Remote Sensing Image Recognition Using Deep Belief Network

Min Li(✉)

School of Computer Engineering, Jingchu University of Technology, Jingmen 448000, Hubei, China
49035084@qq.com

Abstract. How to acquire high-dimensional data such as remote sensing image efficiently and accurately has become a research hotpot recent years. Deep learning is a kind of learning method which uses many kinds of simple layers to learn the mapping relation of complex layers. The authors will attempt to apply the deep belief network model (DBN), which is important in deep learning, to remote sensing image recognition. Using the new large-scale remote sensing image data set with abundant changes as the research object, the hierarchical training mechanism of DBNs is studied and compared with CNNS, the results show that the accuracy and speed of DBNs is better than that of CNNS, and more effective information can be obtained.

Keywords: Remote sensing image recognition · DBNs · CNNs

1 Introduction

With the development of remote sensing technology, remote sensing image processing has become the focus of research in many fields such as security, aerospace, medical, scientific research and so on. In 2016, Deng et al. [1] applied a method of remote sensing image classification based on Fisher-BP to improve the efficiency and accuracy of remote sensing image classification. In 2015, Li et al. [2] used SVM, K-means and limit learning to identify and classify the bad geological objects in remote sensing images. In foreign countries, Mantero et al. [3] have put forward a classification method based on the minimum error decision in the literature, which can well deal with the recognition and classification of remote sensing ground reference data. In the early 1970s, many image analysis methods using remote sensing images were developed to analyze each pixel. With the development of remote sensing technology, the spatial resolution becomes more and more fine, the pixels are not isolated, but filled into the image which is full of spatial pattern. For some typical ground use recognition tasks, pixel or even super-pixel, all existing data sets have some limitations, which severely restrict the development of new data-driven Algorithms. With this in mind, Cheng et al. [4] in 2016 presented a large-scale benchmark data set named "nwpu-resc45". This data set overcomes the limitations of the existing data set, such as the small scale of the number of class images and the

W. Li and D. Tang (Eds.): MOBILWARE 2020, LNICST 331, pp. 187–194, 2020.
https://doi.org/10.1007/978-3-030-62205-3_18

total number of images, the lack of scene changes and diversity, and the saturation of classification accuracy, so this paper chooses this data set as the research object of the deep belief network model. Remote sensing image recognition and classification Image recognition is based on the existing information in memory to judge the information into the sensory organs at this time, so as to achieve the re-recognition of the image. The recognition of remote sensing image is related to the recognition of common images. In urban planning and management, the impact of remote sensing image classification is very important [5].

In recent years, deep learning has become a new direction in the field of machine learning. By simulating the multi-level structure of human brain, feature data are extracted from the bottom layer to the top layer in order to find the regularity of data in time and space and improve the accuracy of classification. The classification of remote sensing image is mainly based on the characteristics of remote sensing image of electromagnetic radiation. The classification map can be used as an intermediate result of other applications, such as target detection and recognition, to provide auxiliary information, and as the final result of basic geographic information in other fields such as resource management, disaster relief, urban planning, etc. According to whether the prior knowledge of the data is needed, the remote sensing classification method is divided into two kinds, one is parameterized and the other is non-parameterized. The first method consists of a maximum likelihood classifier (MLC) [6], a minimum distance classifier (MDC) [7], and an Expectation-Maximization (EM) algorithm [8], all of which presupposes the distribution of data. However, there are also data distribution laws are often difficult to predict the situation, such as multi-time and multi-source remote sensing data. Therefore, the second method is more widely used in remote sensing image classification, including decision tree [9], artificial neural network (ANN) [10], Support vector machine (SVM) and so on [11–13]. The classification of commonly used remote sensing images mainly embodies in two aspects. One is unsupervised classification. It is a clustering analysis method, no training samples in advance, no label information, the data needs to be directly modeled. There are EM algorithm, K-MEANS clustering algorithm, self-coding Algorithm and so on. The other is the supervision of classification. It predicts the class attributes of unknown data instances based on the association pattern between the known data attributes and the class attributes. Common Algorithms include Support vector machine, linear regression, neural network, decision tree, and KNN. Deep belief network model.

At present, general machine learning belongs to the category of shallow learning, which is relatively weak in characterizing complex data or features. In this paper, we introduce the Deep Belief Network (DBN), which uses the layered mechanism to improve the training speed and the ability to deal with complex classification problems. The most common network models used in deep learning include SAE, DBN, and CNN. DBN is the most common and classic model.

2 Premilaries

DBN is a probability generation model consisting of a series of RBM units. RBM is a two-layer undirected graph with no connected models between each layer of nodes.

Where the input visible layer v, h, h$'$. It is a hidden layer that represents feature extraction. Visible layer v Visible unit m Hidden layer h Hidden unit n. The unit, and all visible hidden units are usually random two value variable nodes (only values of 0 and 1) whose distribution satisfies the Bernoulli distribution.

RBM is a typical energy-based model consisting of a visible layer and a hidden layer. The joint configuration energy of these two layers is expressed as:

$$E(v, h; \theta) = -\sum_i \sum_j w_{ij} v_i h_j - \sum_i a_i v_i - \sum_j b_j h_j \tag{1}$$

In the formula: $\theta = (W, a, b)$, 3 RBM is a very important parameter; visible layer junction i And hidden layer nodes j The weight value of the connection between w_{ij} Visible layer i The offset value of each node a_i, the hidden layer j The offset value of each node is b_j; visible layer i The status value of each node is v_i, And hidden layer j The status value of each node is h_j.

As shown in Fig. 1, v Indicates the visible layer, h Indicates a hidden layer, W Represents the connection weight between two layers. Among them, the visible layer and the hidden layer, the inter-layer neurons are fully connected, and the intra-layer neurons are not connected.

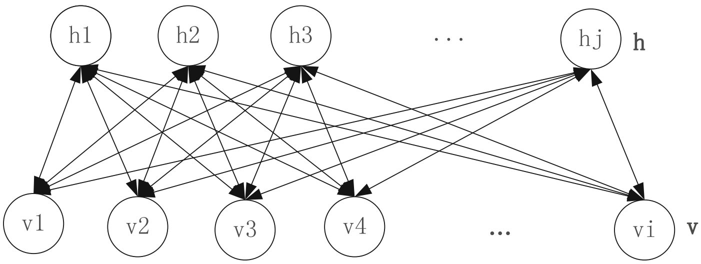

Fig. 1. Schematic diagram of the restricted Boltzmann machine

As a DL model, DBNs have been successfully applied in many fields such as object recognition and speech recognition. For remote sensing images, the pixel layer corresponds to the visible layer, and the feature description factor corresponds to the hidden layer as show in Fig. 2.

The BP algorithm is the most classical algorithm for training neural networks. The Convolutional Neural Network (CNN) used in deep learning algorithms is also trained by similar algorithms. The training methods used by DBNs are quite different from those of traditional neural networks: BP network algorithm for multiple hidden layer networks, the first is that the training time is too long; secondly, the weight adjustment process is from the output layer, the input layer is reversed. To the transmission, when the residual propagates to the first layer, there are many errors, so the weight adjustment is not accurate enough, the algorithm is not efficient, and the training method is not ideal. There are several main problems with the BP training method:

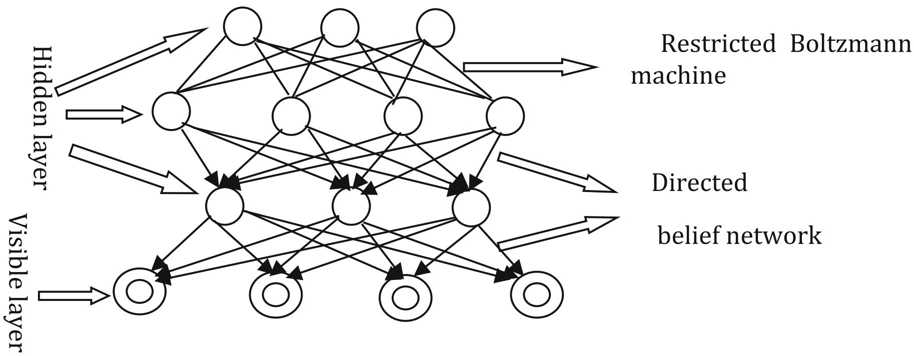

Fig. 2. Deep belief network (DBNs) structure model

(1) BP is the process from the output layer to the input layer. After several hidden layers, the gradient becomes more and more sparse, the error correction signal becomes smaller and smaller, and the weight change is very small.
(2) The weight of the neural network is randomly assigned to the initial value of the backpropagation algorithm. This algorithm may cause the algorithm to converge locally to a minimum.
(3) The marked data can be used for BP algorithm training, because the weight adjustment in the subsequent propagation process must adjust the expected value through the tag value and error, which requires the back-propagation of the neural network in the sample, the actual data is not labeled.

Due to many shortcomings of the traditional neural network training model, DL uses a new training method, which is to establish a multi-layer neural network in unsupervised data. The training time is divided into two steps: the first step is divided into training networks, each training a level. The second step is the parameter adjustment process, which is to monitor the fine adjustment.

The training process of DBN adopts layer-by-layer training. The training only has one layer of RBM at a time. The process of RBM is exactly the same as this training, and the parameters are adjusted separately. After one layer of training, the result is input as the next layer of RBM; so until each layer RBM is trained and this process is called pre-training. After the RBM training is completed, the BP algorithm is used to fine tune according to the tag value of the sample.

3 Remote Sensing Image Recognition Using Deep Belief Network

3.1 Experimental Object and Pretreatment

The experiment selected in this paper is based on the MNIST dataset and the NWPU-RESISC45 dataset is selected [4]. Some remote sensing image is made into subjects. The NWPU-RESISC45 dataset is a publicly available benchmark for Resensing Image Scene Classification (RESISC) created by Northwestern Polytechnical University (NWPU).

In order to get the closest to the standard experiment, the color image was processed, and a batch of remote sensing images similar to figure a were preprocessed and turned into gray image b, and the experimental training set and test set were constructed, As shown in Fig. 3:

a original picture b gray image

Fig. 3. Preprocessing of remote sensing maps

This experiment still uses a double-layer DBNs structure, similar to the MNIST data set, and the two hidden layer units are set to 1000 and 200 respectively, because this paper selects five categories of beaches, clouds, deserts, islands, and lakes. Remote sensing image, so the output layer is 5 units. In the DBNs training phase, set the number of iterations of the two layers of RBM to 200, And the learning rate is set to 0.1. After completing the training of the DBNs, the NN is initialized with the weights learned by the system, and the network parameters are fine-tuned; the iteration number of the NN is 100, the learning rate is 0.1, and the NN partial activation function is set to "Sigmoid". Draw the weight map according to the learned weight as shown in Fig. 4.

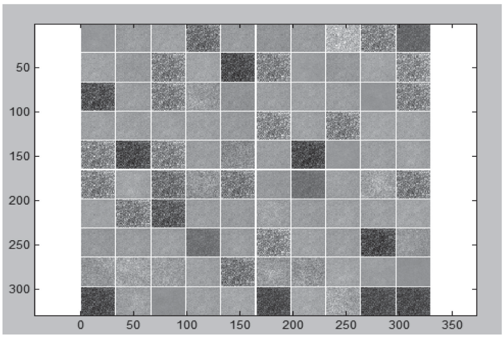

Fig. 4. DBN weight map

The training set and test set are shown in Table 1:

Table 1. Test training data set

Serial number	Total data volume	Training set	Test set
1	70000	60000	10000
2	700	600	100
3	7000	6000	1000

3.2 Analysis of Experiments

This section focuses on comparing the accuracy and rate of DBNs and CNN training the same object. Other parameters of the DBNs model were initialized in this paper as follows: the learning rate of pre-training and fine-tuning was set to 0.05, and the mini-batch size was set to 100.

When training models with the MNIST dataset, compare DBNs with CNN. The influence of iteration times on the recognition and classification accuracy of CNN is shown in Fig. 5.

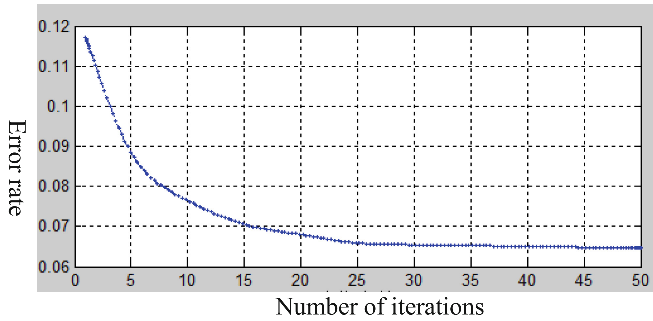

Fig. 5. CNN change with changes in the number of iterations to identify the classification accuracy curve

It can be seen from Fig. 5 that when the number of iterations reaches 50, the recognition error rate of CNN reaches a minimum value. At this time, the recognition classification accuracy in the network is the largest, and the network classification effect is the best. However, when the number of iterations reaches 25, the network has reached a relatively good classification recognition effect. After the number of iterations continues to increase, the curve still has a downward trend, which means that the classification accuracy will still be improved. Improve, but not much improvement. Therefore, in some cases, the number of iterations can be reasonably selected within a certain range

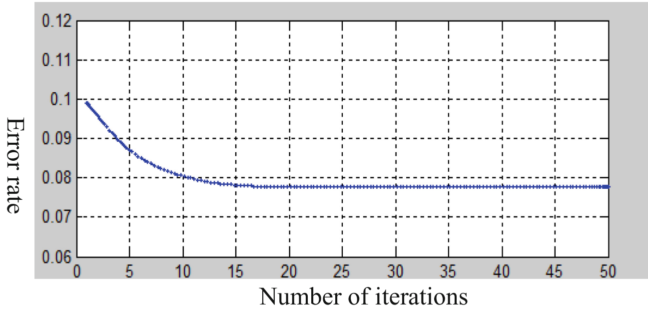

Fig. 6. DBN identifies the accuracy curve of the classification as the number of iterations changes

of variation according to the specific requirements of training time and recognition accuracy.

It can be clearly seen from the trend of the curve in Fig. 6 that when the number of iterations reaches 15, even if the number of iterations continues to increase, the classification accuracy of the DBN network does not increase, and basically goes to a stable value. It is 0.078, that is, the accuracy is as high as 92.2%. This is because the weight of the DBN backward trimming stage is pre-adjusted in the process of forward propagation, avoiding random initialization, and adopting this divergence algorithm to train each layer of RBM, so after dozens of times After the number of iterations, the network can get a good result.

Because DBN is a layered training, and CNN is a training method from the next iteration; A training process of CNN requires alternating convolution and sampling, both of which are time-consuming. Therefore, the forward propagation velocity is relatively low. The backward propagation of CNN includes two processes, deconvolution and ascending sampling, both of which are rather complicated. With multiple levels of CNN, the time of gradient descent is naturally longer. In contrast, DBNs have obvious advantages. As shown in Table 2.

Table 2. DBN and CNN time-consuming comparison

Network model	Total time (seconds)	Error rate
DBN	0.15	0.08
CNN	1.79	0.07

In this experiment, it can be clearly seen that when the DBNs model and the CNN model are training the same object, the training time is shorter, that is, the DBNs have a higher rate in remote sensing image recognition.

4 Conclusion

In this paper, the DBNs model is deeply studied, its principle, method and model structure are analyzed, and the idea structure of RBM is analyzed in detail. The mechanism of DBNs hierarchical training is discussed and compared with CNNS, which greatly reduces the difficulty and training time of model training. The experiment shows that compared with CNNS, DBNs model is basically equal in recognition rate, but has obvious advantage in training rate. The recognition accuracy and speed of the deep belief network model for remote sensing image are better than the traditional neural network learning method. It is also an attempt and improvement to select a new and superior remote sensing scene image data set as the research object of DBNs model in extracting image features.

Acknowledgment. This work was supported by the Education and teaching research project of Jingchu university of technology (No. JX2019-032).

References

1. Deng, J., Zhao, L.C.: Classification of remote sensing images based on Fisher and neural network fusion. Inner Mongolia Norm. Univ. News (Chin. Version Nat. Sci.) **45**(1), 46–49 (2016)
2. Li, X., Zhang, H.: Identification of remote sensing image of adverse geological body based on classification. In: Gong, M., Pan, L., Song, T., Tang, K., Zhang, X. (eds.) BIC-TA 2015. CCIS, vol. 562, pp. 232–241. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-49014-3_21
3. Mantero, P., Moser, G., Serpico, S.B.: Partially supervised classification of remote sensing images through SVM-based probability density estimation. IEEE Trans. Geosci. Remote Sens. **43**(3), 559–570 (2005)
4. Cheng, G., Han, J., Lu, X.: Remote Sensing Image Scene Classification. Benchmark and State of the Art. Proc. IEEE **105**(10), 1865–1883 (2017)
5. Qi, L., Yong, D., Xin, N.: Remote sensing image classification based on DBN model. Comput. Res. Dev. **51**(9), 1911–1918 (2014)
6. Hagner, O., Reese, H.: A method for calibrated maximum likelihood classification of forest types. Remote Sens. Environ. **110**(4), 438–444 (2007)
7. Alberga, V.: A study of land cover classification using polarimetric SAR parameters. Int. J. Remote Sens. **28**(17), 3851–3870 (2007)
8. Kban, K.U., Yang, J., Zhang, W.: Unsupervised classification of polarimetric SAR images by EM algorithm. IEICE Trans. Commun. **E90-B**(12), 3632–3642 (2007)
9. Pal, M., Mather, P.M.: An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sens. Environ. **86**(4), 1145–1161 (2003)
10. Heermann, P., Khazenic, N.: Classification of multispectral remote sensing data using a back-propagation neural network. IEEE Trans. Geosci. Remote Sens. **30**(1), 81–88 (1992)
11. Lardeux, C., Frison, P., Tison, C., et al.: Support vector machine for multifrequency SAR polarimetric data classification. IEEE Trans. Geosci. Remote Sens. **47**(12), 4143–4152 (2009)
12. Nin, X., Ban, Y.F.: Multi-temporal RADARSAT-2 polarimetric SAR data for urban land-cover classification using an object-based support vector machine and a rule based approach. Int. J. Remote Sens. **34**(1), 1–26 (2013)
13. Niu, X., Ban, Y.F.: A novel contextual classification algorithm for multitmporal polarimetric SAR data. IEEE Geosci. Remote Sens. Lett. **11**(3), 681–685 (2014)