Research on High Resolution Satellite Image Classification Algorithm Based on Convolution Neural Network

Gaiping He

Abstract

Artificial neural network is a kind of artificial intelligence method to simulate the function of human brain, and deep learning technology can establish a depth network model with hierarchical structure on the basis of artificial neural network. Deep learning brings new development direction to artificial neural network. Convolution neural network is a new artificial neural network method, which combines artificial neural network and deep learning technology, and this new neural network is widely used in many fields of computer vision. Modern image recognition algorithm requires classification system to adapt to different types of tasks, and deep network and convolution neural network is a hot research topic in neural networks. According to the characteristics of satellite digital image, we use the convolution neural network to classify the image, which combines texture features with spectral features. The experimental results show that the convolution neural network algorithm can effectively classify the image.

Keywords: High Resolution Satellite, Image Classification, Convolution Neural Network, Clustering Algorithm.

Introduction

The basis of image recognition technology is analyzing the main features of images. Complex image recognition requires the integration of image information in multiple levels, in order to integrate the scattered image feature points. Similarly, how to extract and classify feature points in the process of image processing is a hot spot in the research of image recognition in recent years.

Thematic information obtained from satellite images, the traditional method is based on visual interpretation. This approach requires researchers to have a wealth of knowledge and experience, and they need to spend a lot of time to observe the satellite. So the labor intensity of scientific research personnel is great, and the information acquisition cycle is too long.

In order to solve these problems since 1970, people use the computer to get the information of remote sensing image. At that time, people mainly use statistical pattern recognition method for image analysis, such as maximum likelihood method, spectral feature, and so on. Although these methods have played a promoting role in the development of computer classification of remote sensing images, the texture information of target object is ignored result in some features cannot be distinguished in the image classification. With the development of satellite technology, satellite sensors have been able to obtain high-resolution digital satellite images. In these high resolution images, the details of the target feature is clear, showing a lot of texture and structure information, which can be used as an important basis for classification and identification of features. Therefore, it is an important way to improve the classification

1Xi'an International University, Xi'an, Shaanxi, 710077 China
accuracy of remote sensing image by combining the feature of spectral features with texture features. However, the difficulty of the problem is how to use mathematical methods to extract and describe the texture features of the object, and combine these features with spectral features for remote sensing image classification.

In this paper, we use convolution neural networks to analyze the classification method of high resolution satellite images. Compared with tradition deep neural network, convolution neural network can simulate the human visual nervous system, and through the local impression to form the final cognitive results. From the angle of the algorithm itself, CNN uses shared parameters between convolution layers, which not only reduces the required memory size, but also reduces the number of parameters and the need of training, improving the performance of the algorithm. Another machine learning algorithm does not have this advantage.

The Proposed Methodology

**Convolution Neural Network.** Convolution neural network is a commonly used algorithm in depth learning, which is inspired by the human visual system. After many years of development, a multi-layer neural network for image classification is formed. The classic convolution neural network consists of four layers:

- Convolution neural layer.
- Rectified Linear Units layer.
- Pooling layer.
- Normalized layer.

![Classical Convolution Neural Network Structure](image)

Figure 1. Typical CNN Architecture.

Figure 2. Classical Convolution Neural Network Structure.
Besides a multilayer feed forward network based on BP network, at present, there are some other types of artificial neural network:

- **Hopfield Network**: This network is a double layer symmetric fully connected loop network structure. But this network is only suitable for the binary input data and does not have the ability to learn.
- **Adaptive Resonance Theory Network**: This is an unsupervised learning network which adopts the environmental information through self-organization.
- **Kohonen Network**: This kind of network is also called self organizing feature map model, which uses unsupervised learning method to cluster image data.
- **Recurrent Networks**: In a cyclic network, the output value is fed back to the same unit or the previous layer.
- **Counter Propagation Network**: Establish a self organizing query table which is used as a function of approximate computation and classification based on blended learning.
- **Radial Basis Function Network**: This is a kind of special three-layer feed forward error back propagation network, which can be used to train a lot of learning algorithms, including step by step hybrid learning algorithm.

**Image Block Similarity Computation.** Firstly, feature extraction is extracted for each image block, then we determine whether the two image blocks are corresponding according to the Euclidean distance between the features. The characteristic expression of the traditional thinking has a serious impact on the accuracy of the final similarity calculation. In the image super resolution task, because the low resolution image blocks are different from the high resolution image blocks, the image features cannot be extracted directly. However, there are many kinds of choices to calculate the similarity of image blocks using depth learning tools.

- **Pair-Wise Model**: This model is first applied into field of image retrieval, and the image features extracted from this model have better classification ability. In this model, the contrastive-loss function can be computed as follows:

  \[
  E = \frac{1}{2N} \sum_{n=1}^{N} y_n d_n^2 + (1 - y_n) \max(m - d_n, 0)^2 , \quad \text{and} \quad d_n = \| a_n - b_n \|_2
  \]

  where \( y_n \) is the \( n \)-th input training sample's label, if \( y_n = 1 \), then two kind of image blocks are corresponding to each other.

![Figure 3. The Architecture of Pair-Wise Model.](image-url)
Two-Route Model. Channel model is to directly convert image block similarity into two classification problems. The two channels in the input data are respectively stored in the low resolution image block and the high resolution image block, and data are entered into the network for processing at the same time. Compared with Pair-Wise model, two-route model does not extract the feature points of each image separately. Two-route model determine all information offered from two image blocks from the beginning.

For the two-route model, we use sparse representation algorithm to generate a dictionary. High resolution image blocks in a training sample is processed on the basis of mean normalization, next, we use a dictionary learning algorithm to produce a high resolution image block dictionary:

\[
(D_{ht}, \partial) = \arg\min_{Y_h} \| Y_h - D_{ht} \partial \|^2, \text{ s.t. } \| \partial \|_0 \leq L
\]

Where \( D_{ht} \) is high resolution image block dictionary, \( \partial \) is sparse representation, and \( L \) is the maximum sparsity constraint. The learning process of the high resolution image block dictionary and the low resolution image block dictionary is the key point of the whole algorithm.

Convergence Condition. There are several factors that can be used as a criterion for the end of training.

- Training error: \( \rho \leq \varepsilon \).
- Training gradient: \( \nabla \beta \leq \delta \).
- Cross validation.

Where \( \rho \) is an arbitrary error function, \( \varepsilon \) and \( \delta \) are two small real number. In general, the error of the training set will be gradually reduced along with the network training.
Conclusion

Convolution neural network introduced in this paper is a new artificial neural network method, which combines artificial neural network and deep learning technology, and this new neural network is widely used in many fields of computer vision. In the face of high resolution satellite images, the traditional image classification method has been unable to accurately extract the image feature points. Nevertheless, CNN, which is inspired by the human visual system, can simulate the human visual nervous system, and form the final cognitive results through the local impression. From the angle of the algorithm itself, CNN uses shared parameters between convolution layers, which not only reduces the required memory size, but also reduces the number of parameters and the need of training, improving the performance of the algorithm. We fully analyze the characteristics and advantages of CNN, and two different models are proposed in this paper: Pair-wise model and Two-route model which introduce a deep learning tool for high and low resolution image blocks' similarity learning. A large number of experimental results and literature show that, the methods introduced in this paper can accurately extract the features of high resolution image elements, and classify the images effectively.

References