cover in Miyun District, Beijing city using SVM, BP and DNN algorithms.

(a) false color image combined by b4,b3 and b2.  
(b) classification result of SVM algorithm.  
(c) classification result of BP algorithm.  
(d) classification result of DNN algorithm.
3.3. Discussion
Comparing sub-graph(a) and sub-graph(b)–(d) in Fig.1, it can be seen that the texture of the classification results of sub-graph(d) is basically the same as that of sub-graph(a). This indicates that DNN algorithm shows higher classification accuracy in Miyun urban area.

There are obvious differences in Miyun district, Beijing city by comparing sub-graph(a) to sub-graph (b) and sub-graph(c) in Fig.1. This means that the classification accuracy of SVM and BP neural network is lower than that of DNN, and a large number of vegetation is improperly divided into urban area or bare soil. The texture of sub-graph(a) and sub-graph(d) is closely same, especially in Miyun district, Beijing city. The main cause of the phenomenon occur is that the DNN algorithm can take advantage of multi-layer network structure to learn the characteristics difference between vegetation with urban area or bare soil. So the classification results of DNN is higher than that of SVM and BP algorithm. According Fig.1, the classification results of DNN is highest in three algorithms, and its texture of classification results is closely same as the actual land cover condition. All above this also can be found in table2 by comparing the overall accuracy.

It can be seen from Table2 that the classification accuracy of SVM, BP, KNN and DNN algorithms is 82.60%, 85.07% and 96.53% respectively. Experimental data show that the overall SVM classification algorithm is the lowest, and DNN algorithm has the highest classification accuracy, which is 10.39 percentage points higher than SVM classification accuracy.

Based on the multi-layer neural network structure, the DNN algorithm can learn the difference of various samples more accurately through a large amount of sample information. Combined with the Adam optimization method, the optimal solution of the loss function can be found, and its classification accuracy reaches 96.53%. By comparing sub-graph (a) and sub-graph (d) in Fig.1, there is a high consistency in texture, and the classification results are basically consistent with the actual ground cover types.

Kappa coefficient is an important index to evaluate the consistency and reliability of classification results. Table2 shows that Kappa coefficient of DNN algorithm is the highest, reaching 0.95, indicating a good consistency of classification results.

Generalization is another important index to evaluate the advantages and disadvantages of the algorithm. In the case of neural network, the algorithm has a high classification accuracy on the training samples, but performs poorly on the test samples, that is to say, the phenomenon of "over-fitting" occurs. The comparison curves between training accuracy and testing accuracy are shown in Fig.2. It is seen from Fig.2 that the values and changing trends of the two curves are very closely. This means the testing accuracy is consistent with the training accuracy. Fig.2 indicate that there is no "over-fitting" phenomenon in this algorithm, and also means the neural network structure designed in this paper has good generalization.

![Figure 2. Comparison curves of accuracy between training data and testing data.](image-url)
4. Conclusion

In this paper, the deep neural networks is applied to the classification of medium-resolution remote sensing images. Experimental results show that the overall classification accuracy of DNN algorithm reaches 96.53%, 10.39 percentage points higher than that of SVM algorithm, and the classification accuracy is also higher than that of BP neural network. The Kappa coefficient of this algorithm is also the highest among the three algorithms, and it reaches 0.95. Deep neural networks can take advantage of multi-layer network structure, learn from data, accurately depict various data features, and achieve the high-precision classification of remote sensing image. It can be seen that the deep neural networks is effective algorithm in the classification of medium resolution remote sensing image.

The further study expect to focus on the regions which the land cover types are more complex than that in this paper to utilize DNN algorithm’s advantages. The more validations of algorithm need to be done, and CNN network is also to be employ to improve the precision of classification.

References

[1] Canty M J. Boosting a fast neural network for supervised land cover classification. [M]. Oxford: Pergamon Press, Inc., 2009.

[2] Huang C, Davis L S, Townshend J R G. An assessment of support vector machines for land cover classification [J]. International Journal of remote Sensing, 2002, 23 (4): 725-749.

[3] Tseng M H, Chen S J, Hwang G H, et al. A genetic algorithm rule-based approach for land-cover classification [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2008, 63 (2): 202-212.

[4] Huan Yu, Zhang Shuqing, Kong Bo, et al. Optimal segmentation scale selection for object-oriented remote sensing image classification [J]. Journal of Image and Graphics, 2010, 15(2): 352-360.

[5] Fu Weifeng, Zou Weibao. Review of remote sensing image classification based on deep learning[J]. Application Research of Computers, 2018, 35(12):3521-3525.

[6] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets [J]. Neural Computation, 2006, 18(7): 1527-1554.

[7] FU Wen-bo, SUN Tao, LIAN Ji, YAN Bao-wei, FAN Fu-xin. Review of principle and application of deep learning[J]. Computer Science, 2018, 45(S1):11-15+40.

[8] JIAO Li-Cheng, YANG Shu-Yuan, LIU Fang, WANG Shi-Gang, FENG Zhi-Xi. Seventy Years Beyond Neural Networks: Retrospect and Prospect[J]. Chinese Journal of Computers, 2016, 39(08):1697-1716.

[9] Lu Qi, Dou Yong, Niu Xin, Xu Jiaqing, Xia Fei. Remote sensing image classification based on DBN model[J]. Journal of computer research and development, 2014, 51(09):1911-1918.

[10] Wei H , Yangyu H , Li W , et al. Deep Convolutional Neural Networks for Hyperspectral Image Classification[J]. Journal of Sensors, 2015, 2015:1-12.

[11] Liu Dawei, Han Ling, Han Xiaoyong. High Spatial Resolution Remote Sensing Image Classification Based on Deep learning[J]. Acta Optica Sinica, 2016, 36(04):306-314.

[12] MA Kai, LUO Ze. Classification of Remote Sensing Images in Qinghai Lake Based on Convolutional Neural Network[J]. Computer System & Application, 2018, 27(09):137-142.

[13] Lu Mengxi. Research about Remote Sensing Image Classification Based on Deep learning[D]. Harbin Engineering University, 2018.

[14] CHEN Qian, ZHENG Li-juan, LI Xiao-juan, XU Chong-bin, WU Yu, XIE Dong-hai, LIU liang. Water Body Extraction from high-resolution satellite Remote Sensing Images Based on Deep Learning[J]. Geography and Geo-Information Science, 2019, 35(04):43-49.