Research on Convolutional Neural Network Based on Deep Learning Framework in Big Data Education

Xuan Luo(✉)

Wuchang Shouyi University, Wuhan 430064, Hubei Province, China

Abstract. Deep learning is an important part of the development of artificial intelligence. Deep learning has made breakthroughs in many fields (such as image recognition, speech recognition, natural language processing), and has made gratifying achievements in the application of traditional algorithms which are not easy to solve, It includes automatic driverless vehicle, automatic pattern recognition, automatic simultaneous interpretation, commodity image retrieval, handwritten character recognition, license plate recognition, etc. In recent years, with the continuous improvement of research and development personnel's requirements for deep learning development process, the traditional deep learning programming methods can not meet the current needs. The traditional deep learning programming methods will take researchers and developers months or even years to implement the most basic algorithms. At the same time, the traditional deep learning programming methods can not meet the current needs, In this paper, a variety of in-depth learning frameworks, including CAE, have been developed for some of the world's top research institutions. These deep learning frameworks not only provide efficient and fast development models for scientific research institutions and related developers, but also provide several convolutional neural network models for developers to study and improve on the more advanced and perfect convolutional neural network models.

Keywords: Deep learning · Artificial neural network · Convolutional neural network · Deep learning framework · Caffe

1 Related Concepts of Personalized Recommendation System

Deep learning is a branch of machine learning based on artificial neural network. "Artificial neural network" ("Ann") tries to build "artificial" biological neural cells (i.e. neurons) and neural networks, and realize the functions of human brain nervous system in information processing, learning, memory, knowledge storage and retrieval at different levels. Deep learning comes from "artificial neural network", which combines the bottom features to form abstract high-level attribute categories or high-level features, and then finds the distribution feature representation of data [1].

The earliest deep learning framework is recognized as a new cognitive machine proposed by Fukushima K. Neocognitron in 1980. In 1989, in the research of handwritten

M. A. Jan and F. Khan (Eds.): BigIoT-EDU 2021, LNICST 392, pp. 109–114, 2021.
https://doi.org/10.1007/978-3-030-87903-7_16

character set recognition, Yann lecunn1 used back propagation algorithm, MP algorithm and graphics card acceleration, which is similar to the combination of new cognitive machine, weight sharing and convolution neural layer. Since then, this algorithm has become a basic part of many deep learning models. This is also an important milestone in the development of convolutional neural network. Although the algorithm can be successfully implemented, the cost of calculation has exceeded the normal range, so that it can not be applied in the actual research work. Then in 1991, Sepp Hochreiter proposed a solution to the gradient vanishing problem. However, with the development of computational model, SVM (support vector machine) for unknown problem modeling in many ways is more simple and easy to use than neural network, so the development of neural network in the 1990s compared with the slow progress.

2 Overview and Structure of Artificial Neural Network and Convolutional Neural Network

Artificial neural network is to use the connection between neurons, input and output to model, or to study the relationship between data. At present, the most typical artificial neural networks are BP network, Hopfield network, Boltzmann machine, SOFM network and art network [2].

The function of perceptron is to classify and recognize two kinds of attributes. Although the structure of perceptron unit is simple, it has all the elements of artificial neural network. The perceptron is composed of input part and output part. The input and output parts are directly connected. The perceptron can get an output value f (x) by mapping the input value X.

2.1 Feedforward Back Propagation Algorithm

The iterative process of feedforward back propagation algorithm is mainly divided into the following two steps: 1. Input a series of samples and output a result through forward propagation; 2. Adjust the weights in the neural network by calculating the deviation. In the whole calculation process, feedforward and back propagation are carried out alternately until the deviation is within a threshold or the number of iterations reaches a certain requirement. First, explain what the variance loss function is. For a problem with Class C and N training samples, the loss function is defined as:

$$E^N = \frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{c}(t_k^n - y_k^n)^2 \tag{1}$$

Because the error of the whole data set is obtained by summing all the errors in each class, when calculating the feedback function, we need to focus on a single mode. The nth class error of the feedback function is:

$$E_n = \frac{1}{2}\sum_{k=1}^{c}(t_k^n - y_k^n) = \frac{1}{2}\left\|t^n - y^n\right\|_2^2 \tag{2}$$

2.2 Convolution Neural Network

Generally speaking, using multi-layer back propagation network to train data can get better results. The traditional multi-layer back propagation network needs to collect the effective information in the data through manual input. It needs to go through a complex preprocessing process to get the feature variables, and then classify and recognize the above processed feature variables through the definition of a trainable classifier. This kind of traditional multi-layer back propagation network can play the role of classification [3].

Compared with the traditional shallow neural network, the structure of convolutional neural network is much more complex. Each layer of neurons uses local connection to connect, neurons share connection weight and use reduced sampling in time or space to make full use of the characteristics of data itself. These characteristics determine that the dimension of convolution neural network is greatly reduced compared with traditional neural network, Thus, the complexity of the calculation is reduced. Convolutional neural network is divided into two processes, namely convolution and sampling. Convolution is mainly to extract the upper data and abstract, while sampling is to reduce the dimension of the data.

Activation function is an important part of neural network. Through the function transformation of the input of neural network, the appropriate output is obtained. Generally, the activation function injects nonlinear factors into the neural network with poor linear expression ability, which makes the data separable in the nonlinear condition, and also can sparse expression of data, and more efficient data processing (Fig. 1).

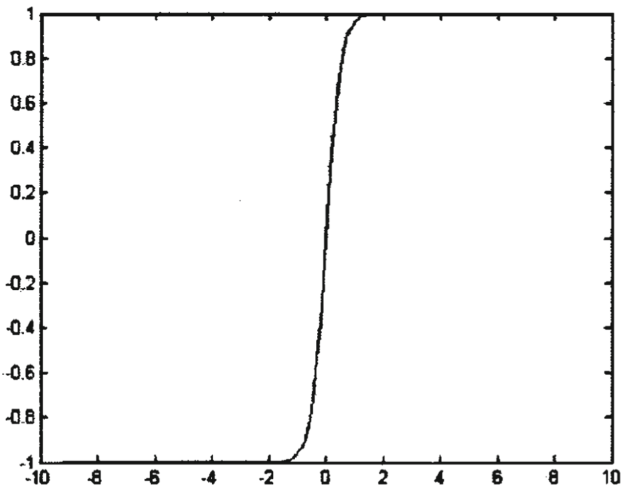

Fig. 1. Tanh activation function

3 The Construction and Principle of Caffe Deep Learning Framework

Cafe deep learning framework provides a pure and changeable framework for multimedia scientists and developers. Cafe framework includes the most advanced deep learning algorithm and a series of reference models. O Cafe framework is written in Python language and MATLAB language, and calls the C + library of Berkeley University, It is used to train and deploy common convolutional neural networks and other effective depth algorithms for common structures. When cafe is applied to commercial and Internet scale media, it needs CUDA GPU for computing. It can process 40 million images (each image is about 25 ms) on independent K40 or Titan GPU every day. The cafe framework realizes the model representation of extraction, can carry out simulation experiments and support more efficient development, and can seamlessly transform with the prototype deployment platform in the cloud environment.

3.1 Overview of Caffe Deep Learning Framework

The key problem of multimedia data analysis is to find the effective representation of perceptual input - image, sound wave, touch and so on. The manual feature extraction method of the input data has entered a stable stage in recent years, but the way of extracting features in the structure of depth composition by convolution has been steadily rising. Depth model has been better than manual feature representation in many fields, and it can be extracted in those areas where manual features are imperfect.

Large scale visual recognition is the most concerned area. Caffe, the deep learning framework used in this paper, has reached a very high level. These convolutional neural networks, or CNNs, are trained differently in the back propagation of convolutional filter layer and other layers such as correction layer and pool layer. According to the early results of data classification in the 1990s, these models are far superior to all known large-scale visual recognition algorithms, and have been applied to large-scale businesses, such as Google, Facebook, Baidu for image understanding and search [4].

3.2 The Characteristics of Caffe In-Depth Learning Framework

Caffe provides a complete toolkit for training, testing, fine tuning, and developing models, and it has well-documented examples for these tasks. Similarly, it is an ideal starting point for researchers and other developers to enter cutting-edge machine learning, which makes it available for industrial development in a short time. Modularity: Cafe follows the principle of modularity as much as possible, which makes the new data format, network layer and loss function easy to expand. The network layer and loss function have been defined, and a large number of examples show how these parts form an identification system for different situations.

Separation of representation and Implementation: the definition of cafe model has been written into configuration file in protocl buffer language. Cafe supports network construction in any directed acyclic graph. According to the instantiation, cafe keeps the memory needed by the network and extracts the memory from the host or the bottom of GPU. Only one function needs to be called to convert between CPU and GPU.

3.3 Training of Convolutional Neural Network Model Based on Caffe

MNIST data set is a large binary handwritten numeral data set, which is a standardized data set in machine learning and deep learning. The MNIST dataset contains 60000 training samples and 10000 test samples. Mnst dataset is a kind of black-and-white image which keeps the aspect ratio of the original 20 × 20 pixel handwritten character image in NIST dataset and expands the center of NST image to 28 × 28 pixel. The special database SD-3 and SD-1 of NIST dataset are composed of SD-3 as training set and SD-1 as test set. The SD-3 data set is clearer than the SD-1 data set, because the SD-3 data set is collected from the employees of the census and Census Bureau of the United States, while the SD-1 data set is collected from high school students in the United States. It is concluded that the training results should be independent of the selected data set. Therefore, mnst data set selects 30000 samples from SD-3 and SD-1 data sets respectively. The test set selects 5000 samples from each of the two data sets. A total of 60000 training samples were collected from 250 writers, and it can ensure that the writers of the training set and the test set do not coincide.

In this paper, we use the relu activation function instead of the original sigmoid activation function. Inspired by alexnet, we add the activation function in the convolution layer and the downsampling layer to make the network converge more quickly. There is no normalization level processing, which will reduce the training speed of the network. If mnst is normalized, errors may occur.

4 Conclusion

Today, with the development of artificial intelligence, deep learning has realized many tasks for scientific research and traditional industries that traditional methods can't accomplish. In recent years, deep learning has become the focus of scientific research. Although there are still many problems to be solved, the breakthrough of deep learning in image recognition and speech recognition has brought hope to scientists and related researchers. Google, Baidu and other research centers have developed a number of deep learning products, and have been applied in many civil fields. Many efficient and practical deep learning development tools are updated at an unprecedented speed. Convolutional neural network is the best in the field of deep learning. Most of the deep learning algorithms and network structure are convolutional neural network. Convolution neural network has been developing rapidly in the past twenty years. After 2010, it has developed more rapidly, and has made many achievements in the field of scientific research and business application (such as shopping picture recognition, Google self driving car, voice assistant, etc.).

Convolutional neural network has become a hot topic in deep learning. At present, there are many deep learning frameworks for developers and researchers to use. There are still many problems in the field of convolutional neural network and artificial intelligence that need to be further studied by researchers. The neural network of handwriting recognition data set can do a lot of follow-up work.

References

1. Yu, K., Jia, L., Chen, Y., et al.: Yesterday, today and tomorrow of deep learning. Comput. Res. Develop. **50**(9), 1799–1804 (2013)
2. Ding, S.: Basis of Artificial Neural Network. Harbin Engineering University Press (2008)
3. Ao, D.: Application of Unsupervised Feature Learning Combined with Neural Network in Image Recognition. South China University of Technology (2014)
4. Xuke: Application of Convolutional Neural Network in Image Recognition. Zhejiang University (2012)