Research on Image Recognition Method of Convolutional Neural Network with Improved Computer Technology

Xiaohong Li¹,*, Xiangfeng Lv¹

¹School of Information Science and Engineering, Tianjin TianShi College, Tianjin, China, 301700

*Corresponding author e-mail: 214@tianshi.edu.cn

Abstract. It has been a goal of artificial intelligence to make computers recognize objects and have the vision similar to human beings. After years of development, considerable progress has been made, but it is not satisfactory. It can be said that deep convolutional network is now the best algorithm for image recognition, which is also the reason why this paper decides to adopt deep convolutional network algorithm. The convolutional neural network uses shared parameters between the convolutional layers, which not only reduces the required memory size, but also reduces the number of parameters to be trained and improves the performance of the algorithm. The difficulty of image recognition is hard to find a way to from the natural environment the boundary shape, texture, angular point, and image characteristics such as concept, and the image itself are highly susceptible to the influence of natural environment and changes in perspective, dimension, twist deformation, interference, light, shade, difficulties of the difference of background doping and within the class, make the computer more difficult to get about abstraction and understand the expression of natural images.

Keywords: Improved Technology, Computer Technology, Convolutional Neural Network, Image Recognition Method

1. Introduction
In order to test the performance and accuracy of Image Recognition algorithm, the researchers created a accurate and sufficiently Large Image database ImageNet, on the basis of the database platform, the annual large-scale Visual Challenge Image. net Large Scale Visual Recognition Challenge (hereinafter referred to as ILSVRC), this is the largest Image Recognition, basic can represent the highest level in the field of computer vision, 2012 years ago, Image Recognition error rate has been high (about 26%). The classical convolutional neural network algorithm does not deeply explore the algorithm. Image recognition is a technology that compares the stored information (the information stored in memory unit) with the current input information (the information received by the senses at that time), USES the computer to process the image, analyzes the context, and understands and recognizes various objects. Due to the complexity and particularity of image acquisition and image itself, image recognition technology and methods have been the focus of research in computer vision.
and related fields. At present, image recognition technology has been widely used in unmanned driving, smart city, intelligent transportation, logistics management, search engine, network communication, virtual reality, medical health, remote sensing imaging and artificial intelligence and other fields\[1\].

In addition, the researchers based on CNKI philosophy and the humanities, social science, information technology pool, according to the keywords "computer-aided translation, machine and auxiliary teaching, CAT teaching, computer-aided translation course, machine and auxiliary" to retrieve the papers about computer aided translation teaching, including meetings, journals, master's, doctoral thesis. The retrieval date was January 1992 to January 2016. The results showed that there were 96 articles on COMPUTER-aided translation teaching, which were divided into undergraduate English major teaching, undergraduate non-English major teaching, master's teaching and doctoral teaching. At present, most image recognition methods are realized by two steps: feature extraction and classification recognition. "Feature extraction" is mainly aimed at the characterization of image contents for specific recognition tasks. It is usually necessary to train specific marked data sets to extract image features. The quality of feature extraction directly affects the accuracy of final classification recognition. "Classification recognition" usually adopts the more mature classification algorithm in the field of machine learning to make category prediction judgment for instances of unknown class labels\[2\].

2. Basic concepts of convolutional neural networks

2.1. The convolutional neural layer
Convolution is a commonly used algorithm in image recognition. It means that each pixel in the output image is obtained by the weighted average of the pixels in the corresponding small area of the input image, and this area is called the convolution kernel. In general, the convolution kernel is square, so it is expressed in terms like m by m, where m is the length of the region. The convolutional neural layer, in fact, carries out convolution operation on every point of the image, and the convolution kernel is taken as the training parameter. The convolutional neural layer can be regarded as an "abstract" operation on the input image. After several processing, the "eigenvalue" of the image can be extracted\[3\]. However, such machine learning methods are highly task-related and often require manual design of image features. It is doubtful whether they can maintain their original performance when they are separated from specific data sets and oriented towards general natural image recognition tasks. On the other hand, the computational cost of such methods is high, and it is still difficult to realize real-time image recognition with phased manual marking features. Therefore, some researchers devote themselves to the study of Deep Learning in which "self-learning features" from original natural images rather than "artificial design features" in advance\[3\].

2.2. ReLU nonlinear function
In the structure of the classic, using neural network activation function is Rectified function, in the convolutional neural network, the nerve layer called ReLU commonly. In the literature, ReLU is demonstrated in detail. In the classical convolutional neural network algorithm, ReLU is also compared with the traditional activation function, and the conclusion is that ReLU can reduce the training time and improve the algorithm performance. Deep convolutional networks generally require a large amount of data for training, so that it is almost impossible for traditional activation functions to cooperate with the convolutional neural layer to complete the training. In this case, ReLU is almost the best option. In this paper, if not specifically, each convolutional neural layer is then treated with ReLU. The biggest difference between deep learning and most machine learning algorithms is feature extraction. Deep learning is like a "black box", all depend on the depth study of feature extraction algorithm of automatic processing, and deep learning often USES the multi-layer network structure step by step the underlying characteristics of the image combination form of high-level characteristics, from local information abstraction to the high-level semantic information, gradually forming a
multi-layer transmission, abstract and iterative refinement process step by step; The whole processing process does not require manual marking of features or additional manual intervention. Feature extraction and classification recognition are connected in tandem to complete end-to-end image recognition, which also saves cost for expensive and time-consuming feature design and extraction\[^4\].

2.3. **Pooling layer**

After the input image is processed by the convolutional neural layer and ReLU, each pixel in the image contains information of a small area around it, resulting in information redundancy. If the image containing redundant information continues to be used, not only the algorithm performance will be reduced, but also the translation invariance of the algorithm will be destroyed. In the fields of pattern recognition, computer vision, artificial intelligence, machine learning and deep learning, image processing and recognition technology has been the earliest attempt and exploration of researchers. From the early recognition of specific objects in the image, to the later classification of objects and to the present semantic analysis and understanding of the image, the core content of the image recognition task still lies in the image feature extraction. Therefore, the image recognition methods under different task backgrounds are all feature extraction methods based on feature engineering\[^5\].

2.4. **The Normalization Layer**

The normalized layer is a layer of nerve designed to make an image more contrasting, and the effect is similar to "adding contrast" to an image. Obviously, an algorithm is needed to calculate the "average", and then adjust each pixel of the image according to certain rules, so that the body of the image can be more distinguishable from the background. At present commonly used algorithm is usually Local Response Normalization (hereinafter referred to as the LRN), this algorithm can effectively improve the subject part and other parts of the degree of differentiation. However, LRN is not necessary, and the improvement of effect is not very obvious. Therefore, in the use of deep convolutional network, LRN is generally used only when the convolution kernel is large, that is, the processing is relatively "rough". According to the information content contained in the image itself, the commonly used image features can be divided into two categories: global feature and local feature. In traditional machine learning, most of the image recognition method is to dig deeper into the global or local features used to identify, thus also spawned image recognition method based on color features, image recognition method based on texture feature and image recognition method based on shape feature "and image recognition method based on spatial relationship characteristics, etc. These methods require sample learning to extract the underlying features of the image for feature coding, fusion and aggregation. Finally, the classification of unmarked data is obtained by using the more mature classification algorithm in machine learning, so as to complete the task of classification and recognition\[^6\].

3. **Design of network structure**

3.1. **Introduced structure**

In recent years, the convolutional neural network (CONVOLUtional neural network) has developed rapidly and gained a lot of praise in industry and academia. It evolved from a biological concept and is a variant of Multilayer Perceptron (MLP), which was first used to simulate the visual behavior of the cerebral cortex and then excelled in computer visual recognition tasks. In the literature, the classical convolutional neural network USES a large convolution kernel of $11 \times 11$. Although this can effectively "extract" images, the training parameters are too many, which limits the performance of the overall algorithm. In order to improve the performance of the algorithm and increase the depth of the neural network, a convolutional layer with a smaller convolution kernel is used in the newly introduced structure. At the same time, in order to adjust the algorithm conveniently and give consideration to the expansibility of the algorithm, the introduced structure should be able to be
superposed directly without introducing new neural layers. To this end, the new structure introduced in this article is shown in Figure 1:

![Figure 1. The structure introduced in this article.](image)

3.2. Overall structure of the network
In general, common convolutional neural networks are composed of input layer, convolutional layer, activation layer, pooling layer, full connection layer and the final output layer. From input to output, the relationship between layers of the convolutional neural network is established through different computational neural nodes to transmit input information layer by layer. The continuous convolution pooling structure decays, deduces and gathers the feature signals of the original data, maps them to the hidden layer feature space, and then the full connection layer classifies the output according to the extracted features. In the initial experimental stage, we tried to use the structure introduced in Figure 1 for 4 times, but the training time was too long. The structure shown in Figure 1 was used three times, so the training time was acceptable, but the effect was not satisfactory. Therefore, in the final design, a neural layer of a 7×7 convolution kernel was considered to replace the original structure, which could not only ensure the training time was not too long, but also effectively reduce the error rate. After the convolutional neural layer, traditional fully connected neural networks and Softmax regression were used, which are also the classic constructs used in the literature. In addition, Softmax output is not a single image recognition classification, but the highest output probability, which makes it easier to measure the accuracy of the algorithm.

4. Experimental results and analysis
In the convolutional layer, one or more convolution kernels (also known as Filter) extract pixel-level image features through convolution operation, and the result of convolution operation constitutes the feature mapping relationship from input to output after the mapping transformation of activation function. Each convolution kernel traverses the entire feature map with a mechanism similar to the sliding window. The convolution kernel gathers and fuses the feature information of each small area to complete the characterization of a small local area of the image. To implement the convolutional neural network using Caffe, data pretreatment is required, and then corresponding files are configured according to the designed network structure. Caffe can be used for training and test results. After the algorithm test results of this paper are obtained, some representative algorithms and their results of ILSVRC are compared, as shown in Table 1.
In Table 1, except for the algorithm shown in the last row (ISI), depth convolutional network is used. It can be seen that the results obtained by the convolutional neural network are quite ideal. Among them, the top-ranked algorithms all used the neural layer with a smaller convolution kernel: VGG used all the neural layers of the 3×3 convolution kernel, while GoogLeNet designed a module with a structure more complex than the one in this paper and used a deeper network structure, and the results obtained were the best that could be achieved before. Since the early days of artificial neural network, the mapping relationship from input to output is established among the interconnected neural nodes through activation functions. The activation layer mainly sets the activation function in the convolutional network. Its essence is a function mapping, which maps and transforms the input data and provides the nonlinear modeling capability of the network. During the calculation, the size of the original data is not changed element by element, that is, the size of the input and output data is the same.

5. Conclusion
The new convolutional neural network structure introduced in this paper can effectively improve the accuracy of image recognition and has good scalability. Compared with the classical convolutional neural network, not only the effect of the algorithm is improved, but also the training parameters are reduced and the training time is shorter. Although there is still some gap from the current level of the world's advanced algorithms, the accuracy rate of the algorithm is still much higher than that of traditional image processing algorithms. In the future, we will continue to conduct in-depth research in this area. Convolution neural network is a kind of inspired by the mechanism of cerebral cortex visual neurons signal transmission and form the deep learning of neural network model, it relies on its own local receptive field, parameter sharing, and other unique characteristics of sparse connection and down sampling, showed a strong feature in all kinds of recognition task since the study and classification recognition ability. In particular, when the input is a two-dimensional image, its continuously superimposed layered network structure can spread the original feature information of the image layer by layer and abstract it layer by layer, thus achieving a certain degree of translation, rotation, scale and deformation invariance, which is of extensive research and application value in many fields such as computer vision.

Acknowledgments
Fund project: National Undergraduate Innovation and Entrepreneurship Training Project (No.:202010859003).
References

[1] Zhang Dai-lei, HUANG Da-nian, ZHANG Chong. Application of BP Neural Network Based on Genetic Algorithm optimization in density interface Inversion [J]. Journal of Jilin University (Earth Science edition). 2017.

[2] Wu Yubo. Development and Application of Image Recognition Technology [J]. Electronic Technology and Software Engineering. 2017.

[3] Zhou Feiyan, Jin Linpeng, DONG Jun. Review of research on convolutional neural networks [J]. Acta Computerica. 2017.

[4] Song Junhui, WANG Li-Gang, HUA Jian, GAO Menglong, FAN Lin-Lin, Wang Xun. Application of meteorological Satellite Imagery in Weather Forecast in Yunfeng Watershed [J]. Agricultural Science and Information. 2016.

[5] Luo Jian, JIANG Min, LIU Xing, ZHOU Long. Rgb-d object recognition in multiscale convolutional recurrent neural networks [J]. Research in Computer Applications. 2017

[6] Li Yandong, HAO Zongbo, LEI Hang. Research Overview of convolutional neural networks [J]. Computer Applications. 2016.