Research on optimization of deep learning algorithm based on convolutional neural network

Luo Yiyue*, Fan Yu², Chen Xianjun³

Haikou College of Economics, Haikou, Hainan, 570203, China
tg667788@xzcstudio.com

Abstract: Convolutional neural network is a representative one of deep learning algorithms. It plays an important role in improving the learning level of artificial intelligence. Therefore, this article discusses the optimization of convolutional neural networks and deep learning algorithms. The optimization of deep learning algorithms and its application in actual image recognition are discussed for reference.

1. Introduction
In recent years, in the context of the continuous development of computer and big data technology, deep learning algorithms have been more widely used and have now become an important tool for intelligent processing of various information. After the continuous improvement of deep learning algorithms, convolutional neural networks have also been rapidly developed. Most of the current deep learning algorithms are implemented based on convolutional neural networks, and they play an important role in solving the difficulty of artificial intelligence to deal with specific rule tasks. Therefore, research on the optimization of deep learning algorithms for convolutional neural networks undoubtedly has important practical significance.

2. Overview of convolutional neural networks

2.1. Basic flow of convolutional neural network
In the convolutional neural network, it has a certain degree of displacement and rotation invariance through local receptive fields, weight sharing and downsampling, and significantly reduces training time and difficulty. In the convolutional neural network, after detecting the input signal, the convolution uses a 3×3 convolution kernel to slide. In the sliding process of the convolution kernel, weight sharing occurs, and its parameters are equal to the size of the convolution kernel. After the convolution step is over, the pooling layer starts to run. By taking the maximum value of the result in the feature map, it reduces the complexity of the calculation and makes the characteristics of the target signal more clearly expressed [1].

2.2. Main application of convolutional neural network
Since the main way for humans to obtain information is vision, image recognition technology came into being. With the advent of the big data era, the previous image recognition technology has been unable to meet actual needs. Therefore, as a representative of deep learning, convolutional neural networks have been gradually applied in image recognition [2]. Through neural network training, it can effectively
identify common natural images without artificial feature extraction. It has been widely used in the fields of face recognition, text detection and target tracking [3].

3. Basic deep learning optimization algorithm for convolutional neural networks

3.1. Gradient descent method
The traditional neural network model usually consists of multiple layers, where the layer can be interpreted as a specific structure that receives input, generates output, and performs parameter transformation. It can minimize the loss function by adjusting the linear transformation parameters. Since the gradient descent method uses the entire training set for calculation, the direction of each update is determined and can converge to the extreme point. However, due to its long learning time, when the amount of data in the training set is large, it tends to take up a lot of memory [4].

3.2. Implement logic in deep learning framework
In contrast, in single-layer neural network training, the gradient calculation is relatively simple, and the error loss function can also be directly expressed in the form of a weight function. The gradient calculation is mainly realized based on the back propagation algorithm. It applies the chain rule to accumulate and superimpose the local gradients of all paths between two nodes. Although this involves exponential calculation, after adopting the calculation idea of differentiable programming, the calculation difficulty will be greatly reduced. Specifically, differentiable programming can be divided into forward phase and reverse phase. In the forward phase, training examples are input to the convolutional neural network for weighting calculation. In the reverse phase, the gradient of the loss function with respect to different weights is used to update the weights [5].

3.3. Adaptive optimization algorithm in deep learning
In deep learning, the learning law is one of the most important hyper-parameters, and it must be controlled within a reasonable range. Too fast or too slow will affect the training effect. In order to flexibly adjust the learning law according to the actual situation, an adaptive learning law is generated, which can use different learning laws according to different parameters to ensure that the entire training process converges faster or has a higher accuracy. There are several commonly used adaptive optimization algorithms, including stochastic gradient descent, momentum-based gradient descent, AdaGrad, RMSProp and Adam.

4. Further optimization of deep learning algorithms

4.1. Optimization method based on data filtering

4.1.1 Data set. The ability of depth learning algorithm is usually detected by MNIST handwritten character data set (Figure.1) in the optimization of depth learning algorithm. In practical work, the method of constructing MNIST classification data set into segmentation data set is used. The handwritten character picture is spliced into the whole image.
4.1.2 Optimization method based on gradient level. The stochastic gradient descent method is one of the commonly used methods in convolutional neural network training. In the calculation of this method, the errors produced by different training samples are also different, which makes the gradients different. Usually, in the data screening process, the pre-training model inspection method is used to calculate the gradient magnitude of each training sample, and the importance of the picture is determined according to the calculation result.

Specifically, random, maximum gradient and non-maximum and maximum gradient three methods are commonly used data screening methods. In the training process of convolutional neural network, as the number of iterations increases, the gradient of the sample changes quite obviously. The gradient contribution of most samples will not exist, and some samples that maintain a larger gradient generally have the problem of poor overall quality. Therefore, the non-maximum and maximum gradient method is often used for screening to increase the stability of training samples.

4.1.3 Active incremental optimization method. As deep learning algorithms will face increasing amounts of data, they must be able to fully respond to newly emerging data, and the trained system should have the ability to make appropriate changes to ensure that they can respond to new data in new data. The content is learned to avoid the time cost of training new models. However, because the problem of catastrophic forgetting often occurs after the neural network adds new data, an active incremental optimization method arises from this. This method is similar to transfer learning.

In the application process of the active incremental optimization method, the stability of the candidate samples must first be analyzed. It can perform enhancement processing on the individuals in the candidate samples, convert the data individuals into several data blocks, and then use the pre-training model to predict the data blocks. In the prediction process, the data may be misjudged due to the prediction probability. This also requires the definition of the entropy of each data block. At the same time, it should be noted that since the data enhancement process will cause noise in the data set, it is necessary to randomly select the shape of the candidate data and train the selected data to improve robustness.

4.2 Optimization method based on network structure
On the premise that the data set has been determined, the design and optimization of the network structure need to be carried out. The design of the network structure needs to be selected for different data types to ensure feature extraction and the increase of gradient return paths. In deep learning algorithms, multi-tasking schemes frequently appear, making multi-path learning a problem worth studying.

4.2.1 Multi-path optimization scheme. According to the relevant research of the researchers, different levels of networks have different characteristics. In terms of noise, the shallow network gradient is
similar to brown noise, while the deep network is similar to white noise, and the training difficulty is greatly increased after the gradient is reduced. Since the image is locally correlated, it can be assumed that the gradient is similar to it. If the gradient is close to white noise, it can be preliminarily inferred that the updated gradient is meaningless. On this basis, further analysis of the corresponding structure can be realized, and the correlation can be analyzed according to the gradient flow.

In order to further improve the experimental results of the algorithm, usually a layer skip structure is added to the residual network at the initial stage of training to perform a layer skip connection. The experimental results show that the performance difference is obvious after the layer jump connection is performed. In the initial training stage, better results have been achieved on the training set and the verification set. It should be noted that in order to ensure that the layer-jumping structure can still achieve optimal results at the end of training, more data needs to be added to avoid the complexity of the layer-jumping structure from affecting the generalization performance.

4.2.2 Multi-scale optimization scheme.
According to the relevant research theory, a visual system should be able to accurately identify and calculate the shape, direction and color of objects, and can effectively identify them from more complex images. Therefore, the multi-scale method is needed to describe the operator of the image.

When using deep learning algorithms for visual recognition, the features of each layer are different. Generally speaking, the higher the level, the more complex the focus, and the final loss function is usually only located in the last layer, which makes different layers have different learning rates, thus, the concept of multi-scale loss function came into being. Through the combined calculation of the correlation function, the multi-scale scheme can be effectively controlled.

In order to retain more image features, a large number of filters are usually used in the middle layer. Therefore, the feature maps of the middle layer are also relatively large. Although the problem of dimensionality in the design of convolutional neural networks has been solved, the problem of spatial and channel redundancy still exists. Therefore, small convolution kernel and separable convolution are used to solve these redundant problems. At the same time, it is necessary to add structural recombination feature map and carry out weighted average calculation to get better results.

4.3. MobileNet algorithm and improvement experiment
MobileNet algorithm is mainly based on AM-Softmax classifier, its expression is as follows.

\[
L_{AMS} = -\frac{1}{n} \sum_{i=1}^{n} \log \frac{e^{x(w^T_{y_i} f_i - m)}}{e^{x(w^T_{y_i} f_i - m)} + \sum_{j=1, j \neq y_i}^{e} e^{xW^T_{j} f_i}}
\]

Among them, \(m\) is the additional margin introduced to improve the classification accuracy.

In the test process, first, input the corresponding picture (as shown in Figure 2) into the detection classifier, and perform affine transformation on its feature points to obtain the facial image, and adjust the image, and finally, the picture Equalize the histogram to get the final result (Figure 3).
It should be noted that because multiple faces are easily detected in the actual face detection process, in practical applications, it is usually necessary to use the support of big data technology to establish a larger high-quality data set to avoid the problem of false detection as far as possible.

5. Conclusion

At present, depth learning algorithm has become a hot research direction in the field of artificial intelligence. For the important part of artificial intelligence —— image recognition and processing, convolution neural network is the most commonly used tool. Therefore, it is necessary to study and optimize the deep learning algorithm based on convolution neural network to ensure that the depth learning technology is more widely used and to promote the further development of artificial intelligence.

Acknowledgments

Fund Project 1: supported by Hainan Provincial Natural Science Foundation of China: Research on the Key Technology of Deep Learning Face Recognition CNN Based on Edge Computing (Project Number: 619QN249)

Fund Project 2: Scientific Research Project of Higher Education Institutions in Hainan Province: Research and Application of Deep Learning Face Recognition System Based on RaspberryPi (Project Number: Hnky2019-74)

Fund Project 3: Scientific Research Project of Higher Education Institutions in Hainan Province: Research on Poisson 3D Surface Reconstruction Based on Distributed Multi-Grid Numerical Algorithm (Project Number: Hnky2018ZD-11)

References

[1] Gong Chengjun, Pu Yuqing. Intelligent recognition method based on convolutional neural network [J]. Computer and Digital Engineering, 2020, 48(11): 2633-2636+2749.

[2] Chang Ying, Chang Dajun. Research on improved convolutional neural network solder joint defect recognition algorithm[J]. Laser Technology, 2020, 44(06): 779-783.

[3] Liu Tianbao. Research on face detection algorithm based on convolutional neural network[J]. Software Guide, 2020,19(10):66-70.

[4] Cheng Jinfu, Zhang Wanhe. Research and application of deep learning in image recognition[J]. Electronic World, 2020(19): 48-49.

[5] Wang Tong. The structure and related algorithms of convolutional neural networks in deep learning[J]. Information and Computer (Theoretical Edition), 2020, 32(08): 41-43.