Research on the Application of Computer CNN in Image Recognition

Qiaoye Chai1,*

1Department of Information Technology, Shanxi Finance Vocational College, Taiyuan, Shanxi, China

*Corresponding author e-mail: sxjycqy@sxjrzyxy.com

Abstract. With the development of AI, deep learning (hereinafter referred to as DL) has been applied to many aspects, especially the application of CNN in image recognition (hereinafter referred to as IR). CNN is an important branch of DL, which has been deeply studied in IR. Through CNN, we can choose the right image in the image set, which can also be classified. With the different purposes of IR, the academia has developed a variety of different CNN models, which will better promote the application in IR. By increasing the training data, we can improve the recognition rate and accuracy of CNN, which will improve the efficiency of IR. IR based on CNN has broken the bottleneck of traditional target detection (hereinafter referred to as TD), which has become a mainstream algorithm in the field of TD. Firstly, this paper analyzes the basic theory and related models of CNN. Finally, this paper analyzes the application of CNN in IR.

Keywords: CNN, Image Recognition, Classification

1. Introduction

IR is an important field of AI, which is also an important application field of CNN. DL has been applied to many IR fields, such as face recognition, medical IR, remote sensing image classification and so on, which can quickly recognize the features of small sample data [1]. Through unsupervised and semi supervised learning, CNN can recognize the data in images and videos. Through CNN, we can better study the TD, which can more accurately locate the target of interest [2]. By judging the specific category of each target, the algorithm can give the bounding box, which will lock the relevant features. Therefore, CNN is also used in TD, including intelligent video surveillance, automatic vehicle driving, robot environment perception and so on [3-5].

2. The nature of CNN in IR

2.1. Connectivity
In CNN IR, the connection between convolution layers is called sparse connection. Compared with the full connection in feedforward neural network (hereinafter referred to as NN), the neurons in convolution layer are only connected, which is not all neurons. The sparse connection of CNN IR has regularization effect, which improves the stability of network structure (hereinafter referred to as NS) [6-8]. Meanwhile, CNN avoids over fitting, which reduces the total weight parameters. CNN is conducive to the fast learning of NN, which can reduce the memory overhead. All pixels in the same channel share a set of convolution kernel weight coefficients, which is called weight sharing. Weight sharing distinguishes CNNs from other NNs with local connections. Although the latter uses sparse connections, the weights of different connections are different. Weight sharing is the same as sparse connection, which reduces the total parameters of CNN IR. From the perspective of fully connected network, sparse connection and weight sharing of CNN IR can be regarded as two infinitely strong priors [9-10].

2.2. Biological similarity

In CNN IR, the sparse connection based on receptive field setting has a clear corresponding neuroscience process. Visual cortex cells receive signals from photoreceptors in the retina. But a single visual cortex cell does not receive all the signals from the photoreceptors. Only stimulation in the receptive field can activate the neuron. A large number of visual cortex cells receive signals from the retina and establish visual space by systematically superimposing receptive fields. The nature of weight sharing in CNN IR has no clear evidence in biology, but weight sharing significantly improves the learning effect [11].

3. Related concepts of CNN

3.1. Basic structure

The typical structure of CNN consists of five parts, as shown in Figure 1. CNN is a special deep feedforward network. In the NS, we usually use the network model of combining multi convolution layer and multi pooling layer in the network model, which can make the output more accurate. The scenic spot models in CNN model mainly include LeNet-5, AlexNet, ZF-Net, etc.

![The CNN Architecture](image)

**Figure 1.** The CNN Architecture

3.2. Neocognitron model

Neocognitron model is a bionic structure, which was discovered by Fukushima in the study of cat visual system. Fukushima analyzes the alternating s-cell layer and C-cell layer, which makes up for the concept of Neocognitron hierarchy. In this paper, the s cell layer is simply analyzed, as shown in Figure 2.
3.3. MFR-DenseNet model

MFR-DenseNet is a method to improve the representation ability of DenseNet by adaptively recalibrating between different convolution layer features. Firstly, the model recalibrates the dynamic channel characteristics, which will be gradually introduced into DenseNet. Therefore, the MFR-DenseNet model constructs the channel characteristic CFR-DenseNet. By simulating the interdependencies between the features of different convolution layers, ILFR-DenseNet is constructed. Finally, we can combine CFR-DenseNet and ILFR-DenseNet with the integrated learning method, which is the MFR-DenseNet model, as shown in Figure 3.

3.4. LeNet-5 model

Through the back propagation of training, we can apply CNN to handwritten digit recognition, which is LeNet-1. CNN is simpler than Neocognitron and its extensions. In Neocognitron, layer s alternates with layer C. In CNN, convolution and subsampling layers alternate. In CNN, we use weight sharing, which realizes the concept of receptive domain. The LeNet-5 model is shown in Figure 4.
4. Application of CNN in IR

4.1. The application of image segmentation in TD

R-CNN can use CNN to replace the traditional part of graph feature extraction. Compared with the traditional method, the map of R-CNN has been significantly improved. However, because each image needs 2000 convolution operations, the test speed of the algorithm is too slow, the average processing time of one image is 34 s, and the time cost is large. R-CNN needs different methods, which can be tedious, as shown in Figure 5.

4.2. CNN TD based on Classification

Traditional TD methods include preprocessing, window sliding, and other steps, which have the functions of feature extraction, feature selection and feature classification. We can directly use CNN to binary classify the candidate region generated by each sliding window, which can determine whether it is the target to be detected. This method is called CNN TD based on classification. Compared with traditional TD, CNN TD based on classification has only three steps: window sliding, image classification, post-processing and so on. Window sliding and post-processing are fixed methods. Therefore, we need to improve the feature extraction, selection and classification of CNN, which will improve the accuracy of image recognition.

4.3. TD model based on regression method

Fast R-CNN realizes end-to-end training, which basically realizes real-time detection. At the same time, we need to train two convolutional networks. We adopt a strict "three-step" approach to solve the problem. CNN has its own unique advantages, which only uses one NN to complete the TD.
the Yolo network model, we divide the image into several regions, and each region is only responsible for predicting at most one target, which can be corrected by the border regression method.

![Diagram](image.png)

**Figure 6. Yolo NS**

5. Conclusion

AI has become a reality, which is also the problem we will face. Therefore, we must learn CNN, which can better help us solve the problems in IR. Through CNN, image processing technology will also get comprehensive development, which will promote the rapid development of society.

References

[1] Chen Yaodan, Wang Lianming. Face recognition method based on convolutional NN [J]. Journal of Northeast Normal University, 2016, 48 (2): 70-76.

[2] Li Yandong, Hao Zongbo, Lei hang. Review of convolutional NNs [J]. Computer applications, 2016, 36 (9): 2508-2515.

[3] Liu yuecong, Xu Weisheng, Wu Qidi. Spatiotemporal deep convolution NN and its application in behavior recognition [J]. Computer science, 2015 (7): 245-249.

[4] Lu Hongtao, Zhang Qinchuan, Meng Zihang. Review on the application of deep convolution NN in computer vision [J]. Data acquisition and processing, 2016 (1): 1-17.

[5] Tang Xianlun. IR method based on conditional depth convolution generation countermeasure network [J]. Acta automatica Sinica, 2018, 44 (5): 855-864.

[6] Wang Hongde. Application of convolutional NN IR technology in intrusion detection system [J]. Journal of Dalian Jiaotong University, 2017 (5): 107-110.

[7] Wang Yifei. Application of convolutional NN based on shortcut in face recognition [J]. Journal of electronic measurement and instrument, 2018 (4): 22-23.

[8] Wang Zhen, Gao Maoting. Design and implementation of IR algorithm based on convolutional NN [J]. Modern computer (Professional Edition), 2015 (20): 61-66.

[9] Yang Zixian. Research and application of convolutional NN IR technology based on DL [J].
Henan: China Equipment Engineering, 2018409 (23): 154-157.

[10] Yuan Bin, Zhang Yong. Traffic congestion detection algorithm based on image texture analysis [J]. Journal of Shanghai Institute of shipping science, 2015, 38 (4): 77-81.

[11] Zhou Junyu. Overview of convolutional NN application in image classification and TD [J]. Computer engineering and application, 2017, 53 (13): 34-41.