Research on Deep Convolution Neural Network in Computer Vision

Dong Hao

Ocean University of China, College of Information Science and Engineering, Qingdao City, Shandong Province, 266003

Keywords: Computer Vision, Deep Convolution Neural Network, Application

Abstract: With the development of science and technology, the era of big data has followed, making deep convolution neural networks richer in network structure. Compared with traditional machine learning, it has advantages in feature expression and feature learning. Based on the deep learning algorithm deep convolution neural network model, the computer vision field has made remarkable achievements in recognition ability. This paper mainly discusses the application of deep convolution neural networks in computer vision.

1. Introduction

In machine learning, deep learning plays an important role, and has achieved remarkable improvement in the current research and development, and has gradually become a hot research project. In the article "Science" published by Geoffery Hinton, the concept of deep learning was first proposed, and based on this theory, deep learning was developed, and excellent results were obtained in image classification, language recognition, language processing, etc. have received extensive attention in the academic world. The current convolution neural network model is a content that is developed in object recognition and image classification, promotes the application of enterprise development resources in the Internet field, and further expands the research scale of deep learning systems.

2. Overview of Deep Convolution Neural Networks

Convolution neural networks generally organize important data of two-dimensional input and gradually establish artificial multilayer neural networks. For the whole network, each layer is composed of two-dimensional planes, and in each plane, there are many relatively independent neurons, and adjacent neurons can be connected to each other, but if two nerves the meta is at the same level and cannot be connected.

At present, the development of neural networks is entering a booming stage, and has gradually become an important research content of speech analysis and image recognition. As learning from higher areas is progressive, application changes are also more abundant. In essence, the convolution neural network is the first multi-layer neural network successfully developed. This model algorithm can make the input of multiple signals more convenient. With the gradual increase of learning depth, information learning is becoming a craze. The current convolution neural network can be perfectly applied in many aspects such as speech recognition, image recognition, and speech processing, and the machine learning is developing in a deeper direction. A typical convolution neural network recognition picture model is shown in Figure 1.

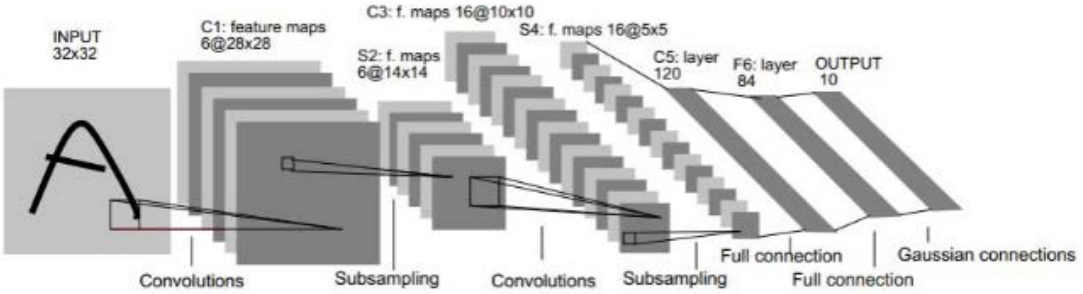

Figure 1. Typical Convolution Neural Network Model

3. Deep convolution neural networks and computer vision

3.1 Pooling operations

The convolution layer can collect features and transplant them into the classifier to develop training for them, so that a final classification result can be obtained. From a theoretical point of view, all the information features collected by the convolution layer are obtained. Porting into the classifier requires a certain amount of computation, especially for large-scale image resolution. For example, if the size of the input image sample is 98×98, let the convolution layer operate on the image through 200 6×6 convolution kernels, and one convolution kernel can output one (98-6+1)×(98)-6+1)=8649 Dimensional vector features. At the end, the convolution layer outputs a vector feature of 8649×200=1729800 dimensions. If you enter this result into the classifier, you will consume huge computing resources while training. However, the image belongs to “staticity”, that is, if the pooled area is m×n, after the final convolution is obtained, the features are divided into multiple independent regions of m×n, and the regions are operated and obtained.

3.2 Image classification

When performing image analysis, by dividing it, different types of operations is generated, this is one type of image classification. This method pays more attention to the integrity of the image, and each image is determined. Therefore, the current image classification calculation operation includes various types of labels carried in many data sets. Like CIFAR-10/100, or Caltech-101/256, but in ImageNet there are a large number of high-resolution label images, the number of images is far more than 14 million. In this case, most of the convolution neural network models basically rely on the input of image data size. When the image is cut, many of the original data in the image will be lost at this time; if the size of the image, the aspect ratio and the like are adjusted to prevent distortion and deformation of the image. Also consider whether the convolution layer will be constrained when inputting the image size, ensuring that it remains fixed during the input dimension.

3.3 Object detection

Compared with image classification, in the field of computer vision, object detection is more complicated. In an image, there may be more than one type of object, but each type of image needs to be separately positioned and recognized. Therefore, if you want to effectively improve the working effect of object inspection, you need to conduct more in-depth research and study on the model. At present, there are generally two problems in the detection of object models for convolution neural networks. One is to summarize and classify the selected regions, and the other is how to select regions. This also reflects the difficulty of object inspection. For the solution of the problem, the content steps are more complicated and complicated, and more requirements are placed on the model standard. Therefore, if the object detection model is better developed, not only the network structure needs to be improved, but also the model training process and training method should be further optimized, so that the object detection can be better integrated into the practical operation.

3.4 Image segmentation

For a picture, there may be many different objects, and it is hoped that each pixel will be predicted and the part to which it belongs (background, animal, person, etc.). In the past period of time, the image threshold segmentation method was the earliest method used in this field, and then the threshold method based on fuzzy set and non-Shannon entropy was gradually derived. However, with the increasing use of deep learning applications, the field of image segmentation has been further developed. Scientists first use the classification network, and based on the data obtained from the image classification training, the image segmentation model is constructed, and the deep features of the network are combined with the shallow features. The deconvolution layer is used to enlarge the original image to obtain more accurate segmentation. As a result, we call this a jump structure.

4. Conclusion

At present, in the deep learning application, only simple reasoning calculation can be performed, and such impressive results can be obtained in the fields of image and speech, which also reflects from the side that if the research on convolution neural networks is intensified Strength can play a greater role in other fields and gradually move toward artificial intelligence.

References

[1] Huang Bin, Lu Jinjin, Wang Jianhua et al. Object Recognition Algorithm based on Deep Convolution Neural Network [J]. Journal of Computer Applications, 2017(12).

[2] Lu Hongtao, Zhang Qinchuan. Review of the Application of Deep Convolution Neural Networks in Computer Vision [J]. Data Acquisition and Processing, 2017(01).