Research on the Application of Deep Learning in Computer Image Recognition

Miaofen Feng1,*
1College of Internet of Things, Jiangxi Teachers College, Yingtan, 335000, China
*Corresponding author e-mail: fengmiaofen@jxsfgz.com

Abstract. Deep learning is a new exploration, which is cited in machine research. It can mainly perform feature extraction and analysis by simulating the human brain. It makes the development of computer image recognition technology further. This article mainly discusses the classification of deep learning network models, and it analyzes the typical applications of deep learning models in image recognition.

Keywords: Deep Learning, Image Recognition, Neural Network

1. Introduction

This paper describes the research of deep learning in the field of image recognition. First, this paper introduces several typical deep learning network models, then it introduces the application of deep learning in the field of image recognition from face recognition, action recognition, and fall detection. Finally, we also discussed the research difficulties and development prospects in this field.

2. Deep learning network model

In the context of big data, deep learning has demonstrated its powerful data analysis and feature extraction capabilities. Many deep network models have been proposed and gradually optimized, showing their own characteristics. The following are several network models commonly used in the field of image recognition.

2.1. Convolutional neural network

CNN is similar to the traditional neural network structure. It is a multi-layer network structure. Each layer is composed of a large number of neurons. It can be regarded as a simple simulation of the human brain neural structure. The convolutional neural network consists of three parts. The first part is the input layer, the second part is a combination of n-layer convolutional layer and pooling layer, and
the third part is a fully connected multilayer perceptron classifier. Figure 1 is the CNN Basic structure diagram [1,2].

![Figure 1. Schematic diagram of the basic structure of CNN](image)

2.2. Deep belief network

DBN is stacked by multiple restricted Boltzmann machines (RBM). The RBM proposed by Hinton is a probability model based on energy theory, and defines a probability distribution through the thermodynamic energy function. DBN training is divided into two stages: pre-training and fine-tuning. In the first stage, each layer performs unsupervised learning, the output of the previous layer is the input of the next layer, and the initial value of the overall network parameter is the trained network parameter value; the second stage is to use the supervised learning pair The network is trained. After this algorithm uses training data to initialize the visual layer, it only takes a few iterations to get an estimate of the model [3].

2.3. Recurrent neural network

Recurrent neural network is a type of recurrent neural network that takes sequence data as input, recursively in the evolution direction of the sequence, and all nodes (recurrent units) are connected in a chain. The timing expansion diagram of RNN is shown in Figure 2.

![Figure 2. Timing expansion diagram of RNN](image)
3. Image recognition

3.1. Development of image recognition technology

In recent years, the application of image recognition technology in our daily lives has become more and more extensive. For example, it is used for medical cytopathological image recognition; used for agricultural crop disease image recognition; used for traffic control image recognition of traffic signs and so on [4]. Generally speaking, our daily life is inseparable from the application and development of image recognition technology. From the perspective of modern technology development, rapid technological innovation will inevitably provide a more comprehensive application platform for image recognition technology, and this will also put forward higher and stricter requirements for image recognition technology.

3.2. Image recognition technology

In our daily life, there are many ways to obtain information from the outside world. For example, the sense of smell to perceive smell, the sense of hearing to listen to the sound, the sense of taste to taste, etc., but in all the ways we use the most commonly used and most indispensable is the vision. Normal people cannot live without walking, eating, reading, etc, all of which use vision. We always say that the eyes are the "window of the soul". From this we can see how important vision is to our lives. It adds color to our lives. The image is the main way to record human visual information, and it plays an important role in our development [5].

Image recognition technology is also the pattern recognition technology of images. Its working mode is to model image information, build models, extract features, and then analyze and process the image, identify and classify the image according to the feature information, and finally achieve what we want The desired effect.

4. Typical applications of deep learning in image recognition

4.1. Face recognition

Face recognition is a technology for identification by comparing facial features. In recent years, as the deep learning boom has advanced, many researchers have combined face recognition with deep learning. Co-supervised by Softmax loss and center loss to train CNN to improve recognition accuracy; CNN-based face recognition system is strictly evaluated, and the impact of different structures of CNN on face recognition performance is quantitatively evaluated [6]. The results show that CNN Feature fusion of different layers in the middle can improve the performance of face recognition.

4.2. Human action recognition

Human action recognition has always been a hot and difficult point in computer vision research, and it has also been applied to many fields, such as virtual reality, human-computer intelligent interaction and other fields. Someone proposed a recognition method that combines acceleration sensors and deep
convolutional neural networks, which can effectively classify the five types of actions of the human body: walking, sitting, lying, running, and standing. This method uses the sliding window folding method to transform the sensor data into a three-channel data format similar to RGB images. The model automatically extracts data features to classify each action; the algorithm achieves a recognition rate of 91.26% on the Actitracker database [7,8]. Others have proposed a method of human action recognition based on genetic algorithm and convolutional neural network, using genetic algorithm to initialize CNN weights to minimize classification errors; using gradient descent algorithm to evaluate CNN classifiers, using the global genetic algorithm The search ability and the local search ability of the gradient descent algorithm find the optimal solution and improve the performance of the classifier.

In the case of big data, grouped convolutional neural networks are more flexible than convolutional neural networks, with fast training speeds and considerable recognition rates. But in the case of similar actions, the feature information extracted by the depth model is largely the same, which is prone to misjudgment of actions [9].

4.3. Fall detection

The incidence of falls is high in the lives of the elderly. The incidence of falls is high in the life of the elderly, so it is of great significance to automatically detect the accidental fall of the elderly, which can clearly reduce the risk of death. Feng et al. proposed a fall detection system based on computer vision, which uses deep learning to analyze fall events in the smart home environment; the foreground human body extracted by background subtraction is used as the input of the classifier, and the classifier output is combined with specific discrimination rules, To make the final decision on whether a fall occurs.

Jokanovic et al. proposed a method of using deep learning to detect human falls. This method automatically captures the complex characteristics of radar echoes and integrates time, frequency and distance information to reduce false alarms. Some scholars also combine color images and depth images to propose a fall detection method that combines target area positioning and deep feature learning; the algorithm combines depth information to achieve target area location, and uses convolutional neural networks to learn deep features to detect falls. The algorithm flow is shown in Figure 3. Show [10].
Compared with RGB images, depth images have the characteristics of less interference noise and accurate scene depth information. Using Kinect to collect depth images, combined with the distance information between the human body in the scene and the collection device, calculates the height and width of the human body through spatial mapping; through pixel conversion, realize the automatic positioning of the target area to optimize the image. Use the optimized depth image of the target area as the input of the neural network, the size of the input layer corresponds to the height and width of the input image, through two layers of convolution, two layers of pooling and one layer of full connection layer, and finally automatically extract 300 features to classify falls and non-falls.

Deep learning is applied to the fall detection and recognition of deep images. The algorithm converges quickly, its recognition rate is significantly improved, and its robustness is strong, which promotes the development of smart medical care in the context of artificial intelligence.

5. Conclusion

Although deep learning has many advantages in image recognition, there are still some problems. The current intelligent technology is still at the primary level of intelligence, and we still have many problems. Therefore, we should continue to study deep learning technology, and strive to make deep learning technology able to recognize and judge like humans.

References
[1] Liu Ji; Sun Rencheng; Song-lin qiao. Application of deep learning in medical image recognition [J]. Journal of Qingdao University (Natural Science Edition), 2018: 72-77+83.

[2] YANG Xiong. Research and Application of Deep Learning in Online Pornographic Image Recognition [J]. Journal of Jiamusi University (Natural Science Edition), 2018: 76-79.

[3] Fang Yanping. Research on Application of Computer Graphics Image Processing Technology in Copper Ore Recognition [J]. World Nonferrous Metals, 2016: 29-30.

[4] WANG Yi-wei. Research on Deep Learning Classification Network and Its Application in Computer Vision [J]. Communications World, 2019: 201-202.

[5] Wang Zhiyu, Li Ningning, Cong Xidong. Research on Application of Computer and Image Recognition Technology in Log Measuring [J]. Forestry Machinery and Woodworking Equipment, 2006: 13-16.

[6] FENG Yun, LIU Yongjun, ZHANG Cuijun, YANG Yang. A study on the application of adaptive learning algorithm in image recognition [J]. Computer Engineering and Applications, 2003: 72-73+80.

[7] Li Chaobo, Li Hongjun, Xu Chen. Application of Deep Learning in Image Recognition [J]. Journal of Nantong University (Natural Science Edition), 2018: 5-13.

[8] ZHAO Jing. Application of Deep Learning in Image Recognition [J]. Journal of Xiangyang Vocational and Technical College, 2019: 80-82.

[9] YI Shidong. Research on Image Recognition Algorithm Based on Deep Learning [J]. Network Security Technology and Application, 2018: 42-44.

[10] WANG Yuyan. Application and Development of Deep Learning in Image Recognition [J]. China Strategic Emerging Industries, 2017: 127-128.