ISSN 2616-5775 Vol. 7, Issue 2: 79-84, DOI: 10.25236/AJCIS.2024.070211

Application of Convolutional Neural Networks in High Score Remote Sensing Image Classification

Jiawei Zhang[a,*], Xuwei Zhang[b], Xu Ji[c]

School of Geomatics and Urban Spatial Information, Beijing University of Civil Engineering and Architecture, Beijing, 102616, China
[a]2108570021102@stu.bucea.edu.cn, [b]1225595395@qq.com, [c]1525822059@qq.com
[*]Corresponding author

Abstract: *Remote sensing image classification is a critical link in remote sensing. Traditional remote sensing image classification is based on shallow structure model algorithms such as SVM and decision tree. However, when faced with high-resolution remote sensing images, due to a large amount of data and complex data features, the recognition accuracy of traditional shallow models has been unable to meet the current needs. When faced with image classification, convolutional neural networks can better cope with image translation, scaling, and other deformations and effectively reduce the errors introduced by the above factors. At the same time, the convolutional neural network can capture and extract complex signal features with the help of a deeper feature processing level to obtain a better classification effect. This paper discusses the application of convolutional neural networks to classify high-resolution remote sensing images and provides some references for remote sensing image processing and target recognition.*

Keywords: *Deep learning; High-scoring remote sensing image classification; Convolutional neural network; Target recognition*

1. Introduction

A remote sensing image is an image that forms the characteristics of ground objects by recording electromagnetic waves of various ground objects. In remote sensing images, the contours and areas of rocks, soil, water, vegetation, buildings, etc. are described by the level of brightness values and the size of pixel values, and it is precisely based on these differences that we can distinguish the ground objects one by one. In the processing and application of remote sensing images, the most important thing is the classification of remote sensing images. The classification of remote sensing images is essentially the use of computers to distinguish various objects in the obtained remote sensing images. The distinction is usually based on recorded spectral information and geospatial information. According to the above basic classification basis, some special methods can be used to construct independent feature spaces, and the pixels in remote sensing images are allocated to each independent feature space to achieve classification [1]. The current mainstream remote sensing impact classification methods are: supervised, unsupervised, and semi-supervised. The current mainstream classification algorithms are support vector machine-based classification algorithms, decision tree-based classification algorithms, random forest-based classification algorithms, and semantic modeling-based classification algorithms. However, in the face of high-resolution remote sensing images, the above single or combined methods cannot obtain better classification results [1].

In this context, the artificial neural network has entered the attention of researchers in remote sensing image classification. The artificial neural network is a typical machine learning model, which is very good at dealing with nonlinear classification problems, and also has strong self-learning and self-organization, which enables it to obtain better performance when dealing with some complex problems. The traditional classification method represented by SVM belongs to the shallow structure, and its typical characteristic is that between the signal input and the signal output, it experiences less linear or nonlinear processing levels (the lowest can be one layer). When dealing with some complex feature extraction and classification tasks, the shallow structure model is difficult to capture the complex features in the signal, so it usually shows obvious limitations [2]. As the amount of data to be learned and classified becomes larger and larger, the traditional shallow structure model is no longer applicable. The artificial neural network, a deep structure model, can better cope with the nonlinear learning of complex features.

When faced with classification problems with complex signals such as high-scoring remote sensing image classification, deep-structure models can achieve higher classification performance. The convolutional neural network is a typical artificial neural network with a multi-layer perceptron for two-dimensional shape perception. From the classification perspective, it can also be regarded as a class with typical structural features in the BP neural network. Thanks to its characteristics of local perception, shared weights, and downsampling in space and time, convolutional neural networks have high uniformity for different deformations such as translation and scaling of images [3]. This paper mainly studies the application of convolutional neural networks in high-resolution remote sensing image classification and hopes to provide some references for remote sensing image processing.

2. Deep Learning and Common Models

Deep learning is a high-level machine learning method derived from the development of machine learning to a certain extent. The idea of deep learning comes from the characteristics of the human brain. When the human brain is in the process of cognition of the external world, it processes and extracts features through a progressively deepened hierarchical model. There is essentially a semantic mapping relationship from low to high between layers. Geoffrey Hinton, who first proposed deep learning as a brain-machine learning method, has the title of "Father of Deep Learning". After years of demonstration and experiments, he proved the feasibility and value of deep learning and thus opened academic and commercial applications of deep learning. So far, deep learning has shown its value in various fields, and it has been proved to be an inevitable trend in the era of big data and artificial intelligence. At present, many high-tech companies worldwide have successively invested in the research and application of deep learning technology, such as Baidu, Google, Microsoft, iFLYTEK, and SenseTime. Microsoft Research Asia released a deep learning-based simultaneous interpretation system in 2012. In 2013, Baidu established the first deep learning research institute in China that was led and operated by an enterprise. In the following year, Google officially acquired DeepMind and entered the field of deep learning in an all-around way. In addition, iFLYTEK in my country has been committed to the application of deep learning in the fields of language translation and semantic recognition. At the same time, SenseTime is specially oriented to images and big urban data and establishes business models based on deep learning to provide business solutions for various fields [3].

The result of deep learning is to obtain a deep network structure consisting of many neurons, and there are also complex associations with different weights between neurons. During the learning process, the connections between the network nodes may appear or disappear, or the weights may change, thereby realizing different network functions. This deep network structure has the characteristics of a neural network, so it can also be called a deep neural network. Common deep learning models include convolutional neural networks, sparse coding, restricted Boltzmann machines, etc.

3. Algorithm Principle of Convolutional Neural Network

As mentioned above, the convolutional neural network is a typical type of deep learning model, and its structure is highly similar to that of human brain neurons. The convolutional neural network is the first network model that combines neural network technology and deep learning to build a multi-layer network structure and complete machine learning tasks. The convolutional neural network applies the idea of gradient backpropagation to continuously optimize the weight of the network and constructs a global learning algorithm with a deep combination of multi-layer filters and classifiers. This method effectively reduces the complexity of the network structure and can deal with problems in machine vision and speech analysis.

The convolutional neural network effectively reduces the learning parameters in the network, which positively affects the training performance of the BP algorithm. The convolutional neural network captures the data features through the feature monitoring layer, which implicitly learns the features, avoiding the complexity and tendency of explicit extraction. At the same time, in the convolutional neural network, all neurons on the feature mapping surface have the same weight, which allows the learning process to be performed in parallel, which greatly improves efficiency. Because of this feature, convolutional neural networks have obvious advantages compared with other networks. The convolutional neural network has a good performance in processing image feature extraction and classification tasks. The main reason is that a part of the image is used as the input data of the outermost layer, which is input into the network, and then passes through the layers between the layers. The digital filter observes and extracts the core features and finally forms the weights between the layers, which can

ISSN 2616-5775 Vol. 7, Issue 2: 79-84, DOI: 10.25236/AJCIS.2024.070211

capture the essential features of the image, such as directional edges, bump vertices, and other key features. Based on this feature, the convolutional neural network can cope well with image stretching, rotation, scaling, and other deformations. This characteristic is more pronounced when the input data is a multi-bit image. At the same time, because of their processing characteristics, images can be directly used as input data for training or classification, and their application is more convenient and efficient [4]. The basic structure of a convolutional neural network is shown in Figure 1.

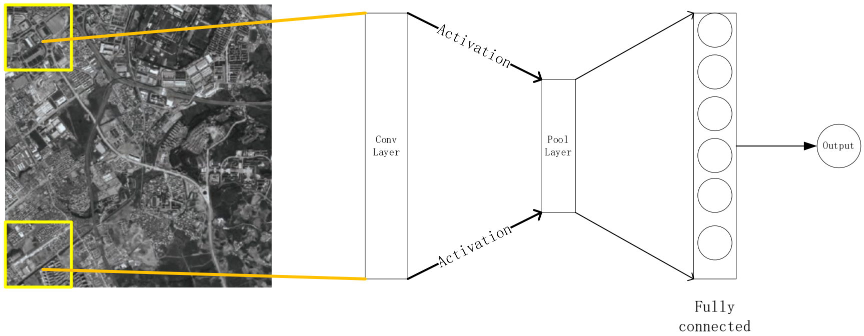

Figure 1: Convolutional neural network

The research and selection of network structure is the core of convolutional neural network research. The convolutional neural network structure is very complex, which makes it time-consuming in the training process. Therefore, in the face of a certain problem or data set, how to design an appropriate network structure to obtain better training and prediction performance is the basic problem to be solved by researchers at present. On this basis, whether it is possible to find a neural network structure with wider adaptability is an advanced research hotspot. For example, is it possible to build a neural network model that can evolve autonomously so that the initially simple neural network can increase layers and optimize weights autonomously with the continuous advancement of the learning process. Until it reaches the training goal and stops actively, and obtains a better neural network performance of neural network models.

4. Application of High-Resolution Remote Sensing Image Classification

4.1 Mainstream Related Research

Fukushima originally proposed the convolutional neural network in Japan in the early 1980s. Based on Fukushima's research, Yann LeCun and other scholars have further improved it and further expanded the former's ideas, clarifying the volume of convolutional neural networks. The application method of the hidden layer of the product neural network. Since then, convolutional neural networks have been further generalized and simplified based on Seven Behnke and Patrice Simard et al. By 2012, convolutional neural networks were well optimized and achieved intentional performance on well-known image libraries MNIST, HWDB, etc. So far, convolutional neural networks have been deeply applied in speech and image fields and have made great contributions to specific work fields, including machine vision, remote sensing image processing, etc. [5].

4.2 Application Advantages of Convolutional Neural Network over Ordinary Fully Connected Network

For convolutional neural networks, the advantage lies in recognizing audio and images. The neural network can obtain mostly complex classification surfaces based on its multi-layer backpropagation training mode. In the traditional pattern recognition process, a feature filter specially designed for a certain feature is often used for feature extraction, and a feature variable is obtained by eliminating irrelevant parameters. Then, by inputting these feature variables into the trainable classifier, the classifier gradually has the ability to recognize such features. In this process, a multi-layer, fully connected network can be used as a classifier for training [7-10]. At this time, if the hand-designed feature filter is removed, the image is directly input into the network as data, and the previous layers are trained through the backpropagation algorithm to form a feature filter, and the same classification effect can be obtained obviously,. The structure of the fully connected neural network can be seen in Figure 2.

ISSN 2616-5775 Vol. 7, Issue 2: 79-84, DOI: 10.25236/AJCIS.2024.070211

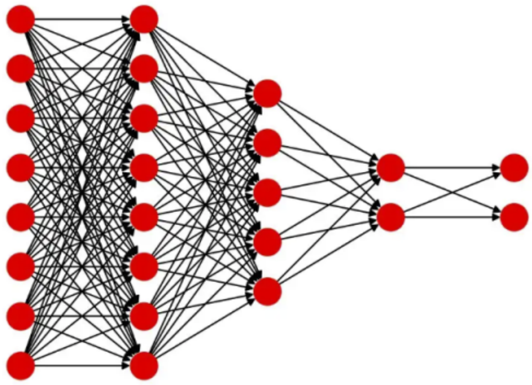

Figure 2: Fully connected network

However, there are typical problems in practical applications of fully connected networks:

(1) In the face of image or audio data, the amount of data is usually large, and it contains a larger group of variables. This makes the fully connected network face a large number of neurons. If there are not enough data samples, the weights between neurons will be too scattered, leading to the overfitting problem [11-15].

(2) Ordinary fully connected networks do not distinguish between data categories. For example, to identify two highly similar images, the difference between the two is only a slight displacement, and the rest are exactly the same. At this point, the fully connected network cannot perceive this subtle difference and cannot optimize for these features during training.

(3) The fully connected network does not pay attention to the locality of the training data.

In the face of audio and video data with a typical one-dimensional or two-dimensional structure, the convolutional neural network can perform local perception and combine weight sharing and downsampling from the spatio-temporal angle to realize the extraction of local features. Displacement, deformation, and other situations can be better dealt with [6].

4.3 High-resolution Remote Sensing Image Classification Training

The high-resolution remote sensing image classification problem is a complex feature extraction problem. This paper chooses to apply a convolutional neural network to solve this problem. Convolutional Neural Networks have three typical layers:

(1) Convolutional Layer

Each neuron in the convolutional layer is associated with a local area of the previous layer, which can be regarded as a feature observation and extractor for this area. Each neuron can capture a certain feature of the region, such as directed edges, vertices, etc. At the same time, this local connection method of the convolutional layer results in fewer parameters, and the training performance is significantly improved [5]. The mathematical principle of the convolutional layer can be expressed as Equation 1. Where l represents the location of the convolutional layer, k represents the convolution kernel, and b is the bias. f is the selected activation function. Mj is the input data from the previous layer.

$$x_j^l = f(\sum_{i\in M_j} x_i^{l-1} \cdot k_{ij}^l + b_j^l) \tag{1}$$

(2) Sampling Layer

A sampling layer follows the convolutional layer. The sampling layer downsamples the image data output by the convolutional layer to further reduce its resolution. This layer can be represented as Equation 2 [16-20]. Where down() is the sampling function used, β is the sampling coefficient of the sampling function, and b and f still represent the bias and activation functions, respectively.

$$x_j^l = f(\beta_j^l down(x_j^{l-1}) + b_j^l) \tag{2}$$

(3) Fully Connected Layer

The fully connected layer mainly classifies the sampled features and completes this learning through the hidden layer and the output layer to realize forward propagation. The result is output externally through the output layer.

There are two main stages of training for the high-resolution remote sensing image classification problem studied this time: forward propagation and backward propagation.

(1) Forward Propagation

This training phase mainly extracts a sample from the sample set, sends it to the network, passes through a series of pre-designed convolutional layers, sampling layers, and fully connected layers, and finally obtains the actual output. The process can be expressed as Equation 3.

$$O_P = F_n(\cdots(F_2(F_1(XW_1)W_2)\cdots)W_n) \tag{3}$$

(2) Backward Propagation

The error is mainly back-propagated to correct the previous training results in this stage. Through forward propagation, an actual output $\mathrm{O_P}$ can finally be obtained. By comparing with the ideal control data $\mathrm{y_p}$, the difference between the two can be found, and by minimizing the error, the network weights of the previous sequence are adjusted accordingly to achieve Reverse active tuning of the network. Its process can be expressed as Equation 4.

$$E_P = \frac{1}{2}\sum_j(y_{pj} - O_{pj})^2 \tag{4}$$

In essence, the application problem of high-resolution remote sensing image classification mainly relies on two basic assumptions. First, ensure that the input of neurons is limited and as few as possible so that the gradient can be more step-by-step propagation in multiple layers. Second, hierarchical local connections have good priors. That is, they can be confirmed without a posteriori. This structure is very friendly to visual recognition [6]. If the entire network happens to be under ideal parameter conditions, its gradient optimization can finally achieve the desired effect.

5. Conclusion

This paper gives an overview of the remote sensing image classification problem and studies its work goals and traditional methods. Afterward, the limitations of traditional classification algorithms were discussed, and convolutional neural networks were introduced to deal with more complex, high-scoring remote sensing image classification problems. High-scoring remote sensing images have more parameters, more complex data features, and poor classification accuracy under traditional methods such as SVM. This paper discusses the application feasibility of a convolutional neural network in high-resolution remote sensing image classification and discusses the corresponding network construction and training. It is hoped that the research in this paper can be helpful for remote sensing image processing.

References

[1] Song Yixuan, Song Fei, Jin Lei, Lei Tao, Liu Gang, Jiang Ping, Peng Zhenming. An attention cut classification network for fine-grained ship classification in remote sensing images[J]. Remote Sensing Letters, 2022, 13(4).

[2] Li Yansheng, Dang Bo, Zhang Yongjun, Du Zhenhong. Water body classification from high-resolution optical remote sensing imagery: Achievements and perspectives[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 187.

[3] Zhao Chuanpeng, Qin Cheng-Zhi. Identifying large-area mangrove distribution based on remote sensing: A binary classification approach considering subclasses of non-mangroves[J]. International Journal of Applied Earth Observation and Geoinformation, 2022, 108.

[4] Devi, N. Bharatha, Kavida, A. Celine, Murugan, R. Feature Extraction and Object Detection Using Fast-Convolutional Neural Network for Remote Sensing Satellite Image[J]. Journal of the Indian Society of Remote Sensing, 2022(prepublish).

[5] Uss Mykhail, Vozel Benoit, Lukin Vladimir, Chehdi Kacem. Exhaustive Search of Correspondences between Multimodal Remote Sensing Images Using Convolutional Neural Network[J]. Sensors, 2022, 22(3).

[6] Shi Cuiping, Zhang Xinlei, Sun Jingwei, Wang Liguo. Remote Sensing Scene Image Classification Based on Self-Compensating Convolution Neural Network[J]. Remote Sensing, 2022, 14(3).

[7] Thepade Sudeep D., Dindorkar Mayuresh R. Fusing deep convolutional neural network features with Thepade's SBTC for land usage identification[J]. Engineering Science and Technology, an International Journal, 2022, 27.

[8] Zhang W, Tang P, Zhao L. Remote sensing image scene classification using CNN-CapsNet[J]. Remote

ISSN 2616-5775 Vol. 7, Issue 2: 79-84, DOI: 10.25236/AJCIS.2024.070211

Sensing, 2019, 11(5): 494.

[9] Soni Ashish, Koner Radhakanta, Villuri Vasant Govind Kumar. Fusion of Dual-Scale Convolution Neural Network for Urban Building Footprints[J]. Ain Shams Engineering Journal, 2022, 13(3).

[10] Alquraish Mohammed M., Khadr Mosaad. Remote-Sensing-Based Streamflow Forecasting Using Artificial Neural Network and Support Vector Machine Models[J]. Remote Sensing, 2021, 13(20).

[11] Bai Bing, Zhao Hongmei, Zhang Sumei, Zhang Xuelei, Du Yabin. Can Neural Networks Forecast Open Field Burning of Crop Residue in Regions with Anthropogenic Management and Control? A Case Study in Northeastern China[J]. Remote Sensing, 2021, 13(19).

[12] Zhang Sihang, Shao Zhenfeng, Huang Xiao, Bai Linze, Wang Jiaming. An internal-external optimized convolutional neural network for arbitrary orientated object detection from optical remote sensing images [J]. Geo-spatial Information Science, 2021, 24(4).

[13] Xie Guanyao, Niculescu Simona. Mapping and Monitoring of Land Cover/Land Use (LCLU) Changes in the Crozon Peninsula (Brittany, France) from 2007 to 2018 by Machine Learning Algorithms (Support Vector Machine, Random Forest, and Convolutional Neural Network) and by Post-classification Comparison (PCC)[J]. Remote Sensing, 2021, 13(19).

[14] Patel, Heena, Upla, Kishor P. A shallow network for hyperspectral image classification using an autoencoder with convolutional neural network[J]. Multimedia Tools and Applications, 2021, 81(1).

[15] Chen Haiwen, Wang Lu, Zhang Lei, Li Yanping, Wu Yuting, Qi Jun. Research on Remote Sensing Image Target Detection Methods Based on Convolutional Neural Networks[J]. Journal of Physics: Conference Series, 2021, 2025(1).

[16] Siesto Guillermo, FernándezSellers Marcos, LozanoTello Adolfo. Crop Classification of Satellite Imagery Using Synthetic Multitemporal and Multispectral Images in Convolutional Neural Networks[J]. Remote Sensing, 2021, 13(17).

[17] Yannic Lops, Arman Pouyaei, Yunsoo Choi, Jia Jung, Ahmed Khan Salman, Alqamah Sayeed. Application of a Partial Convolutional Neural Network for Estimating Geostationary Aerosol Optical Depth Data [J]. Geophysical Research Letters, 2021, 48(15).

[18] Lang Lei, Xu Ke, Zhang Qian, Wang Dong. Fast and Accurate Object Detection in Remote Sensing Images Based on Lightweight Deep Neural Network[J]. Sensors, 2021, 21(16).

[19] Kareem Razia Sulthana Abdul, Ramanjineyulu Anil Gandhudi, Rajan Regin, Setiawan Roy, Sharma Dilip Kumar, Gupta Mukesh Kumar, Joshi Hitesh, Kumar Ankit, Harikrishnan Haritha, Sengan Sudhakar. Multilabel land cover aerial image classification using convolutional neural networks[J]. Arabian Journal of Geosciences, 2021, 14(17).

[20] Aires Filipe, Boucher Élodie, Pellet Victor. Convolutional neural networks for satellite remote sensing at coarse resolution. Application for the SST retrieval using IASI[J]. Remote Sensing of Environment, 2021, 263.