Development of CNN and its Application in Education and Learning

Yiming Zhou[1,a] and Yongmei Su[1,b,*]

[1] Beijing University of Posts and Telecommunications, Beijing, China

[2] Zhengzhou Technical College, Zhengzhou, Henan, China

[a] zhouyiming@bupt.edu.cn, [b] 15038310680@163.com

Keywords: Convolutional neural network, Educational learning materials, Image classification

Abstract: This paper introduces the concept, development process, realization method and educational application of neural network in detail. Under the classification of CNN, LenNet-5, AlexNet, ZFNet and VGGNet have all been introduced and summarized in terms of structure and characteristics. As for local CNN, R-CNN, Fast R-CNN, Faster R-CNN have relevant development and optimization, and this paper will discuss relevant optimized convolutional neural network. Introduced basic part of the knowledge in and after the end of the level, we will discuss that in view of the education of CNN application and research, provide help for the education after class learning, mainly in the children's autonomous learning in the computer classified according to image feature selection so as to achieve the purpose of communicating with children, and even help children to learn image through image tagging - look at the picture and speak.In addition to the education of parents and teachers, the large amount of learning data provided by computers will provide more learning materials for children and help them to learn and classify. Organizing your thinking is easier to exercise.

1. Introduction

The motive of convolutional neural networks is shown in the figure above，The generation of convolutional neural network is based on three observations. First, the features of the image are generally more important than the whole image and consume less memory. Based on the features, the machine can judge the whole image, or add labels and classifications to the image. Second, the same feature element may appear in different images, such as a bird's beak, in different images but can be classified based on this feature alone. Third, subsampling will not affect or change the features in the image.

1.1 Comprehensive analysis of CNN model

LeNet-5 consists of 8 layers. Input layer: N 32x32 training samples. The input image size is 32x32, larger than the letters in MNIST database. C1 layer is a convolution layer, which is composed of 6 Feature maps. Each neuron in the feature graph is connected to the neighborhood with input 55. The size of the feature graph is 2828, which prevents input connections from falling out of bounds. The S2 layer is a sub-sampling layer with six 1414 feature graphs. Each element in the feature graph is connected to the 22 neighborhood of the corresponding feature graph in C1. C3 layer is also a convolution layer. It also delves the convolution layer S2 through the convolution kernel of 5x5, and then the characteristic map obtained is only 10x10

  DOI: 10.25236/sser.2019.134

neurons. S4 is a lower sampling layer, consisting of 16 5x5 size feature maps, each unit of which is connected to the 2x2 neighborhood of the corresponding feature map in C3. The S4 layer has 32 trainable parameters (1 factor and a bias for each feature map) and 2000 connections. C5 layer is a convolution layer with 120 features. The F6 has 84 units connected to the C5 layer and has 10,164 trainable parameters.

AlexNet Is the enlarged version of LeNet, the input is a 224x224 image, after 5 convolutional layers, 3 full connection layers (including a classification layer), to reach the final label space.Layer 1: these are filled blocks and boundary features.Intermediate layer: learn some texture features. Higher level: near the level of the classifier, the shape features of the object can be clearly seen. The last layer: classification layer, is completely different attitude of the object, according to different objects show different attitude characteristics. The learning process of objects is from the edges to the parts to the whole. ZFNet is much like AlexNet, except it's optimized. VGGNet has explored the relationship between the depth of convolutional neural network and its performance. By repeatedly stacking 33 small convolutional kernels and 22 maximum pooling layers, VGGNet has successfully constructed a convolutional neural network with a depth of 16-19 layers.[1-3]

1.2 Regional convolutional neural network

R - CNN is in order to achieve the purpose of identifying objects, that is to say, for any one image, we want to be able to use the R - CNN method to partition the edges of the object and background. Constitute two parts, and Fast R-CNN and Faster R-CNN is by improving the original model to make the process more rapidly, the former is analysis in different regions of the calculation of convolution layer, the latter is to overcome the complexity of training on the pipeline.[4,5] After being able to identify the object in the image, that is, after distinguishing the object from the background, identifying the precise position of the object is the next step.

2. Realization and application based on convolution neural network and education industry

2.1 Realization based on convolution neural network

NumPy was used to build CNN of convolutional neural network：Three modules were constructed, namely convolution layer, ReLu activation function and maximum pooling. After the image is read, it is converted to grayscale image, and then a filter is prepared for the convolution layer to perform convolution operation with the input image. The ReLU graph applies the activation function to the feature graph of each layer, and finally performs pooling.[6,7] Pooling is generally divided into a variety of pooling methods, in addition to the maximum pooling, there are average pooling, summation pooling, and so on, the refined view is as follows：

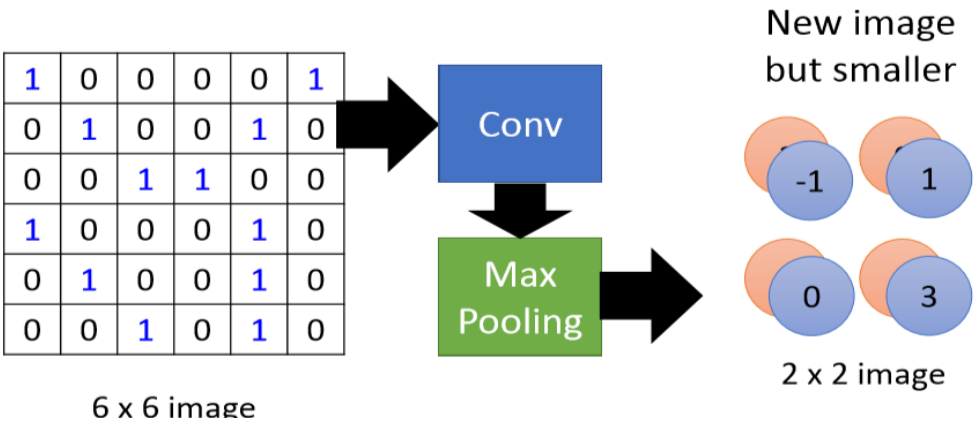

Figure 1 CNN—Convolution + Max Pooling

The second implementation method is based on Pytorch. After the convolution layer and pooling layer, the image should have been convolved and pooled several times. When entering the next activation layer, the image should first enter the batch standardization layer, which can help accelerate the convergence speed.Under the operation of the two implementation methods, the accuracy of above 0.97 can be achieved.[8]

2.2 Application of CNN in the field of education

CNN is mostly applied in computer vision and natural language processing. In addition, the achievements of machine learning can help children's education enlightenment.The application classification of CNN is mainly divided into three categories. The first category is image classification. In this category, the goal we want to achieve is to output the most likely tag after model processing when given an image, which can truly and accurately reflect the content contained in the image. The application classification of CNN is mainly divided into three categories. The first category is image classification. In this category, the goal we want to achieve is to output the most likely tag after model processing when given an image, which can truly and accurately reflect the content contained in the image.

The second type of application is image detection. The principle of the above image recognition is mainly the recognition of features in the image by convolution neural network. The set of features of some things can be used to judge the label of things. On this basis, it is more difficult to detect the image, that is, it is necessary to classify the image and circle the object with a rectangle. R-CNN is the framework generated on this demand, but it takes a long time to detect, about two seconds to detect a photo. However, the subsequent optimization models Fast R-CNN and Fster rR-CNN greatly reduced the time and improved the accuracy.[9] Finally, SSD with high accuracy and speed was developed, which met the requirements of most cases. The extended application is image segmentation, which is equivalent to the collection of multiple image detection. Its purpose is to separate various objects in the image with different colors. From FCN model to DeepLab framework to CRF as RNN of Oxford University, the accuracy of image segmentation is also improving. The educational application in this respect should be more helpful to art students and designers than to children. In terms of painting and design, the cognition of object position and three-dimensional state is more important and crucial. The educational application in this respect should be more helpful to art students and designers than to children. In terms of painting and design, the cognition of object position and three-dimensional state is more important and crucial. Especially in the field of image

segmentation, the primary and secondary relations of objects can be identified by separating multiple objects and finding the relation. After removing complex background, the position relation of objects can be clearly seen. Help for children and students will be mainly in the construction of thinking. It can effectively help students form a three-dimensional view in their thinking, such as in the three-dimensional geometry questions in the college entrance examination.[10]

The third kind of application is to look at the picture. As the name implies, once a picture appears, the goal of the machine should be to describe the specific content of the picture in words. The principle of implementation is mainly the relationship between the content of multiple tags. The possible relationship is calculated by the model and the predicted relationship is extracted and judged by features. The main application area is image search, which is related to education and can further explore the field of education for the blind. For blind people, access to information is through touch and hearing. However, the speed of the blind to obtain information is still very slow. The conversion of visual images into text and then into voice output will greatly shorten the speed of the blind to obtain information. In further development, devices that help blind people see the world could start with this principle. Another area of extension is text-to-image. Similarly, the principle that can be used in helping blind people is image to text, so that people with hearing problems can further explore this area. Through the received text input, relevant image information and action features are extracted from the database to build a complete picture. It can be seen from the above main application fields that different fields have different applications in the direction of education. According to different principles, there can be various application developments, and the application of convolutional neural network in education is highly targeted. The basic principle is the very basic convolution and pooling reaction mentioned above.

3. Conclusion

Convolutional neural network has some advantages that traditional technologies do not have. For example, it has a good fault-tolerant ability and can process information in a complex environment, and can also conduct relevant processing when the samples are not complete. And its generalization ability is obviously better than other methods. The convolutional neural network makes full use of the characteristic information contained in the data itself by combining the local sensing area and the Shared weight. This is a deep learning model. It is commonly used in computer vision and natural language recognition, but further applications have not been developed. For example, in the field of education. Due to its strong pertinence, it has application potential for different objects based on different principles, such as the help for the disabled, and the help for people learning to draw.

References

[1] Wu Yue, Hassner Tal, Kim Kanggeon, Medioni Gerard, Natarajan Prem (2017) Facial Landmark Detection with Tweaked Convolutional Neural Networks. IEEE transactions on pattern analysis and machine intelligence, 10.1109/TPAMI.2017.2787130

[2] Byra Michal, Galperin Michael, Ojeda-Fournier Haydee (2018) Breast mass classification in sonography with transfer learning using a deep convolutional neural network and color conversion. Medical physics, 10.1002/mp.13361

[3] Lee Min Beom, Hong Hyung Gil; Park Kang Ryoung (2017) Noisy Ocular Recognition Based on Three Convolutional. Neural Networks. Sensors (Basel, Switzerland), 10.3390/s17122933

[4] Kunrong Zhao, Tingting He, Shuang Wu, Songling Wang, Bilan Dai, Qifan Yang, Yutao Lei (2018) Application research of image recognition technology based on CNN in image location of environmental monitoring UAV. Springer Journal, 10.1186/s13640-018-0391-6

[5] Sengur Abdulkadir, Akbulut Yaman, Guo Yanhui, Bajaj Varun (2017) Classification of amyotrophic lateral sclerosis disease based on convolutional neural network and reinforcement sample learning algorithm. Health information science and systems, 10.1007/s13755-017-0029-6

[6] Kim D H, MacKinnon T (2017) Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. Clinical radiology, 10.1016/j.crad.2017.11.015

[7] Jinwei Wan, Bo Chen, Bin Xu, Hongwei Liu, Lin Jin (2019) Convolutional neural networks for radar HRRP target recognition and rejection. EURASIP Journal on Advances in Signal Processing, 10.1186/s13634-019-0603-y

[8] Mingyuan Xin, Yong Wang (2019) Research on image classification model based on deep convolution neural network. EURASIP Journal on Image and Video Processing, Vol.2019 (1), pp.1-11

[9] Kim Tackeun, Heo Jaehyuk, Jang Dong-Kyu (2018) Machine learning for detecting moyamoya disease in plain skull radiography using a convolutional neural network. EBioMedicine, 10.1016/j.ebiom.2018.12.043

[10] Phan Anh Viet, Nguyen Minh Le, Nguyen Yen Lam Hoang, Bui Lam Thu (2018) DGCNN: A convolutional neural network over large-scale labeled graphs. Neural networks:the official journal of the International Neural Network Society, 10.1016/j.neunet.2018.09.001